

UNIVERSIDADE ESTADUAL DE CAMPINAS (UNICAMP)  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E CIÊNCIA DA  
COMPUTAÇÃO (IMECC)

*EM*  
ANÁLISE DISCRIMINANTE EM MEDIDAS REPETIDAS

*EM*  
EMMA NORMA CAMBILLO MOYANO

ORIENTADORA : PROFA. DRA. REGINA C. C. P. MORAN

CAMPINAS - SÃO PAULO - BRASIL

Março de 1990

## ANÁLISE DISCRIMINANTE EM MEDIDAS REPETIDAS

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pela Sra. EMMA NORMA CAMBILLO MOYANO e aprovada pela Comissão Julgadora.

Campinas , 19 de Março de 1990 .

Profa. Dra.



REGINA C. C. P. MORAN

Orientadora

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da Computação, UNICAMP , como requisito parcial para obtenção do Título de Mestre em Estatística.

A Tomás e Lília ,

Com Amor.

## AGRADECIMENTOS

Quero agradecer a todos que contribuíram para a realização do presente trabalho.

Em particular,

A Professora Regina C. C. P. Moran ; orientadora pela constante orientação , apoio e segurança que me transmitiu para vencer as dificuldades do trabalho .

Aos Professores do Departamento de Estatística do IMECC , pelo ensinamentos recebidos.

A Todos os colegas e amigos de la Universidad Nacional de San Marcos -UNMSM e do Ministerio de Agricultura - INAF pela amizade e incentivo constantes.

A Todos os colegas e amigos de Campinas que de uma forma ou de outra contribuíram para a realização deste trabalho.

A Marina , Olga pela amizade .

## SUMÁRIO

A análise discriminante trata dos problemas de discriminação e classificação. Na discriminação está voltada a descobrir e testar diferenças entre grupos, na classificação leva a regras bem definidas utilizadas para classificar novos indivíduos.

Medidas repetidas refere-se a dados constituídos de valores de uma mesma variável aleatória, observada sobre o mesmo indivíduo em ocasiões distintas. Os modelos de médias e polinomiais são utilizados no estudo do comportamento dessa classe específica de vetor multivariado.

Tanto a análise de discriminante quanto a de medidas repetidas tem sido estudadas por muitos anos. Esta tese está concentrada no estudo das regras de classificação de máxima verossimilhança, sob hipótese de multinormalidade e mesma matriz de variância e covariância dentro dos grupos, no contexto de medidas repetidas. Assim, nos dados originais, através de simples modelos de média as regras de classificação são derivadas para o caso de  $g$  grupos e investigada sua especificidade para duas estruturas da matriz de variância e covariância, a uniforme e a seriada. Em cada uma delas, e também investigada a especificidade da regra de classificação para o caso particular de paralelismo, estruturação adicional nos primeiros momentos.

Ainda, o mesmo estudo é feito em transformações das variáveis originais, através de ajustes de modelos saturados que

produzem um igual número de coeficientes aqui chamados variáveis alternativas , às quais, por sua vez , correspondem a estruturas da matriz de variância e covariância induzida pela estruturação nas variáveis originais .

Os resultados obtidos, exibem um grau de especificidade, em cada caso considerado , cuja análise evolui para entendimentos mais minuciosos dos aspectos da análise discriminante neste contexto .Para o caso de dois grupos , a função discriminante foi também particularizada e estes resultados são utilizados para efeito de interpretação , na sua forma analítica e via exemplo.

Para um modelo de médias fixado, na estrutura uniforme ficou evidente que um mesmo valor do coeficiente de correlação pode corresponder à situações totalmente distintas quanto à dispersão dos perfis, e portanto este valor deve ser examinado em função das variabilidades entre e dentro . Mantida a relatividade das magnitudes das fontes de variabilidade entre e dentro , duas são as situações que favorecem a discriminação : valores grandes de variabilidade dentro quando conjugados com valores baixos do coeficiente de correlação ou valores pequenos de variabilidade dentro quando conjugados com valores altos do coeficiente de correlação. No caso seriado , fixada a escala de variabilidade , valores baixos do coeficiente de correlação facilitam a discriminação.

## INDICE

Cap. 1	INTRODUÇÃO .....	01
Cap. 2	DISCRIMINAÇÃO NAS VARIÁVEIS ORIGINAIS.....	06
2.1	No modelo de médias.....	06
2.2	Para estruturas uniforme e seriada.....	11
Cap. 3	DISCRIMINAÇÃO NAS VARIÁVEIS ALTERNATIVAS.....	21
3.1	No modelo polinomial saturado.....	22
3.2	Para estruturas uniforme e seriada.....	29
Cap. 4	ALGUMA INTERPRETAÇÃO PARA DOIS GRUPOS.....	42
4.1	Interpretação nos dados originais.....	43
i)	$\Sigma$ Uniforme .....	43
ii)	$\Sigma$ Seriada.....	47
4.2	Interpretação nos coeficientes polinomiais	49
i)	$\Sigma$ Uniforme.....	49
ii)	$\Sigma$ Seriada.....	51
4.3	O caso de dois grupos via exemplo.....	53
4.4	Considerações em casos reais.....	57
	APÊNDICES.....	61
	A Surgimento das diferentes estruturas da matrizes	
	de variância e covariância.....	62
A.1	$\Sigma$ uniforme .....	62
A.2	$\Sigma$ seriada.....	64
A.3	$\Sigma$ seriada de Kenward.....	68
	B Cálculo das inversas das matrizes de variância e	

covariância dos dados originais .....	71
B.1 $\Sigma$ uniforme.....	71
B.2 $\Sigma$ seriada .....	72
B.3 $\Sigma$ seriada de Kenward .....	73
C Matrizes de variância e covariância dos coeficientes ortonormais e suas inversas.....	75
C.1 $\Sigma$ Uniforme.....	75
C.2 $\Sigma$ seriada.....	76
C.3 $\Sigma$ seriada de Kenward.....	80
D Programas.....	84
D.1 Geração dos perfis individuais.....	84
D.2 Gráfico dos perfis individuais.....	88
REFERÊNCIAS BIBLIOGRÁFICAS.....	90



# CAPITULO 1

## INTRODUÇÃO

O propósito deste capítulo é colocar o problema em estudo no contexto das áreas envolvidas.

Na análise multivariada figura como um tópico importante a análise discriminante, técnica voltada à examinar o quanto é possível distinguir entre membros de vários grupos, com base em observações feitas sobre eles. A análise provê testes de diferenças, regras de alocação e estimativas de probabilidade de alocação correta .

Análise discriminante ou discriminação e classificação é portanto uma extensão da análise de variância ordinária à observações multivariadas(MANOVA) apresentando os problemas próprios acima mencionados. Na fase de classificação leva a regras bem definidas, as quais são utilizadas para classificar novos indivíduos. Na primeira fase descobre e testa diferenças entre grupos , os testes envolvidos são idênticos aos de MANOVA . Este trabalho focaliza a questão da classificação explorando as regras de máxima verossimilhança . Em casos de medidas repetidas ( sob



hipótese de multinormalidade e matriz de variância e covariância comum) sob o aspecto da análise discriminante referente à discriminação. Quando os dados tem estrutura de medidas repetidas as regras de classificação podem revelar onde está o poder de separação dos grupos e, particularmente, no caso de dois grupos a função discriminante será estudada.

Estrutura de medidas repetidas refere-se à dados constituídos de valores de uma mesma variável aleatória, observada sobre o mesmo indivíduo em ocasiões distintas. Estes indivíduos provêm de  $g$  grupos ou foram alocados à  $g$  grupos, conforme o carácter observacional ou experimental do problema. Nos dois casos, os objetivos e interpretações se diferenciam. Para efeito de desenvolvimento da metodologia não serão distinguidos.

Esta estrutura de dados aparece no contexto de experimentos de medidas repetidas e estudos longitudinais, o primeiro nome em geral reservado para espaçamentos pequenos entre ocasiões, conforme sugerido (1985).

No caso geral de vetores multivariados as inter-relações entre variáveis componentes são tratadas através da correspondente matriz de variância e covariância. No caso específico de dados provenientes de medidas repetidas, temos uma classe também específica de vetor multivariado por terem as variáveis componentes mesmas propriedades métricas (mesma escala e origem), o quê por sua vez pode sugerir estruturar a matriz de

variância e covariância. As estruturas uniforme e seriada são exploradas.

A análise deste tipo de dados passa por vários enfoques , os modelos de média e polinomiais serão abordados para dentre deles explorar as regras de discriminação.

Uma potencial aplicação dos resultados quanto à classificação de um novo indivíduo , pode ser divisada em estudos prospectivos nos quais os indivíduos são agrupados pela ocorrência ou não de determinado evento ao término do período de observação.

Por exemplo na prevenção da toxemia gravídica através da pre-eclampsia, gestantes acompanhados mês a mês quanto à evolução da pressão arterial agrupadas ao final da gestação nos grupos com pre-eclampsia e sem pre-eclampsia. Uma nova paciente pode passar então pelo procedimento de classificação , mês a mês , segundo as regras derivadas para cada dos conjuntos de ocasiões (as derivadas para os dois primeiros meses, tres primeiros meses, até os nove meses completos ).

A utilização dos resultados , pode ser de interesse, ainda que não haja sentido para classificação de novos indivíduos. Isto porque a técnica de discriminação , como desenvolvida neste trabalho , ilumina aspectos das diferenças dos grupos : relativos à variabilidade e às diferenças nos perfis médios.

O trabalho será estruturado em quatro capítulos.

No capítulo dois o modelo de médias usual é considerado e as regras de classificação derivadas e particularizadas para matrizes de variância e covariância uniforme e seriada, sendo tratado mais especificamente o caso de dois grupos e dentro deste a ocorrência de paralelismo.

O terceiro capítulo segue a mesma linha de desenvolvimento do capítulo dois sobre as variáveis originais transformadas, via modelos polinomiais saturados, ou seja um igual número de variáveis alternativas. Estas últimas despertam maior interesse no contexto de medidas repetidas.

No quarto capítulo os resultados para o caso de dois grupos, para as duas estruturas de matriz de variância e covariância exploradas, são reapresentados sob enfoque interpretativo. Um exemplo de resultados simulados é utilizado para ilustrar os aspectos considerados. Conclusões indicativas são apresentadas e podem servir para orientar expectativas quanto ao poder de separação dos grupos.

Em termos gerais o estudo da especificidade das regras de discriminação no caso de medidas repetidas, para as matrizes estruturadas aqui consideradas, e no caso mais simples de dois grupos, objetiva combinar esta técnica ao esforço de modelagem usual. De fato neste contexto além de ajustar curvas e detectar diferenças, os padrões de mudança ao longo do tempo estão sendo caracterizados. No ajuste polinomial as constantes refletem as

médias gerais ao longo das ocasiões , o coeficiente de primeiro grau a taxa de mudança das respostas ,o de segundo grau a alteração desta tendência e assim por diante. Na caracterização da regra discriminante , a função ganha significado relativo à esses padrões de mudança revelando o papel destes componentes na separação entre os grupos. Dentro da análise discriminante , o tratamento das variáveis transformadas quando tratando de medidas repetidas , passa então a focalizar menos as observações nos diferentes tempos e mais o padrão de mudança ao longo do tempo.

A questão de estimação não foi explorada , e vale ainda salientar que o trabalho trata modelos saturados. No futuro , os resultados aqui derivados serão explorados do ponto de vista da estimação , e particularmente do ajuste polinomial com covariáveis e portanto com polinômios de grau menor e / ou distintos.

## CAPÍTULO 2

### DISCRIMINAÇÃO NAS VARIÁVEIS ORIGINAIS

Neste capítulo a função discriminante é estudada em vetores multinormais aqui chamados discriminadores, correspondentes à observação de uma mesma variável sobre o mesmo indivíduo ao longo de um certo número de ocasiões, isto é, no contexto de medidas repetidas.

Aqui, a modelagem é considerada nas variáveis originais e a especificidade da função discriminante corre por conta do modelo na seção 2.1. Na seção 2.2 os resultados para o modelo de médias são particularizados para estrutura uniforme da matriz de variância e covariância e também para estrutura seriada, sendo considerados o paralelismo em cada caso.

#### 2.1 NO MODELO DE MÉDIAS

Na análise multivariada de perfis no caso de  $g$  grupos, Singer e Andrade (1986), quando as médias são descompostas segundo modelo de médias de Bock (1975), temos,

$$E(X) = A E \quad (2.1.1)$$

onde,

$$X_{n \times p} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(g)} \end{bmatrix}$$

e,

$$X_{n_j \times p}^{(j)} = \begin{bmatrix} x_{j1} \\ x_{j2} \\ \vdots \\ x_{jn_j} \end{bmatrix}$$

$$x_{j\alpha} = (x_{j\alpha 1} \ x_{j\alpha 2} \ \dots \ x_{j\alpha p})'$$

e,

$$A_j = \begin{bmatrix} 1_{n_1} & 0 & \dots & \dots & 0 \\ 0 & 1_{n_2} & \dots & \dots & 0 \\ \vdots & \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & \dots & 1_{n_g} \end{bmatrix}$$

onde,

$$1_{n_j} = [1 \ 1 \ \dots \ 1]'$$

e o vetor de dimensão  $n_j$ ;  $j = 1, 2, \dots, g$ .

O vetor p-variado  $x_{j\alpha}$  correspondente à observação dos discriminadores sobre o  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo, é modelado por,

$$x_{j\alpha} = \tau + \theta_j + \varepsilon_{j\alpha} \quad (2.1.2)$$

$$\alpha = 1, \dots, n_j, j \in G = \{1, 2, \dots, g\}$$

com,

$\tau$  é o vetor de médias ao longo das ocasiões

$\theta_j$  é o vetor do efeito do  $j$ -ésimo grupo do qual provem o indivíduo.

$\varepsilon_{j\alpha}$  é o vetor dos erros, e por hipótese

$$\varepsilon_{j\alpha} \sim N_p(0, \Sigma)$$

De (2.1.1.) e (2.1.2) ,



$$[u] = \begin{bmatrix} \mu'_1 \\ \mu'_2 \\ \vdots \\ \mu'_g \end{bmatrix} = \begin{bmatrix} (\tau + \theta)' \\ (\tilde{\tau} + \tilde{\theta}_1)' \\ \vdots \\ (\tau + \theta)' \\ \tilde{\tau} \quad \tilde{\theta}_g \end{bmatrix} \quad (2.1.3)$$

onde,

$\mu'_j$  é o vetor de médias do j-ésimo grupo.

Neste caso a regra de classificação de máxima verossimilhança, Mardia et al. (1979), aloca  $x$  (uma nova observação) ao grupo  $j$ , onde  $j \in G$  é o valor de  $i$ , que minimiza a distância de Mahalanobis:

$$\min_{i \in G} (x - \mu'_i)' \Sigma^{-1} (x - \mu'_i) = (x - \mu'_j)' \Sigma^{-1} (x - \mu'_j) \quad (2.1.4)$$

No decorrer do trabalho o grupo ao qual corresponde a mínima distância é indicado por  $j$ .

No caso considerado, por (2.1.3), o cálculo do mínimo para o caso de  $g$  grupos se reduz à,

$$\begin{aligned} \min_{i \in G} (x - (\tau + \theta))' \Sigma^{-1} (x - (\tau + \theta)) = \\ = (x - (\tau + \theta))' \Sigma^{-1} (x - (\tau + \theta)) \end{aligned}$$

Particularmente no caso de dois grupos a regra de classificação é dada por , aloca-se  $\tilde{x}$  ao grupo 1 se,

$$(\mu_{\tilde{1}} - \mu_{\tilde{2}})' \Sigma^{-1} (\tilde{x} - \mu_{\tilde{2}}) > 0 \quad (2.1.5)$$

e aloca-se  $\tilde{x}$  ao grupo 2 em outro caso; onde  $\mu_{\tilde{}}$  é a media geral.

Sob o modelo (2.1.2) esta fica reduzida à : aloca-se  $\tilde{x}$  ao grupo 1 se,

$$(\theta_{\tilde{1}} - \theta_{\tilde{2}})' \Sigma^{-1} (\tilde{x} - \mu_{\tilde{2}}) > 0 \quad (2.1.6)$$

e aloca-se  $\tilde{x}$  ao grupo 2 em outro caso , pois por (2.1.3)

$$\mu_{\tilde{}} = \tau_{\tilde{}} + \frac{\theta_{\tilde{1}} + \theta_{\tilde{2}}}{2}$$

•

$$\mu_{\tilde{1}} - \mu_{\tilde{2}} = \theta_{\tilde{1}} - \theta_{\tilde{2}}$$

denotando a diferença entre os efeitos de grupo por  $d_{\tilde{}}$ ,

$$d_{\tilde{}} = (d_1, d_2, \dots, d_p)'$$

onde  $d_k$  é a diferença entre os efeitos de grupo na  $k$ -ésima ocasião, (2.1.6) pode ser expressa como, aloca-se  $\tilde{x}$  ao grupo 1 se,

$$\tilde{d}' \Sigma^{-1} (\tilde{x} - \tilde{\mu}) > 0 \quad (2.1.7)$$

e aloca-se  $\tilde{x}$  ao grupo 2 em outro caso, dependendo por um lado da magnitude da diferença dos efeitos de grupo e por outro dos desvios da observação com relação à média geral dos grupos através da matriz de variância e covariância.

## 2.2 PARA ESTRUTURAS UNIFORME E SERIADA

Devido a existência de algumas estruturas frequentes para a matriz de variância e covariância, neste contexto entre elas a uniforme, e.g., Geisser (1963) e a de correlação serial, e.g., Cole e Grizzle (1966), a função discriminante será estudada nestes dois casos.

No modelo de médias de Bock (1975), com matriz de variância e covariância uniforme, que se pode escrever como,

$$\Sigma = \sigma^2 I + \sigma_\alpha^2 1 1'$$

onde  $\sigma_\alpha^2$  é a variabilidade dentro do indivíduo ao longo das ocasiões e  $\sigma^2$  é a variabilidade entre indivíduos dentro de cada ocasião. Detalhes do surgimento da matriz estão no apêndice A.1.

A inversa é dada por :

$$\Sigma^{-1} = \frac{1}{\sigma^2} \left( \mathbb{I} - \frac{\sigma_\alpha^2}{\sigma^2 + p\sigma_\alpha^2} \mathbf{1}\mathbf{1}' \right)$$

O cálculo da inversa encontra-se no apêndice B.1.

O cálculo do mínimo para o caso de  $g$  grupos (2.1.4) é,

$$\text{Mín}_{i \in G} \frac{1}{\sigma^2} \left( \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right)' \left( \mathbb{I} - \frac{\sigma_\alpha^2}{\sigma^2 + p\sigma_\alpha^2} \mathbf{1}\mathbf{1}' \right) \left( \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right) =$$

$$\frac{1}{\sigma^2} \left\{ \sum_{k=1}^p \left( x_k - \mu_{jk} \right)^2 - \frac{\sigma_\alpha^2}{\sigma^2 + p\sigma_\alpha^2} \left[ \sum_{k=1}^p \left( x_k - \mu_{jk} \right) \right]^2 \right\} \quad (2.2.1)$$

onde,

$x_k$  é a  $k$ -ésima observação do novo indivíduo

$\mu_{jk}$  é a média do  $j$ -ésimo grupo na  $k$ -ésima ocasião.

No caso específico de dois grupos a regra de classificação (2.1.7) se reduz a aloca-se  $x$  ao grupo 1 se

$$\frac{1}{\sigma^2} \left\{ \sum_{k=1}^p \left[ d_k - \frac{p \sigma_\alpha^2}{\sigma^2 + p \sigma_\alpha^2} \bar{d} \right] (x_k - \mu_k) \right\} > 0 \quad (2.2.2)$$

e aloca-se  $x$  ao grupo 2 em outro caso, onde  $\mu_k$  é a média geral na  $k$ -ésima ocasião

Se os perfis são paralelos isto é quando as diferenças dos efeitos dos grupos permanecem constantes de ocasião a ocasião ou seja ,

$$d_i = d = \bar{d}$$

onde  $\bar{d}$  é a média da diferença entre os efeitos dos grupos.

A regra de classificação é dada por, aloca-se  $x$  ao grupo 1 se,

$$\frac{d}{\sigma^2 + p \sigma_\alpha^2} \sum_{k=1}^p (x_k - \mu_k) > 0 \quad (2.2.3)$$

em caso contrário ao grupo 2.

Outra estrutura que a matriz de variância e covariância pode assumir é de correlação serial ,Singer e Andrade (1986) .

$$\Sigma = \frac{\sigma_u^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \dots & \rho^{p-2} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & \dots & 1 \end{bmatrix}, \quad |\rho| < 1, \sigma_u^2 > 0$$

onde  $\rho$  é a correlação entre duas ocasiões consecutivas.  $\sigma_u^2$  é variabilidade entre indivíduos corrigida por fatores que envolvem a estrutura de correlação dentro do indivíduo, o estimador de máxima verossimilhança é obtido por Hearne et al.(1983). Detalhes do surgimento da matriz são dados no apêndice A.2.

A inversa é dada por :

$$\Sigma^{-1} = \frac{1}{\sigma_u^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \dots & \dots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & -\rho & 1 \end{bmatrix}$$

e seu cálculo dado no apêndice B.2.

O cálculo do mínimo para o caso de  $g$  grupos, (2.1.4)

é dada por,

$$\text{Mín}_{i \in G} \begin{bmatrix} x \\ \sim \end{bmatrix} - \begin{bmatrix} \mu_i \\ \sim \end{bmatrix} \Sigma^{-1} \begin{bmatrix} x \\ \sim \end{bmatrix} - \begin{bmatrix} \mu_i \\ \sim \end{bmatrix} =$$

$$\begin{aligned}
&= \frac{1}{\sigma_U^2} \left( \sum_{k=1}^p (x_k - \mu_{jk})^2 + \rho^2 \sum_{k=2}^{p-1} (x_k - \mu_{jk})^2 \right. \\
&\quad \left. - 2\rho \sum_{k=1}^{p-1} (x_{k+1} - \mu_{j(k+1)}) (x_k - \mu_{jk}) \right) \quad (2.2.4)
\end{aligned}$$

a qual se reduz à

$$\text{Min}_{x \in G} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_j \\ \sim \end{matrix} \right]' \Sigma^{-1} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_j \\ \sim \end{matrix} \right] = \frac{1}{\sigma_U^2} \sum_{k=1}^p (x_k - \mu_{jk})^2$$

se a correlação entre ocasiões é zero, um caso particular do modelo uniforme quando não há variação dentro do indivíduo.

Particularmente no caso de 2 grupos a regra de classificação (2.1.7) é dada por: aloca-se  $x$  ao primeiro grupo se

$$\begin{aligned}
&\frac{1}{\sigma_U^2} \left( \sum_{k=1}^p d_k (x_k - \mu_j) - \rho \sum_{k=2}^{p-1} (d_{k-1} - \rho d_k + d_{k+1}) (x_k - \mu_k) \right. \\
&\quad \left. - \rho d_2 (x_1 - \mu_1) - \rho d_{p-1} (x_p - \mu_p) \right) > 0 \quad (2.2.5)
\end{aligned}$$

em caso contrário aloca-se ao grupo 2.

Se os perfis são paralelos,  $d_i = d$ , fica expressa por,

$$\frac{d}{\sigma_u^2} \left[ (1 - \rho) \sum_{k=1}^p (x_k - \mu_k) - \rho (1 - \rho) \sum_{k=2}^{p-1} (x_k - \mu_k) \right] > 0 \quad (2.2.6)$$

e se além do paralelismo,  $\rho = 0$ , então fica reduzida a

$$\frac{d}{\sigma_u^2} \sum_{k=1}^p (x_k - \mu_k) > 0$$

a mesma regra obtida no caso uniforme sob paralelismo quando não há variação dentro dos indivíduos.

Como uma forma de introduzir o padrão de correlação seriado na estrutura dentro do indivíduo, Kenward (1985) apresenta outra estrutura para a matriz de variância e covariância. detalhes do surgimento da matriz, são dados no apêndice A.3.



$$\Sigma = \frac{\sigma_{\alpha}^2}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \dots & 1 \end{bmatrix} + \sigma^2 \mathbf{1} \mathbf{1}', \quad \begin{array}{l} |\rho| < 1 \\ \sigma^2 > 0 \\ \sigma_{\alpha}^2 > 0 \end{array}$$

Para efeito de comparação com o caso uniforme, a notação das fontes de variação entre indivíduos e dentro do indivíduo foi mantida, porém  $\sigma_{\alpha}^2$  e  $\sigma^2$  devem ser estimados dentro deste modelo e só coincidem com o caso uniforme quando  $\rho = 0$  e os papéis de  $\sigma^2$  e  $\sigma_{\alpha}^2$  são trocados. A inversa é dada por

$$\Sigma^{-1} = \frac{1}{\sigma_{\alpha}^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 + \rho^2 & -\rho & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix} - w \mathbf{A}$$

Os detalhes da manipulação algébrica encontram-se no apêndice B.3

$$\mathbf{A} = \begin{bmatrix} (1-\rho)^2 & (1-\rho)^3 & (1-\rho)^3 & \dots & (1-\rho)^3 & (1-\rho)^2 \\ (1-\rho)^3 & (1-\rho)^4 & (1-\rho)^4 & \dots & (1-\rho)^4 & (1-\rho)^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (1-\rho)^2 & (1-\rho)^3 & (1-\rho)^3 & \dots & (1-\rho)^3 & (1-\rho) \end{bmatrix}$$

$$w = \frac{\sigma^2}{\sigma_\alpha^2 (\sigma_\alpha^2 + v \sigma^2)}$$

$$v = p(1-\rho)^2 + 2\rho(1-\rho)$$

O cálculo do mínimo para o caso de  $g$  grupos, (2.1.4) é dada por :

$$\text{Min}_{i \in G} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right]' \Sigma^{-1} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right] =$$

$$= \frac{1}{\sigma_\alpha^2} \left[ \sum_{k=1}^p \left( x_k - \mu_{jk} \right)^2 + \rho^2 \sum_{k=2}^{p-1} \left( x_k - \mu_{jk} \right)^2 - 2\rho \sum_{k=1}^{p-1} \left( x_{k+1} - \mu_{j(k+1)} \right) \right]$$

$$\left[ x_k - \mu_{jk} \right]^2 - w(1-\rho)^2 \left[ (x_1 - \mu_{j1}) + (x_p - \mu_{jp}) \right] + (1-\rho) \sum_{k=2}^{p-1} \left[ x_k - \mu_{jk} \right]^2 \quad (2.2.7)$$

a qual se reduz à

$$\text{Min}_{i \in G} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right]' \Sigma^{-1} \left[ \begin{matrix} x \\ \sim \end{matrix} - \begin{matrix} \mu_i \\ \sim \end{matrix} \right] = \frac{1}{\sigma_\alpha^2} \sum_{k=1}^p \left( x_k - \mu_{jk} \right)^2 -$$

$$- w \left[ (x_1 - \mu_{j1}) + (x_p - \mu_{jp}) + \sum_{k=2}^{p-1} (x_k - \mu_{jk}) \right]^2$$

se a correlação entre ocasiões é zero.

Particularmente no caso de 2 grupos a regra de classificação (2.1.7) é dada por, aloca-se  $x$  ao primeiro grupo se,

$$\begin{aligned} & \frac{1}{\sigma_\alpha^2} \left[ \sum_{k=1}^p d_k (x_k - \mu_{jk}) - \rho \sum_{k=2}^{p-1} (d_{k-1} - \rho d_k + d_{k+1}) (x_k - \mu_{jk}) \right. \\ & \left. - \rho d_2 (x_1 - \mu_1) - \rho d_{p-1} (x_p - \mu_p) \right] - \\ & w (1 - \rho^2) \left[ 1 - \rho \left[ \sum_{k=2}^{p-1} (x_k - \mu_k) + \sum_{k=2}^{p-1} d_k \right] \right. \\ & \left. + \rho^2 \sum_{k=2}^{p-1} d_k \sum_{k=2}^{p-1} (x_k - \mu_k) \right] > 0 \end{aligned} \quad (2.2.8)$$

em caso contrário aloca-se ao grupo 2.

Se os perfis são paralelos,  $d_k = d$ , fica expressa por,

$$\frac{d}{\sigma_\alpha^2} \left[ (1 - \rho) \sum_{k=1}^p (x_k - \mu_k) - \rho (1 - \rho) \sum_{k=2}^{p-1} (x_k - \mu_k) \right] -$$

$$w(1 - \rho)^2 \left[ \left( 1 - \rho \sum_{k=2}^{p-1} (x_k - \mu_k) \right) \left( 1 - (p-2)\rho d \right) \right] > 0$$

e se além do paralelismo,  $\rho = 0$  então fica reduzida a .

$$\frac{d}{\sigma_\alpha^2} \left[ \sum_{k=1}^p (x_k - \mu_k) - \frac{\sigma^2}{\sigma_\alpha^2 + \rho \sigma^2} \right] > 0$$

Como esperado, a modelagem de Kenward (1985) leva à funções discriminantes que se comparadas ao caso seriado anterior (2.2.4) e (2.2.5), ganham uma parcela que através do fator multiplicativo  $w$  envolve a variação "entre" e "dentro", de forma análoga ao caso uniforme. Quando  $\rho = 0$ , como as fontes de variação tem papéis trocados nos fatores multiplicativos em relação ao caso uniforme, parece indicar que o modelo de Kenward tem pouco respaldo em fenômenos reais quando  $\rho = 0$ , pois deixa as inter-relações por conta dos indivíduos.

## CAPÍTULO 3

### DISCRIMINAÇÃO NAS VARIÁVEIS ALTERNATIVAS

Neste capítulo a função discriminante é estudada depois de feito o ajuste de curvas polinomiais, isto é via transformações das variáveis originais em modelos saturados, ou seja o grau do polinômio é o máximo. É produzido portanto um igual número de variáveis: os coeficientes polinomiais, aqui chamados de variáveis alternativas. É amplamente conhecido no contexto das medidas repetidas, por questões de caracterizar padrão de mudança ao longo do tempo e possivelmente reduzir dimensionalidade enquanto aumenta interpretabilidade, o uso dos coeficientes polinomiais, Wishart (1938), Potthoff e Roy (1964), Rao (1965, 1966, 1967), Grizzle e Allen (1969), etc. são exemplos de contribuições importantes nesta linha.

Aqui serão abordados apenas modelos saturados, por questão de simplicidade.

Na seção 3.1 o ajuste via polinômios é apresentado e a regra de classificação nas variáveis alternativas derivada. Na seção 3.2 as regras são particularizadas para as estruturas uniforme e seriada da matriz de variância e covariância (sempre a estruturação relativa aos dados originais). Em cada caso é

tratado o paralelismo.

É interessante notar que as estruturas particulares das matrizes de variância e covariância dos coeficientes polinomiais, correspondentes aos casos uniforme e seriado, tem implicações que serão investigadas quando as estimativas dos coeficientes foram obtidas pelo ajuste via covariáveis conforme proposto por Rao (1965) e estudado em Kenward (1985,1986).

### 3.1 NO MODELO POLINOMIAL SATURADO

Na análise de curvas de crescimento, no caso de  $g$  grupos, Potthoff e Roy (1964) introduzem o seguinte modelo.

$$E(X) = A E T' \quad (3.1.1)$$

com  $X$  e  $A$  denotadas no capítulo 2 tomando  $E_{g \times p}$  como a matriz dos coeficientes associados aos polinômios naturais.

$$E = \begin{bmatrix} \xi_1^1 \\ \xi_1^2 \\ \vdots \\ \xi_1^g \\ \xi_2^1 \\ \xi_2^2 \\ \vdots \\ \xi_2^g \\ \vdots \\ \xi_g^1 \\ \xi_g^2 \\ \vdots \\ \xi_g^g \end{bmatrix} = \begin{bmatrix} \xi_{10} & \xi_{11} & \dots & \xi_{1(p-1)} \\ \xi_{20} & \xi_{21} & \dots & \xi_{2(p-1)} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \xi_{g0} & \xi_{g1} & \dots & \xi_{g(p-1)} \end{bmatrix}$$

e  $T$  a matriz dos tempos

$$T = \begin{bmatrix} t_1^0 & t_1^1 & \dots & t_1^{p-1} \\ t_2^0 & t_2^1 & \dots & t_2^{p-1} \\ \dots & \dots & \dots & \dots \\ t_p^0 & t_p^1 & \dots & t_p^{p-1} \end{bmatrix}$$

$x_{\sim j\alpha}$  o vetor p-variado correspondente à observação dos discriminadores sobre o  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo ou seja,

$$x_{\sim j\alpha} = (x_{j\alpha 1}, x_{j\alpha 2}, \dots, x_{j\alpha p})' \quad \alpha = 1, \dots, n_j$$

$$j \in G$$

onde de (3.1.1.) temos,

$$x_{\sim j\alpha} = T \xi_{\sim j} + \varepsilon_{\sim j\alpha} \quad (3.1.2)$$

e

$$x_{j\alpha k} = \xi_{j0} t_k^0 + \xi_{j1} t_k^1 + \dots + \xi_{j(p-1)} t_k^{p-1} + \varepsilon_{jk\alpha}$$

$\varepsilon_{\sim j\alpha} (px1)$  é o vetor p-variado dos erros e por hipótese,

$$\varepsilon_{\sim j\alpha} \sim N_p \left( 0, \Sigma \right) \quad (3.1.3)$$

então de (3.1.2) e (3.1.3),

$$x_{\sim j\alpha} \sim N_p \left( \bar{\Pi} \xi_{\sim j}, \Sigma \right) \quad (3.1.4)$$

Considere a transformação,

$$x_{\sim j\alpha}^* = (\bar{\Pi}, \bar{\Pi})^{-1} \bar{\Pi}' x_{\sim j\alpha}$$

e obvio que,

$$x_{\sim j\alpha}^* \sim N_p \left( \xi_{\sim j}^*, \Sigma^* \right) \quad (3.1.5)$$

onde

$$\Sigma^* = (\bar{\Pi}, \bar{\Pi})^{-1} \bar{\Pi}' \Sigma \bar{\Pi} (\bar{\Pi}, \bar{\Pi})^{-1}$$

E neste caso a regra de classificação para os  $g$  grupos, (2.1.4) fica dada por : aloca-se  $\xi$  os novos coeficientes (obtidos pelo ajuste saturado ao novo individuo) ao grupo  $j$  ; onde  $j \in G$  é o valor de  $i$  que minimiza a distância de Mahalanobis

$$\text{Mín}_{i \in G} \left[ \begin{matrix} \xi \\ \xi \end{matrix} - \begin{matrix} \xi_{\sim i} \\ \xi_{\sim i} \end{matrix} \right]' \Sigma^{*-1} \left[ \begin{matrix} \xi \\ \xi \end{matrix} - \begin{matrix} \xi_{\sim i} \\ \xi_{\sim i} \end{matrix} \right] = \left[ \begin{matrix} \xi \\ \xi \end{matrix} - \begin{matrix} \xi_{\sim j} \\ \xi_{\sim j} \end{matrix} \right]' \Sigma^{*-1} \left[ \begin{matrix} \xi \\ \xi \end{matrix} - \begin{matrix} \xi_{\sim j} \\ \xi_{\sim j} \end{matrix} \right] \quad (3.1.6)$$

Particularmente no caso de dois grupos a regra de classificação (2.1.5) é dada por, aloca-se  $\xi$  ao grupo 1 se,



$$\left( \begin{matrix} \xi_1 \\ \xi_2 \end{matrix} \right)' \Sigma^{-1} \left( \begin{matrix} \xi_1 \\ \xi_2 \end{matrix} - \begin{matrix} \xi^* \\ \xi^* \end{matrix} \right) > 0 \quad (3.1.7)$$

onde

$$\xi^* = \frac{\xi_1 + \xi_2}{2}$$

e aloca-se  $\xi$  ao grupo 2 em outro caso.

Seja

$$U = P S$$

a decomposição de Gram-Schmidt para a matriz das potências dos tempos, onde

$$P = \begin{bmatrix} \Phi_0^{(1)} & \Phi_1^{(1)} & \dots & \Phi_{p-1}^{(1)} \\ \Phi_0^{(2)} & \Phi_1^{(2)} & \dots & \Phi_{p-1}^{(2)} \\ \dots & \dots & \dots & \dots \\ \Phi_0^{(p)} & \Phi_1^{(p)} & \dots & \Phi_{p-1}^{(p)} \end{bmatrix}$$

é a matriz  $p \times p$  dos polinômios ortogonais de Fisher - Yates (1963) dados na tabela de Bock (1975) pg. 584, e  $S$  é uma matriz triangular superior.

A reparametrização polinômial é dada por, (3.1.1),

$$E(x_{\sim j \alpha}) = P \sum_{\sim j} \xi_{\sim j} \quad (3.1.8)$$

ou seja,

$$E(x_{\sim j \alpha}) = P \beta_{\sim j}$$

onde  $\beta_{\sim j}(\text{pxl})$  é o vetor dos coeficientes associados aos polinômios ortonormalizados do  $j$ -ésimo grupo, daqui temos,

$$\beta_{\sim j} = \sum_{\sim j} \xi_{\sim j} \quad (3.1.9)$$

Considere

$$y_{\sim j \alpha} = P' x_{\sim j \alpha} \quad (3.1.10)$$

onde  $y_{\sim j \alpha}$  é o vetor dos coeficientes ortonormais associados às curvas ajustadas ao  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo, de (3.1.4) e (3.1.8),

$$y_{\sim j \alpha} \sim N_p(\beta_{\sim j}, P' \Sigma P) \quad (3.1.11)$$

pois  $P'P = I$ .

Neste caso o cálculo do mínimo para os g grupos, (2.1.4) é dada por,

$$\text{Mín}_{\beta \in G} (\tilde{y} - \tilde{\beta}_i)' (\tilde{P}' \tilde{\Sigma} \tilde{P})^{-1} (\tilde{y} - \tilde{\beta}_i) = (\tilde{y} - \tilde{\beta}_j)' (\tilde{P}' \tilde{\Sigma} \tilde{P})^{-1} (\tilde{y} - \tilde{\beta}_j) \quad (3.1.12)$$

Particularmente no caso de dois grupos a regra de classificação (2.1.5) é dada por : aloca-se  $\tilde{y}$  ao grupo 1 se,

$$(\tilde{\beta}_1 - \tilde{\beta}_2)' (\tilde{P}' \tilde{\Sigma} \tilde{P})^{-1} (\tilde{y} - \tilde{\beta}^*) > 0 \quad (3.1.13)$$

onde  $\tilde{\beta}^*$  é o vetor de médias ,

$$\tilde{\beta}^* = \frac{\tilde{\beta}_1 + \tilde{\beta}_2}{2}$$

em caso contrário ao grupo 2, dependendo por um lado da magnitude da diferença dos coeficientes ortonormais dos grupos e por outro dos desvios dos coeficientes ortonormais da curva ajustada ao novo indivíduo com relação à média dos coeficientes ortonormais , através da matriz de variância e covariância das variáveis transformadas conforme (3.1.10).

É óbvia a equivalência entre a regra de classificação para os g grupos quando são considerados os discriminadores (variáveis originais) e quando são considerados os coeficientes

ortonormais ,de fato por (3.1.10) temos de (3.1.8) e (3.1.4) que

$$\left( \underset{\sim}{y} - \underset{\sim}{\beta} \right)' \left( \underset{\sim}{P}' \underset{\sim}{\Sigma} \underset{\sim}{P} \right)^{-1} \left( \underset{\sim}{y} - \underset{\sim}{\beta} \right) = \left( \underset{\sim}{x} - \underset{\sim}{P} \underset{\sim}{\beta} \right)' \underset{\sim}{\Sigma}^{-1} \left( \underset{\sim}{x} - \underset{\sim}{P} \underset{\sim}{\beta} \right)$$

A adoção dos coeficientes ortonormais se deve às correspondentes estruturas da matriz de variância e covariância. Para as estruturas consideradas da matriz de variância e covariância , as regras de discriminação como a representação gráfica dos coeficientes ganham em simplicidade .

A perda em interpretabilidade depende por (3.1.9) do grau do polinômio através da matriz de reparametrização \$ . Lembrando que nos coeficientes naturais a interpretabilidade de termos de maior grau , e a estrutura das inter-relações entre os coeficientes naturais , são elementos complicados. Ou seja, nos coeficientes naturais a simplicidade das variáveis (coeficientes) pode estar associada a estruturas de inter-relações fortes. Já , nos coeficientes ortonormais as inter-relações com os naturais são facilmente estabelecidas (3.1.9) e pode haver grande simplificação das estruturas de inter-relações . De fato, como será mostrado a seguir , no caso uniforme os coeficientes ortonormais são não correlacionados. No caso seriado, os coeficientes de ordem par são não correlacionados com os coeficientes de ordem ímpar.

### 3.2 PARA ESTRUTURA UNIFORME E SERIADA.

O cálculo do mínimo para o caso de  $g$  grupos quando é considerada a matriz de variância e covariância uniforme é ; de (3.1.12)

$$\text{Min}_{i \in G} (y - \beta_{i'})' \begin{bmatrix} \frac{1}{\sigma^2 + \sigma_\alpha^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma^2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{\sigma^2} \end{bmatrix} (y - \beta_{i'}) =$$

$$\frac{1}{\sigma^2 + \sigma_\alpha^2} (y_0 - \beta_{j0})^2 + \frac{1}{\sigma^2} \sum_{k=1}^{p-1} (y_k - \beta_{jk})^2 \quad (3.2.1)$$

onde  $\beta_{jk}$  é o coeficiente dos polinômios ortonormais de  $k$ -ésimo grau no  $j$ -ésimo grupo. A matriz da distância generalizada neste caso, no lado esquerdo da igualdade (3.2.1) é o resultado bastante conhecido e utilizado da diagonalidade da matriz de variância e covariância dos estimadores dos coeficientes polinômiais, detalhes no apêndice C.1.

Particularmente para o caso de dois grupos a regra de classificação (3.1.13) é, aloca-se  $y$  ao grupo 1 se,

$$\frac{1}{\sigma^2 + \sigma_\alpha^2} (\beta_{10} - \beta_{20})(y_0 - \beta_0^*) + \frac{1}{\sigma^2} \sum_{k=1}^{p-1} (\beta_{1k} - \beta_{2k})(y_k - \beta_k^*) > 0 \quad (3.2.2)$$

e aloca-se  $y$  ao grupo 2 em outro caso.

Estabelecimento da relação das regras de classificação quando temos coeficientes naturais e quando temos coeficientes ortogonais normalizados, de (3.1.9) e (3.1.11).

$$y = S \xi$$

$$\text{Min}_{i \in G} (S\xi - S\xi_i)' (P' \Sigma P)^{-1} (S\xi - S\xi_i) = (\xi - \xi_j)' (U' \Sigma U)^{-1} (\xi - \xi_j)$$

Particularmente em (3.1.13),

$$(S\xi_1 - S\xi_2)' (P' \Sigma P)^{-1} (S\xi - S\xi^*) = (\xi_1 - \xi_2)' (U' \Sigma U)^{-1} (\xi - \xi^*) > 0$$

O cálculo do mínimo para o caso de  $g$  grupos quando é considerada a matriz de variância e covariância de padrão de correlação serial, é de (3.1.12),

$$\text{Min}_{i \in G} (y - \beta_i)' (\Sigma_\beta)^{-1} (y - \beta_i) = (y - \beta_j)' (\Sigma_\beta)^{-1} (y - \beta_j) \quad (3.2.3)$$

onde  $\Sigma_{\beta}$  é a matriz de variância e covariância dos coeficientes, pois neste caso,

$$\Sigma_{\beta} = P' \Sigma P$$

tem como elemento da diagonal

$$\Sigma_{\beta_{(i+1)(i+1)}} = cte \left[ \sum_{k=1}^p P_i^2(t_k) + \sum_{k=2}^p \rho P_i(t_k) P_i(t_{k-1}) + \sum_{k=1}^{p-1} \rho P_i(t_k) P_i(t_{k+1}) + \right. \\ \left. \sum_{k=3}^p \rho^2 P_i(t_k) P_i(t_{k-2}) + \sum_{k=1}^{p-2} \rho^2 P_i(t_k) P_i(t_{k+2}) + \dots + \right. \\ \left. + \rho^{p-1} P_i(t_1) P_i(t_p) + \rho^{p-1} P_i(t_1) P_i(t_p) \right]$$

e

$$\Sigma_{\beta_{(i+1)(j+1)}} = cte \left[ \sum_{k=1}^p P_i(t_k) P_j(t_k) + \sum_{k=2}^p \rho P_i(t_k) P_j(t_{k-1}) + \sum_{k=1}^{p-1} \rho P_i(t_k) P_j(t_{k+1}) \right. \\ \left. + \sum_{k=3}^p \rho^2 P_i(t_k) P_j(t_{k-2}) + \sum_{k=1}^{p-2} \rho^2 P_i(t_k) P_j(t_{k+2}) + \dots + \right.$$

$$\left. \begin{aligned} & + \rho^{p-1} P_i(t_p) P_j(t_1) + \rho^{p-1} P_i(t_1) P_j(t_p) \end{aligned} \right\} \\
 \text{cte} = \frac{\sigma_u^2}{1 - \rho^2} \tag{3.2.4}$$

$$i = 0, 1, \dots, p-1$$

$$j = 0, 1, \dots, p-1$$

Os detalhes da manipulação algébrica encontram-se no apêndice C.2

E ainda, os elementos da matriz correspondentes às covariância entre as estimativas dos coeficientes um de ordem par e outro de ordem ímpar são nulos, de fato em (3.2.4), tomando  $i=2l$  e  $j=2m+1$ , tem-se

$$\Sigma_{\beta_{(2l+1)(2m+1)}} = \text{cte} \left[ \begin{aligned} & \sum_{k=2}^p \rho P_{2l}(t_k) P_{2m+1}(t_{k-1}) + \sum_{k=1}^{p-1} \rho P_{2l}(t_k) P_{2m+1}(t_{k+1}) + \\ & + \sum_{k=3}^p \rho^2 P_{2l}(t_k) P_{2m+1}(t_{k-2}) + \sum_{k=1}^{p-2} \rho^2 P_{2l}(t_k) P_{2m+1}(t_{k+2}) + \\ & + \dots + \rho^{p-1} P_{2l}(t_p) P_{2m+1}(t_1) + \rho^{p-1} P_{2l}(t_1) P_{2m+1}(t_p) \end{aligned} \right]$$



$$l = 0, 1, \dots, (p-2)/2 \quad \text{se } p \text{ é par,}$$

$$m = 0, 1, \dots, (p-2)/2$$

$$l = 0, 1, \dots, (p-1)/2 \quad \text{se } p \text{ é ímpar}$$

$$m = 0, 1, \dots, (p-3)/2$$

onde pela construção de  $P$ , Fisher e Yates (1963), nota-se que

$$\sum_{k=v+1}^p \rho^2 P_{2l}(t_k) P_{2m+1}(t_{k-v}) = - \sum_{k=1}^{p-v} \rho^2 P_{2l}(t_k) P_{2m+1}(t_{k+v}),$$

$$v = 1, 2, \dots, p-1$$

Identificando-se desta forma o seguinte padrão para a matriz de variância e covariância,

$$\left( \begin{matrix} \Sigma \\ \beta \end{matrix} \right)^{-1} = \begin{bmatrix} b_{11} & 0 & b_{13} & 0 & b_{15} & 0 & \dots & 0 \\ 0 & b_{22} & 0 & b_{24} & 0 & b_{26} & \dots & b_{2p} \\ b_{31} & 0 & b_{33} & 0 & b_{35} & 0 & \dots & 0 \\ 0 & b_{42} & 0 & b_{44} & 0 & b_{46} & \dots & b_{4p} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & b_{p2} & 0 & b_{p4} & 0 & b_{p6} & \dots & b_{pp} \end{bmatrix}$$

se  $p$  é par, e

$$\left( \Sigma_{\beta} \right)^{-1} = \begin{bmatrix} b_{11} & 0 & b_{13} & 0 & b_{15} & 0 & \dots & b_{1p} \\ 0 & b_{22} & 0 & b_{24} & 0 & b_{26} & \dots & 0 \\ b_{31} & 0 & b_{33} & 0 & b_{35} & 0 & \dots & b_{3p} \\ 0 & b_{42} & 0 & b_{44} & 0 & b_{46} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p1} & 0 & b_{p3} & 0 & b_{p4} & 0 & \dots & b_{pp} \end{bmatrix}$$

se  $p$  é ímpar.

A inversa dada por

$$\Sigma_{\beta}^{-1} = P \cdot \Sigma^{-1} P$$

tem como elemento da diagonal,

$$\left( \Sigma_{\beta} \right)^{-1}_{(i+1)(i+1)} = \frac{1}{\sigma_u^2} \left[ \sum_{k=1}^p P_l^2(t_k) - 2\rho \sum_{k=1}^{p-1} P_l(t_k) P_l(t_{k+1}) + \rho^2 \sum_{k=2}^{p-1} P_l^2(t_k) \right] \quad (3.2.5)$$

e

$$\left[ \begin{matrix} \Sigma \\ \beta \end{matrix} \right]^{-1}_{(l+1)(j+1)} = \frac{1}{\sigma^2_U} \left[ \begin{matrix} \rho^2 \sum_{k=2}^{p-1} P_{l(t_k)} P_{j(t_k)} - \rho \sum_{k=2}^p P_{l(t_k)} P_{(t_{k-1})} - \\ - \rho \sum_{k=1}^{p-1} P_{l(t_k)} P_{j(t_{k+1})} \end{matrix} \right]$$

$$l = 0, 1, \dots, p-1$$

$$j = 0, 1, \dots, p-1$$

Os detalhes da manipulação algébrica encontram-se apêndice C.2.

E ainda, em (3.2.5) tomando  $i = 2l$  e  $j = 2m+1$ , tem-se

$$\left[ \begin{matrix} \Sigma \\ \beta \end{matrix} \right]^{-1}_{(2l+1)(2m+1)} = \frac{1}{\sigma^2_U} \left[ \begin{matrix} \rho^2 \sum_{k=2}^{p-1} P_{2l(t_k)} P_{2m+1(t_k)} - \rho \sum_{k=2}^p P_{2l(t_k)} P_{2m+1(t_{k-1})} - \\ - \rho \sum_{k=1}^{p-1} P_{2l(t_k)} P_{2m+1(t_{k+1})} \end{matrix} \right]$$

onde pela construção de  $P$ , Fisher e Yates (1953), nota-se que,

$$\sum_{k=2}^{p-1} P_{2l}^{(t_k)} P_{2m+1}^{(t_k)} = 0$$

$$\sum_{k=2}^p P_{2l}^{(t_k)} P_{2m+1}^{(t_{k-1})} = - \sum_{k=1}^{p-1} P_{2l}^{(t_k)} P_{2m+1}^{(t_{k+1})}$$

$l = 0, 1, \dots, (p-2) / 2$   
 $m = 0, 1, \dots, (p-2) / 2$      se  $p$  é par ,

$l = 0, 1, \dots, (p-1) / 2$   
 $m = 0, 1, \dots, (p-3) / 2$      se  $p$  é ímpar ,

Identificando-se desta forma o seguinte padrão para a inversa da matriz de variância e covariância,

$$\left[ \Sigma_{\beta} \right]^{-1} = \begin{bmatrix} a_{11} & 0 & a_{13} & 0 & a_{15} & 0 & \dots & 0 \\ 0 & a_{22} & 0 & a_{24} & 0 & a_{26} & \dots & a_{2p} \\ a_{31} & 0 & a_{33} & 0 & a_{35} & 0 & \dots & 0 \\ 0 & a_{42} & 0 & a_{44} & 0 & a_{46} & \dots & a_{4p} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & a_{p2} & 0 & a_{p4} & 0 & a_{p6} & \dots & a_{pp} \end{bmatrix}$$

se  $p$  é par , e

$$\left( \begin{array}{c} \Sigma \\ \beta \end{array} \right)^{-1} = \begin{bmatrix} a_{11} & 0 & a_{13} & 0 & a_{15} & 0 & \dots & a_{1p} \\ 0 & a_{22} & 0 & a_{24} & 0 & a_{26} & \dots & 0 \\ a_{31} & 0 & a_{33} & 0 & a_{35} & 0 & \dots & a_{3p} \\ 0 & a_{42} & 0 & a_{44} & 0 & a_{46} & \dots & 0 \\ \vdots & & & & & & \ddots & \vdots \\ \vdots & & & & & & & \vdots \\ a_{p1} & 0 & a_{p3} & 0 & a_{p5} & 0 & \dots & a_{pp} \end{bmatrix}$$

se  $p$  é ímpar, as quais podem ser rearranjadas como abaixo, evidenciando que os coeficientes de ordem par são não correlacionados com os coeficientes de ordem ímpar.

$$\left( \begin{array}{c} \Sigma \\ \beta \end{array} \right)^{-1} = \begin{bmatrix} \Sigma_{\beta_{\text{par}}}^{-1} & 0 \\ 0 & \Sigma_{\beta_{\text{ímp.}}}^{-1} \end{bmatrix}$$

onde  $\Sigma_{\beta_{\text{par}}}^{-1}$  é  $(p/2, p/2)$  se  $p$  é par ou  $\Sigma_{\beta_{\text{ímp.}}}^{-1}$  é  $\left(\frac{p-1}{2}, \frac{p-1}{2}\right)$  se  $p$  é ímpar, que correspondem a rearranjar o vetor de diferenças dos coeficientes como

$$\left[ \begin{array}{c} y - \beta_j \\ \sim \\ \sim \end{array} \right] = \left[ \begin{array}{c} y_{\sim \text{par}} - \beta_{\sim \text{par}} \\ \hline y_{\sim \text{ímp.}} - \beta_{\sim \text{ímp.}} \end{array} \right]$$

reduzindo o cálculo do mínimo (3.2.3) a,

$$\begin{aligned}
\min_{\beta \in G} & \left( \begin{array}{c} y_{\text{par}} - \beta_{\text{par}} \\ y_{\text{imp}} - \beta_{\text{imp}} \end{array} \right)' \begin{pmatrix} \Sigma_{\beta_{\text{par}}}^{-1} & 0 \\ 0 & \Sigma_{\beta_{\text{imp}}}^{-1} \end{pmatrix} \begin{pmatrix} y_{\text{par}} - \beta_{\text{par}} \\ y_{\text{imp}} - \beta_{\text{imp}} \end{pmatrix} = \\
& = \begin{pmatrix} y_{\text{par}} - \beta_{\text{par}} \end{pmatrix}' \Sigma_{\beta_{\text{par}}}^{-1} \begin{pmatrix} y_{\text{par}} - \beta_{\text{par}} \end{pmatrix} + \\
& + \begin{pmatrix} y_{\text{imp}} - \beta_{\text{imp}} \end{pmatrix}' \Sigma_{\beta_{\text{imp}}}^{-1} \begin{pmatrix} y_{\text{imp}} - \beta_{\text{imp}} \end{pmatrix} \quad (3.2.5)
\end{aligned}$$

Para o caso particular de 2 grupos, (3.1.13) aloca  $y$  ao primeiro grupo se

$$\begin{aligned}
& \begin{pmatrix} \beta_{\text{par}} - \beta_{\text{par}}^* \\ \beta_{\text{imp}} - \beta_{\text{imp}}^* \end{pmatrix}' \begin{pmatrix} \Sigma_{\beta_{\text{par}}}^{-1} & 0 \\ 0 & \Sigma_{\beta_{\text{imp}}}^{-1} \end{pmatrix} \begin{pmatrix} y_{\text{par}} - \beta_{\text{par}}^* \\ y_{\text{imp}} - \beta_{\text{imp}}^* \end{pmatrix} > 0 \quad (3.2.6) \\
& \begin{pmatrix} \beta_{\text{par}} - \beta_{\text{par}}^* \end{pmatrix}' \Sigma_{\beta_{\text{par}}}^{-1} \begin{pmatrix} y_{\text{par}} - \beta_{\text{par}}^* \end{pmatrix} + \\
& + \begin{pmatrix} \beta_{\text{imp}} - \beta_{\text{imp}}^* \end{pmatrix}' \Sigma_{\beta_{\text{imp}}}^{-1} \begin{pmatrix} y_{\text{imp}} - \beta_{\text{imp}}^* \end{pmatrix} > 0
\end{aligned}$$

em caso contrário ao grupo 2.

A estrutura dos erros modelada em Kenward (1985) e tratada nos dados originais no capítulo anterior, transportada às variáveis alternativas leva aos seguintes resultados.

O cálculo do mínimo (3.2.3) é dada por,

$$\text{Min}_{\beta \in G} (y - \beta)' (\Sigma_{\beta_0})^{-1} (y - \beta) = (y - \beta)' (\Sigma_{\beta_0})^{-1} (y - \beta) \quad (3.2.7)$$

onde,

$$\Sigma_{\beta_0} = P' \Sigma_e P$$

$\Sigma_e$  é a matriz de variância e covariância quando os erros são modelados conforme apresentado por Kenward (1985) e foi dada no capítulo dois, a inversa de  $\Sigma_{\beta_0}$  detalhada no apêndice C.3 tem como elemento da diagonal,

$$\begin{aligned} (\Sigma_{\beta_0})_{(i+1)(i+1)}^{-1} = & \frac{1}{\sigma_\alpha^2} \left[ \sum_{k=1}^p P_i^2(t_k) - 2\rho \sum_{k=1}^{p-1} P_i(t_k) P_i(t_{k+1}) \right. \\ & \left. + \rho^2 \sum_{k=2}^{p-1} P_i^2(t_k) \right] - w(1-\rho)^2 + 2w\rho(1-\rho)^2 \sum_{k=2}^{p-1} P_{ik} + \\ & + w\rho^2(1-\rho)^2 \left[ \sum_{k=2}^{p-1} P_{ik} \right]^2 \end{aligned}$$

e

$$\left( \Sigma_{\beta_0} \right)_{(i+1)(j+1)}^{-1} = \frac{1}{\sigma_\alpha^2} \left[ \rho^2 \sum_{k=2}^{p-1} P_i(t_k) P_j(t_k) - \rho \sum_{k=2}^p P_i(t_k) P(t_{k+1}) - \right. \\ \left. - \rho \sum_{k=1}^{p-1} P_i(t_k) P_j(t_{k+1}) \right] - w(1-\rho)^2 + \\ + w\rho(1-\rho)^2 \left[ \sum_{k=2}^{p-1} P_{ik} + \sum_{l=2}^{p-1} P_{jl} \right] - w\rho^2(1-\rho)^2 \sum_{k=2}^{p-1} P_{ik} \sum_{l=2}^{p-1} P_{jl}$$

$$i = 0, 1, \dots, p-1$$

$$j = 0, 1, \dots, p-1$$

onde pela construção de  $P$ , Fisher e Yates (1963), as covariâncias entre coeficientes um de ordem par e outro de ordem ímpar são nulos. Desta forma a matriz de variância e covariância  $\Sigma_{\beta_0}$  e o correspondente vetor de medias podem ser rearranjados como no caso anterior, e o cálculo do mínimo (3.2.3) é reduzido a,

$$\min_{\beta \in G} \left( \begin{array}{c} y_{\text{par}} - \beta_{\text{par}} \\ \sim_{\text{par}} \quad \sim_{\text{par}} \end{array} \middle| \begin{array}{c} y_{\text{imp}} - \beta_{\text{imp}} \\ \sim_{\text{imp}} \quad \sim_{\text{imp}} \end{array} \right) \left( \begin{array}{c} \Sigma_{\beta_{\text{par}}}^{-1} \quad 0 \\ 0 \quad \Sigma_{\beta_{\text{imp}}}^{-1} \end{array} \right) \left( \begin{array}{c} y_{\text{par}} - \beta_{\text{par}} \\ \sim_{\text{par}} \quad \sim_{\text{par}} \\ \hline y_{\text{imp}} - \beta_{\text{imp}} \\ \sim_{\text{imp}} \quad \sim_{\text{imp}} \end{array} \right)$$



$$\begin{aligned}
&= \begin{pmatrix} y_{\sim \text{par}} - \beta_{\sim \text{par}} \\ \beta_{\sim \text{par}} \end{pmatrix} \Sigma_{\beta_{\sim \text{par}}}^{-1} \begin{pmatrix} y_{\sim \text{par}} - \beta_{\sim \text{par}} \\ \beta_{\sim \text{par}} \end{pmatrix} + \\
&+ \begin{pmatrix} y_{\sim \text{imp}} - \beta_{\sim \text{imp}} \\ \beta_{\sim \text{imp}} \end{pmatrix} \Sigma_{\beta_{\sim \text{imp}}}^{-1} \begin{pmatrix} y_{\sim \text{imp}} - \beta_{\sim \text{imp}} \\ \beta_{\sim \text{imp}} \end{pmatrix} \quad (3.2.8)
\end{aligned}$$

Para o caso particular de 2 grupos, (3.1.13)

$$\begin{aligned}
&\begin{pmatrix} \beta_{\sim 1 \text{par}} - \beta_{\sim 2 \text{par}} \\ \beta_{\sim 1 \text{imp}} - \beta_{\sim 2 \text{imp}} \end{pmatrix} \begin{pmatrix} \Sigma_{\beta_{\sim \text{par}}}^{-1} & 0 \\ 0 & \Sigma_{\beta_{\sim \text{imp}}}^{-1} \end{pmatrix} \begin{pmatrix} y_{\sim \text{par}} - \beta_{\sim \text{par}}^* \\ y_{\sim \text{imp}} - \beta_{\sim \text{imp}}^* \end{pmatrix} > 0 \\
&\begin{pmatrix} \beta_{\sim 1 \text{par}} - \beta_{\sim 2 \text{par}} \end{pmatrix} \Sigma_{\beta_{\sim \text{par}}}^{-1} \begin{pmatrix} y_{\sim \text{par}} - \beta_{\sim \text{par}}^* \end{pmatrix} + \\
&+ \begin{pmatrix} \beta_{\sim 1 \text{imp}} - \beta_{\sim 2 \text{imp}} \end{pmatrix} \Sigma_{\beta_{\sim \text{imp}}}^{-1} \begin{pmatrix} y_{\sim \text{imp}} - \beta_{\sim \text{imp}}^* \end{pmatrix} > 0 \quad (3.2.9)
\end{aligned}$$

Neste modelo comparado ao seriado usual é importante notar que a estrutura de inter-relações, ou seja, independência entre pares e ímpares foi mantida. Ainda, a variabilidade "entre" interfere apenas com a variância do termo constante, detalhes no apêndice C.3, analogamente ao caso uniforme em que a variabilidade "dentro" só afeta o termo constante.

## CAPITULO 4

### ALGUMA INTERPRETAÇÃO PARA DOIS GRUPOS

Dos resultados obtidos para matrizes de variância e covariância estruturadas, os relativos ao caso mais simples de dois grupos são revistos neste capítulo com alguns aspectos de interpretação. Obviamente em casos reais, o poder da interpretação ganha com o conhecimento da natureza específica do problema e no caso de mais grupos as extensões são claras. As funções discriminantes são rerepresentadas e a identificação de seus elementos relativos tanto à magnitude absoluta de separação entre os grupos (diferenças de efeitos ou coeficientes polinomiais), quanto às componentes de variabilidade e inter-relações, e da forma pela qual se combinam na função, é utilizada para interpretação. O coeficiente de correlação é utilizado com efeito comparativo visando corrigir o papel nem sempre positivo desempenhado pelos coeficientes de correlação altos. A análise univariada mais simples da estatística induz à anelar por coeficientes de correlação altos. Uma vez abordada a interpretação na secção 4.1, a secção seguinte trata de ilustrar via exemplo um caso de perfis com interação e para distintas

estruturas da matriz de variância e covariância .

Não se trata de um estudo de simulação , mas o exemplo ilustra confirmativamente os aspectos observados na função.

Ainda a secção 4.3 discute a utilização dos resultados obtidos

#### 4.1 INTERPRETAÇÃO NOS DADOS ORIGINAIS

i)  $\Sigma$  uniforme

a) Quando as diferenças dos efeitos dos grupos mudam de ocasião à ocasião , ou seja os perfis não são paralelos , a regra de classificação (2.2.2) , aloca-se  $x$  ao primeiro grupo se

$$\frac{1}{\sigma^2} \sum_{k=1}^p \left[ d_k - \frac{p \sigma_\alpha^2 \bar{d}}{\sigma^2 + p \sigma_\alpha^2} \right] \left[ x_k - \mu_k \right] > 0 \quad (4.1.1)$$

e ao segundo grupo no caso contrário.

Nota-se neste caso que a regra de classificação se reduz ao produto escalar usual do vetor de desvios da nova observação com relação à média geral, pelo vetor de desvios da

diferença de médias com relação ao seu vetor médio, corrigido pelo fator

$$\frac{\rho \sigma_{\alpha}^2}{\sigma^2 + \rho \sigma_{\alpha}^2} \quad (4.1.2)$$

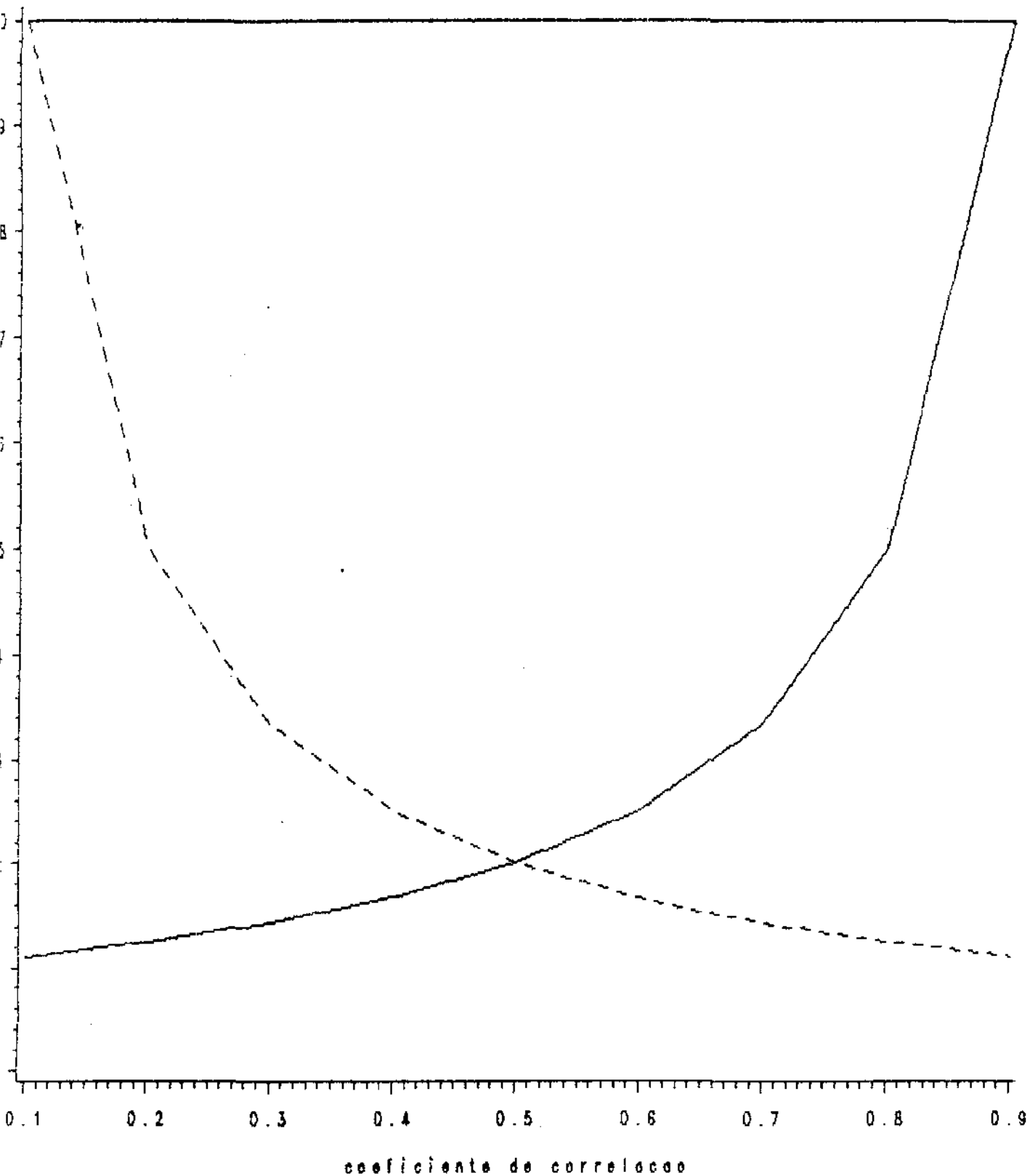
Uma forma de interpretar este fator é relacionando as variabilidades "dentro" e "entre" em função de  $\rho$ , de fato,

$$\sigma_{\alpha}^2 = \frac{\rho}{1 - \rho} \sigma^2 \quad e \quad \sigma^2 = \frac{1 - \rho}{\rho} \sigma_{\alpha}^2$$

Considerando duas situações limite quando o coeficiente de correlação  $\rho$  tende a um, primeiro para a variabilidade "entre" indivíduos ( $\sigma^2$ ) fixa e segundo para a variabilidade "dentro" do indivíduo ( $\sigma_{\alpha}^2$ ) fixa, nota-se que: no primeiro caso  $\sigma_{\alpha}^2$  tende a "explodir" e no segundo  $\sigma^2$  tende a valores pequenos, conforme exemplifica o gráfico 4.1.1. Para valores altos do coeficiente de correlação, as situações consideradas são exemplos de como as fontes "entre" e "dentro", neste caso de  $\Sigma$  uniforme, compõem o coeficiente de correlação. No primeiro caso ( $\sigma^2$  fixa), a variabilidade "dentro" grande relativamente à variabilidade "entre" dificulta a discriminação. No segundo caso ( $\sigma_{\alpha}^2$  fixa), a variabilidade "entre" pequena relativamente à variabilidade

# Variancia total

- 1) fixada variabilidade entre individuos = 1
  - 2) fixada a variabilidade dentro do individuo = 1
- em funcao do coeficiente de correlacao



— variabilidade entre individuos  
- - - variabilidade dentro do individuo

"dentro" facilita a discriminação

O exame da matriz de correlação é portanto informativo neste caso quando combinado com a análise univariada, que obviamente fornece as estimativas das variabilidades entre indivíduos e dentro do indivíduo.

Nestas duas situações distintas, o fator de correção tende a unidade reduzindo a regra de classificação à , aloca-se  $\tilde{x}$  ao primeiro grupo se,

$$\frac{1}{\sigma^2} \sum_{k=1}^p \left( d_k - \bar{d} \right) \left( x_k - \mu_k \right) > 0 \quad (4.1.3)$$

e ao segundo grupo no caso contrário.

b) Quando as diferenças dos efeitos dos grupos permanecem constantes de ocasião a ocasião, ou seja quando há paralelismo.

De (2.2.3) temos a regra de classificação aloca-se  $\tilde{x}$  ao primeiro grupo se

$$\frac{d}{\sigma^2 + p\sigma_\alpha^2} \sum_{k=1}^p \left( x_k - \mu_k \right) > 0 \quad (4.1.4)$$

e ao segundo grupo em caso contrário.

Neste caso a regra de classificação se reduz simplesmente ao produto escalar do vetor de diferenças de médias ( $\bar{d}$ ) pelo de desvios da nova observação com relação à média geral. O papel do vetor  $\bar{d}$  se reduz ao sinal da diferença dos efeitos, indiferente à magnitude. O papel da magnitude da diferença está implícito na média geral via componentes do vetor de desvios da nova observação com relação a média geral; neste também fica implícito o papel das fontes de variabilidade. Novamente, para coeficiente de correlação alto, considerados os dois casos das fontes "entre" e "dentro" fixadas, o poder de separação entre os grupos é menor quando fixada a variabilidade "entre" pois a variabilidade "dentro" cresce e maior quando fixada a variabilidade "dentro" pois a "entre" decresce para mesmos valores de  $p$ .

ii)  $\Sigma$  seriada

a) Quando os efeitos dos grupos mudam de ocasião a ocasião ou seja não permanecem constantes .

A regra de classificação (2.2.5) pode ser reescrita como , aloca-se  $x$  a grupo 1 se

$$\frac{1}{\sigma_u^2} \left[ \left( \sum_{k=2}^{p-1} d_k - \rho (d_{k-1} - \rho d_k + d_{k+1}) \right) (x_k - \mu_k) + \right. \\ \left. + (d_1 - \rho d_2) (x_1 - \mu_1) + (d_p - \rho d_{p-1}) (x_p - \mu_p) \right] \geq 0 \quad (4.1.5)$$

no caso contrário , aloca-se  $x$  ao grupo 2.

O produto escalar neste caso tem o vetor dos desvios das diferenças dos efeitos tomado nas extremidades com relação as diferenças vizinhas corrigidas pelo fator  $\rho$  , e no centro com relação à expressão

$$\rho (d_{k+1} + d_{k-1}) - \rho^2 d_k$$



b) quando não existe mudanças nos efeitos dos grupos de ocasião a ocasião, ou seja estes permanecem constantes.

A regra de classificação foi dada em (2.2.6). Aloca-se  $x$  ao grupo 1 se

$$\frac{d}{\sigma_u^2} \left[ (1-2\rho+\rho^2) \sum_{k=2}^{p-1} (x_k - \mu_k) + (1-\rho)((x_1 - \mu_1) + (x_p - \mu_p)) \right] > 0 \quad (4.1.6)$$

em caso contrário, aloca-se ao grupo 2.

Este caso evidencia o comentado anteriormente pois valores altos de  $\rho$  diminuem o peso relativo do vetor dos desvios da nova observação à média geral. Ainda, note-se a maior magnitude do fator de ponderação nas componentes extremas do vetor de desvios. Como no caso uniforme o vetor de diferenças dos efeitos tem papel de dar sinal à regra e sua atuação quanto à magnitude é indireta via vetor de desvios da nova observação com relação à média geral.

A interpretação do modelo seriado de Kenward (1985) será deixada para a continuação deste trabalho.

#### 4.2 INTERPRETAÇÃO NOS COEFICIENTES POLINOMIAIS

i)  $\Sigma$  uniforme

a) Quando os perfis não são paralelos .

A regra de classificação (3.2.2), aloca-se  $y$  ao grupo 1 se,

$$\frac{1}{\sigma^2 + \sigma_\alpha^2} \begin{bmatrix} \beta_{10} - \beta_{20} \\ \beta_{11} - \beta_{21} \\ \vdots \\ \beta_{1p} - \beta_{2p} \end{bmatrix} \begin{bmatrix} y_0 - \beta_{00} \\ y_1 - \beta_{10} \\ \vdots \\ y_p - \beta_{p0} \end{bmatrix} + \frac{1}{\sigma^2} \sum_{k=1}^{p-1} \begin{bmatrix} \beta_{1k} - \beta_{2k} \\ \beta_{1k} - \beta_{2k} \\ \vdots \\ \beta_{1k} - \beta_{2k} \end{bmatrix} \begin{bmatrix} y_k - \beta_{k0} \\ y_k - \beta_{k0} \\ \vdots \\ y_k - \beta_{k0} \end{bmatrix} > 0 \quad (4.2.1)$$

e caso contrário ao grupo 2.

A regra de classificação neste caso pode ser identificada como o produto escalar comum de dois vetores, o de diferenças entre os coeficientes de mesmo grau corrigidas pelos fatores multiplicativos correspondentes e o de desvios entre os coeficientes do novo indivíduo em relação à média dos coeficientes (mantida a correspondência dos graus). Observa-se que os fatores das diferenças entre os coeficientes de mesmo grau nada mais são que os inversos das variâncias correspondentes a cada um deles, conforme (3.1.11) e (3.1.12).

Nota-se aqui que, o peso da componente relativa ao coeficiente constante é menor que o das componentes relativas aos termos de maior grau.

Ou seja se os perfis não são paralelos, a importância da magnitude das diferenças dos coeficientes de maior grau é inversamente proporcional a variabilidade entre indivíduos, e de forma geral a do termo constante é inversamente proporcional à variação total.

Comparando-se esta regra com a obtida nos dados originais (4.1.1), é de destacar a possibilidade de melhor diagnóstico da separação entre os grupos.

De forma geral o exame desta função discriminante pode revelar, tomada em conta a estrutura de  $\Sigma$  uniforme, a importância das parcelas e portanto dos coeficientes "responsáveis" pela classificação, uma vez que em cada parcela está isolado o fator diferença do coeficiente de cada grau padronizado pela variância do coeficiente. a questão da interpretabilidade ganha então espaço pois, aos termos de cada grau, podem corresponder diferentes significados.

b) Quando os perfis são paralelos.

A regra de classificação (4.2.1) reduzida aloca-se  $\tilde{y}$  ao primeiro grupo se,

$$\frac{1}{\sigma^2 + \sigma_\alpha^2} \begin{bmatrix} \beta_{10} - \beta_{20} \end{bmatrix} \begin{bmatrix} y_0 - \beta_0^* \end{bmatrix} > 0 \quad (4.2.2)$$

e caso contrário ao grupo 2

Se os perfis são paralelos o poder de separação nos grupos depende do sinal da diferença nos coeficientes de grau zero e implicitamente da variância do termo constante que aparece na variação de  $y_0$  em torno de  $\beta_0^*$ .

ii)  $\Sigma$  seriada

a) perfis não são paralelos.

A regra de classificação (3.2.9) é aloca-se  $\tilde{y}$  ao primeiro grupo se,

$$\begin{bmatrix} \beta_{1par} - \beta_{2par} \end{bmatrix}' \Sigma_{\beta par}^{-1} \begin{bmatrix} y_{\beta par} - \beta_{par}^* \end{bmatrix} + \begin{bmatrix} \beta_{1imp} - \beta_{2imp} \end{bmatrix}' \Sigma_{\beta imp}^{-1} \begin{bmatrix} y_{\beta imp} - \beta_{imp}^* \end{bmatrix} > 0 \quad (4.2.4)$$

em caso contrário ao grupo 2.

Neste caso a regra de classificação não se reduz ao produto escalar comum dos vetores de diferenças, mas à soma de dois produtos via as correspondentes inversas da matriz de

variância e covariância. Como mostrado a partir de (3.2.4) até (3.2.6), as estimativas dos coeficientes dos polinômios de ordem par não são correlacionados com os coeficientes de ordem ímpar, como apontou Kenward (1985). Um exame de (3.2.4) e (3.2.6) revela dificuldade de interpretação geral dentro dos blocos correspondentes à variação dentro dos coeficientes de ordem par ou ímpar, no entanto casos específicos de polinômios de grau reduzido podem ser explorados. Dos resultados obtidos nos dados originais, valores altos de  $\rho$  diminuem o poder de separação entre grupos.

b) perfis paralelos

A regra de classificação (3.2.6) é reduzida a, aloca-se  $y$  ao primeiro grupo se,

$$\left[ \beta_{10}^* - \beta_{20}^* \right] a_{11} \left[ y_0 - \beta_0^* \right] > 0 \quad (4.2.5)$$

e ao grupo 2 em caso contrário.

onde,

$a_{11}$  é dada em (3.2.5) para  $i = 1$ .

A diferença dos grupos estará evidentemente no coeficiente de grau zero ( termo constante), mas também é obvio que a correlação com os demais coeficientes pares aparece na medida que a matriz  $\Sigma_{\beta_{par}}$  se distancia do padrão diagonal ou seja

na medida que  $\rho$  cresce.

#### 4.3 O CASO DE DOIS GRUPOS VIA EXEMPLO

Foi considerada uma situação inspirada na literatura de medidas repetidas, como uma forma de ilustrar as interpretações feitas anteriormente quando são ajustadas curvas polinomiais no caso de dois grupos.

As curvas médias de resposta para o primeiro e segundo grupos foram tomadas como as funções de primeiro grau abaixo, que para ocasiões  $t = 1, 2, 3, 4$  têm os correspondentes vetores de média  $\mu_1$  e  $\mu_2$ .

$$E(x_{1\alpha}) = 5 + 0,4t \quad \mu_1 = (5,4 \ 5,8 \ 6,2 \ 6,6)$$

$$E(x_{2\alpha}) = 8,5 - 0,5t \quad \mu_2 = (8,0 \ 7,5 \ 7,0 \ 6,5)$$

A geração exemplificada aqui é de amostras de tamanho dez, em torno dos vetores de médias descritos acima, os detalhes da geração encontram-se no apêndice D.1, e segundo as estruturas da matriz de variância e covariância descritas na tabela 4.3.1

TABELA 4.3.1 Casos considerados das estruturas de  $\Sigma$  para efeito de geração.

Valor de $\rho$	Uniforme		Seriada
	Variabilidade fixada	Variabilidade correspondente	
0,1	$\sigma^2 = 1$ $\sigma_\alpha^2 = 1$	$\sigma_\alpha^2 = 0.111$ $\sigma^2 = 0$	$\sigma_u^2 = 1$
0,5	$\sigma^2 = 1$ $\sigma_\alpha^2 = 1$	$\sigma_\alpha^2 = 1$ $\sigma^2 = 1$	$\sigma_u^2 = 1$
0,9	$\sigma^2 = 1$ $\sigma_\alpha^2 = 1$	$\sigma_\alpha^2 = 0$ $\sigma^2 = 0.111$	$\sigma_u^2 = 1$

Nota :

$\sigma^2$  : variabilidade entre indivíduos.

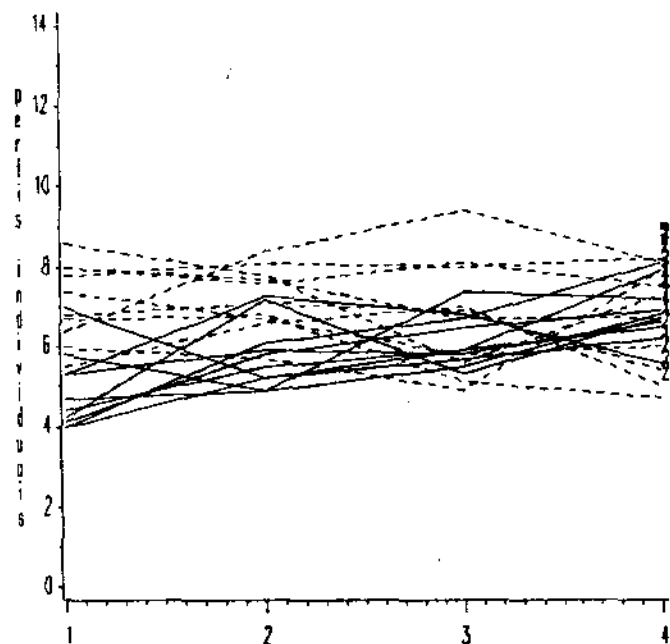
$\sigma_\alpha^2$  : variabilidade dentro do indivíduo.

Variabilidade correspondente : o valor da variabilidade não fixada que correspondente ao valor de  $\rho$  da coluna um, tendo-se em conta que :

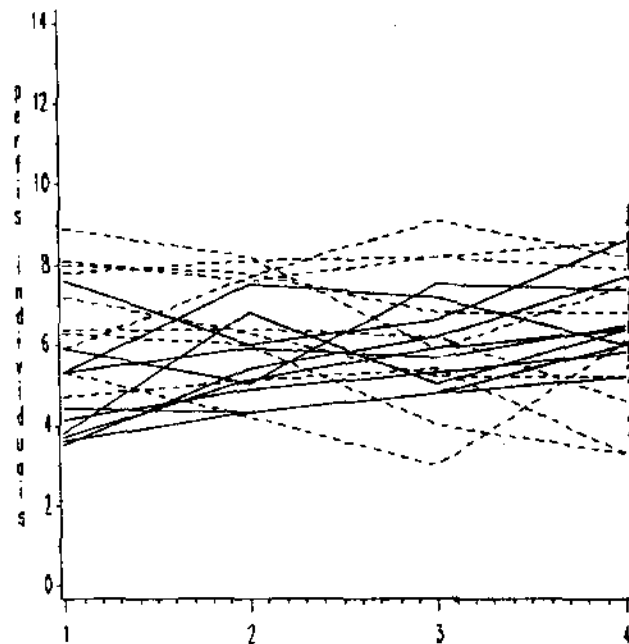
$$\rho^2 = \frac{\sigma_\alpha^2}{\sigma^2 + \sigma_\alpha^2}$$

# perfis individuais dos grupos por ocasio

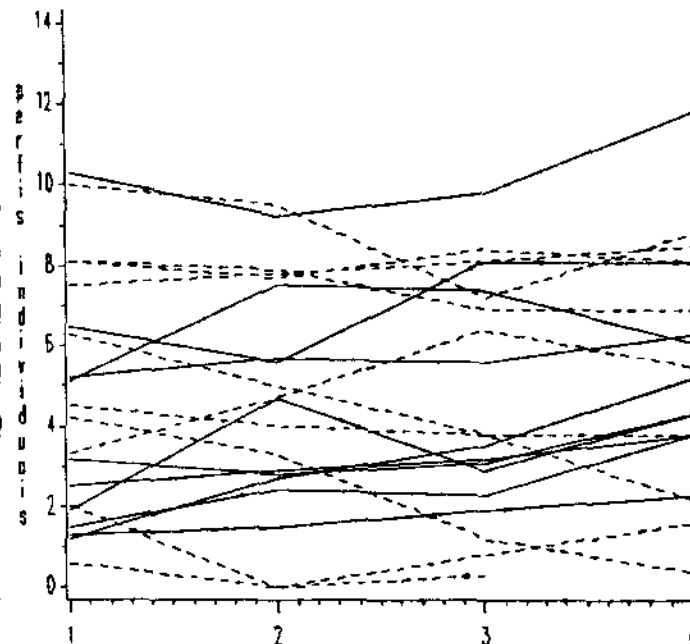
caso uniforme fixada variabilidade entre -1



correlacao 0.1



correlacao 0.5



correlacao 0.9

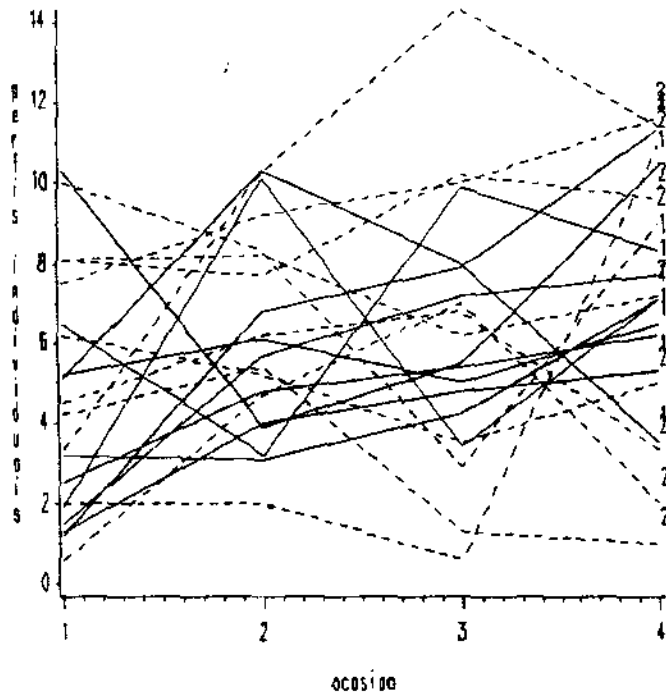
———— grupo 1  
----- grupo 2

grafico 4.3.1

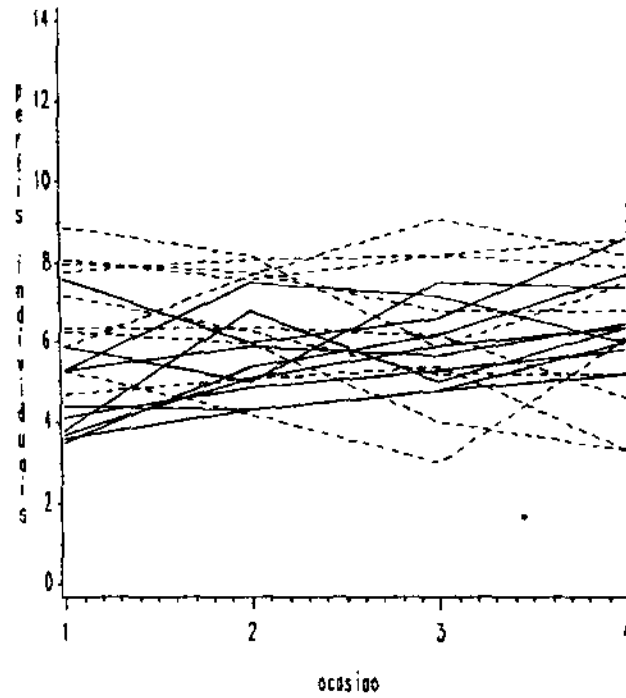


# perfis individuais dos grupos por ocasiao

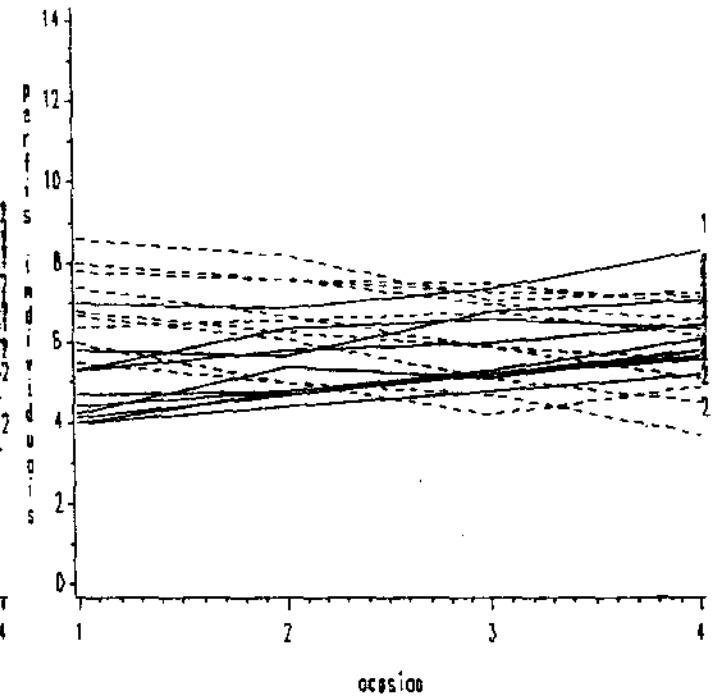
caso uniforme fixada variabilidade dentro =1



correlacao 0.1



correlacao 0.5

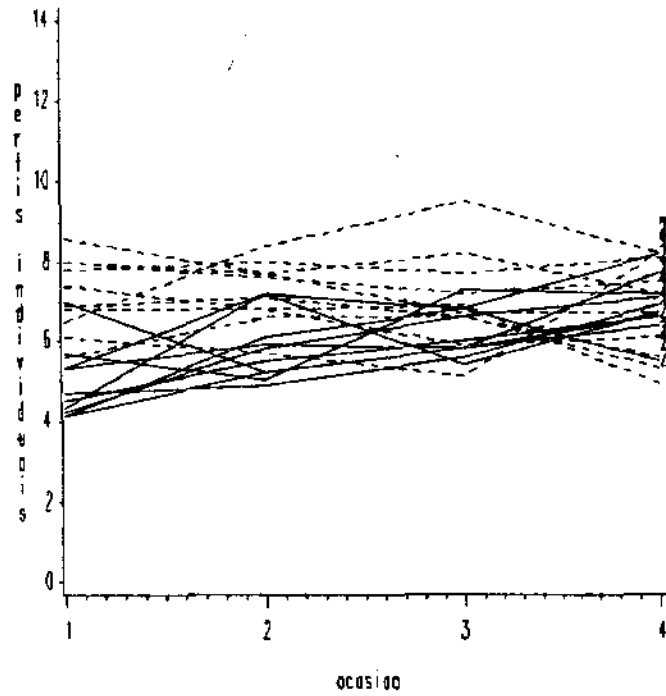


correlacao 0.9

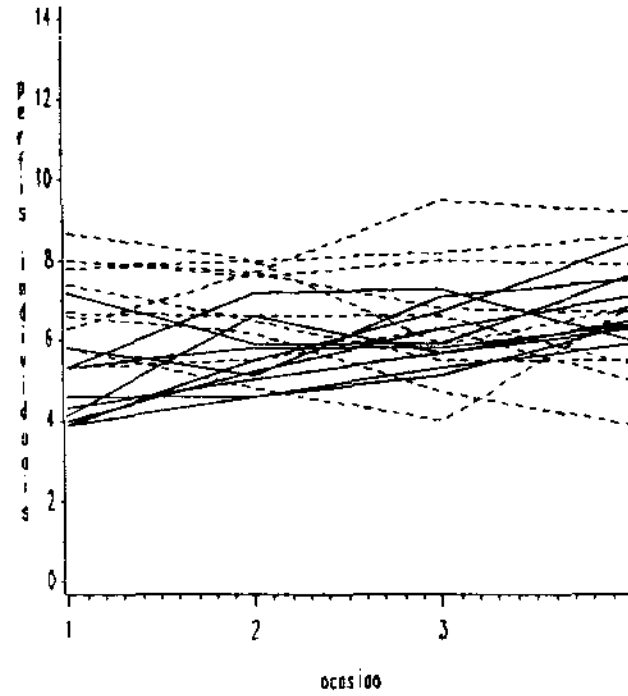
———— grupo 1  
----- grupo 2  
grafico 4.3.2

# perfis individuais dos grupos por ocasiao

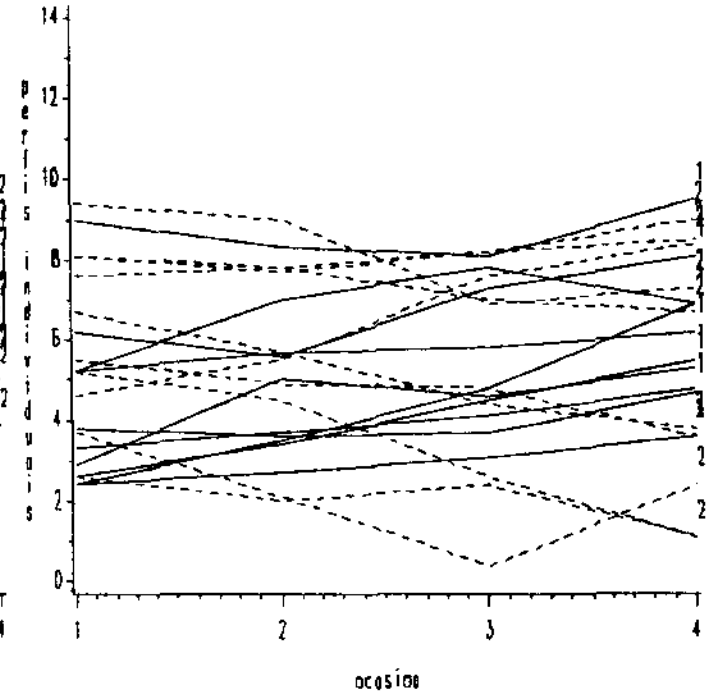
caso seriada fixada escala de variabilidade



correlacao 0.1



correlacao 0.5



correlacao 0.9

———— grupo 1  
----- grupo 2

grafico 4.3.3

Os perfis gerados com a  $\Sigma$  uniforme, fixada a variabilidade entre indivíduos, e, para diferentes valores de  $\rho$ , encontram-se no gráfico (4.3.1). Aprecia-se que a medida que o coeficiente de correlação aumenta os perfis se confundem fazendo mais difícil a discriminação.

Os perfis gerados com a  $\Sigma$  uniforme, fixada a variabilidade dentro do indivíduo, e, para os diferentes valores de  $\rho$ , encontram-se no gráfico (4.3.2). Aprecia-se que a medida que o coeficiente de correlação aumenta é evidente a separação dos perfis para cada grupo, o que levaria a uma melhor discriminação.

Os perfis gerados com a  $\Sigma$  seriada, fixada o valor de  $\sigma_u^2 = 1$ , e para os diferentes valores de  $\rho$  encontram-se no gráfico (4.3.3). Aprecia-se que a medida que o coeficiente de correlação cresce o poder de discriminação diminui.

Comparando os casos estruturados uniforme e seriado, no primeiro para valor da variabilidade dentro do indivíduo fixada, ao aumento do coeficiente de correlação corresponde no primeiro caso a um decréscimo da variação total e no segundo ao comportamento oposto.

Portanto nestas circunstâncias a discriminação é melhor para valores altos de  $\rho$  no caso uniforme e baixos no caso seriado.

O exemplo adotado para ilustração leva a indicação de conclusões de ordem mais geral, uma vez que o padrão de dispersão

de fato independe do modelo das médias. Variação no modelo de médias afeta sem dúvida a questão da discriminação , pois envolve a magnitude de diferença entre os coeficientes , assim a observação acima é válida dentro de um modelo de medias fixo.

A vantagem de utilizar os coeficientes em lugar dos dados originais na classificação é a redução de dimensão , pois ao ajustar curvas polinomiais o que se trata é de explicar a variação dos perfis em função de um menor número de parâmetros ,isto é, levar em consideração o que esta acontecendo em todas as ocasiões com menor número de parâmetros, permitindo estabelecer diferenças entre os grupos, tanto como quando são utilizados os dados originais.

Permite uma maior interpretabilidade dos coeficientes que terão o poder de separação, por exemplo se os grupos se diferenciam pelo coeficiente linear e que a taxa de crescimento medio nos dos grupos são diferentes ; se a diferença esta no coeficiente do termo quadrático a taxa de mudança na taxa de crescimento nos dois grupos são diferentes e assim por diante.

#### 4.4 CONSIDERAÇÕES EM CASOS REAIS

O tratamento da análise discriminante dado neste trabalho suposta válida a multinormalidade, abordada por exemplo por Csorgo (1986); Leslie et al. (1986); Mason e Young (1986) esbarra em pelo menos duas questões: a da estruturação da matriz de variância e covariância e a do modelo polinomial.

Quanto à estruturação da matriz de variância e covariância, os padrões uniforme e seriado, ou ainda outros de natureza não estacionária como por exemplo o de Mansour et al. (1985), envolvem as razões latentes ao problema que tornam plausível levantar hipóteses sobre determinada estrutura os procedimentos inferenciais para não rejeitar a validade do mesmo e a decorrente especificidade, quando aplicável, do método estatístico de análise.

Os testes do padrão uniforme vem desde Box (1950), avançando com a generalização de Huynh e Feldt (1970) e também tratado em textos como de Bock (1975) e discutida por Infante (1989). A estimação do padrão serial pode ser encontrado por exemplo em Hearne et al. (1983), no enfoque clássico da análise multivariada. As questões do padrão seriado ou formulação não estacionárias podem ser tratadas via soluções e metodologia de series temporais.

Quanto à questão do modelo polinomial , desde Potthoff e Roy (1964) passando pelas contribuições de Rao (1965,1966,1967) e discussões aplicadas como em Grizzle e Allen (1969) , a escolha do grau do polinômio e as diferenças entre grupos diagnosticadas via diferenças entre coeficientes vem evoluindo.

Bock (1975) sistematiza a determinação do grau do polinômio nos modelos univariado (  $\Sigma$  uniforme ) e multivariado (  $\Sigma$  geral ) , e os testes em cada caso que levam a validar ou não a situação de paralelismo , neste último caso resultando em possibilidades de diferentes curvas para os grupos. Em cada caso vale evoluir para o ajuste de covariáveis proposto por Rao (1965) e discutido por Kenward ( 1985, 1986) .

Vale notar que nos casos estruturados, este tratamento de ajuste por covariáveis se especializa : no caso uniforme não deve ser utilizado pois o ajuste via polinômios ortogonais resulta em independência entre os coeficientes (3.2.1) , o que não justifica a análise de covariância ; no caso seriado a independência entre os coeficientes de ordem par e os de ordem ímpar sugere adequação do tratamento via covariáveis , sendo o ajuste feito dentro do grupo dos coeficientes pares e dentro do grupo dos coeficientes ímpares . Ainda recorrendo e recuperando o trabalho de Cochran e Bliss (1948) a escolha das covariáveis pode ser mais flexível como corrobora o trabalho de Kenward (1985) em

que os coeficientes para ajuste são subconjuntos dos de esperança nula abrindo a perspectiva de levar em conta as estruturas de covariância. Por exemplo no caso seriado, polinômios do primeiro grau adotado e o modelo saturado sendo de grau dois apenas o termo constante seria corrigido pelo quadrático. Não há correção possível para o termo linear (ou taxa de crescimento) neste caso.

Portanto, no que diz respeito à estimação da função discriminante sobre as variáveis alternativas nos casos estudados, a do caso uniforme não passa pelo ajuste de covariáveis. No entanto, em outras estruturas onde o ajuste por covariáveis seja recomendável, a análise deve ser feita nos coeficientes ajustados. No caso da estrutura seriada a função discriminante nos coeficientes ajustados, dentro de cada grupo de pares e ímpares, se particulariza em função da especificidade da matriz de variância e covariância dos resíduos.

Dadas as possibilidades de diferentes escolhas para os conjuntos de covariáveis não parece vantajoso um tratamento algébrico geral.

Um estudo via simulações comparando o tratamento específico da discriminação quando  $\Sigma$  é estruturado, com o tratamento usual, via pacotes (que ignora este fato) seria útil para quantificar as esperadas perdas no poder de classificação da técnica. Isto se aplica tanto ao tratamento nas variáveis

originais quanto nas alternativas . Dada as limitações do tempo não foi possível realizá-lo.



## APÊNDICE

## APÊNDICE A

Surgimento das diferentes estruturas da matriz de variância e covariância.

### A.1 $\Sigma$ uniforme

O surgimento da matriz de variância e covariância uniforme da secção 2.2, pode ser observado no modelo paramétrico misto apresentado por exemplo em Bock (1975), detalhado a seguir.

$$x_{j\alpha k} = \mu + \lambda_{j\alpha} + \tau_k + \theta_j + (\theta\tau)_{jk} + \varepsilon_{j\alpha k}$$

onde,

$\mu$  = média geral

$\lambda_{j\alpha}$  = efeito do  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo

$\tau_k$  = efeito da  $k$ -ésima ocasião, constante para todo indivíduo

$\theta_j$  = efeito do  $j$ -ésimo grupo

$(\theta\tau)_{jk}$  = efeito da iteração do  $j$ -ésimo grupo e  $k$ -ésima ocasião

$\varepsilon_{j\alpha k}$  = erro específico ao  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo na  $k$ -ésima ocasião.

além disso,

$$\lambda_{j\alpha} \sim N(0, \sigma_\alpha^2) \quad (\text{A.1.1})$$

$$\varepsilon_{j\alpha k} \sim N(0, \sigma^2) \quad (\text{A.1.2})$$

$$\text{Cov}(\lambda_{j\alpha}, \varepsilon_{j\alpha k}) = 0 \quad (\text{A.1.3})$$

$$\text{Cov}(\varepsilon_{j\alpha k}, \varepsilon_{j\alpha k'}) = 0, \quad k \neq k' \quad (\text{A.1.4})$$

isto é,

$$\text{Cov} ( x_{jak}, x_{jal} ) = \begin{cases} \sigma_{\alpha}^2 + \sigma^2 & \alpha = \beta, k=l \\ \sigma_{\alpha}^2 & \alpha = \beta, k \neq l \\ 0 & \alpha \neq \beta \end{cases}$$

De fato para  $\alpha = \beta$  e  $k=l$  por (A.1.1) e (A.1.2)

$$\begin{aligned} V ( x_{jak} ) &= V ( \mu + \lambda_{ja} + \tau_k + \theta_j + (\theta\tau)_{jk} + \epsilon_{ajk} ) \\ &= \sigma_{\alpha}^2 + \sigma^2 \end{aligned}$$

Para  $\alpha = \beta$  e  $k \neq l$

$$\text{Cov} ( x_{jak}, x_{jal} ) = E ( x_{jak} x_{jal} ) - E ( x_{jak} ) E ( x_{jal} )$$

$$E ( x_{j\alpha m} ) = \mu + \tau_m + \theta_j + (\theta\tau)_{jm}$$

$$E ( x_{jak} x_{jal} ) = E \left[ \begin{aligned} & ( \mu + \lambda_{ja} + \tau_k + \theta_j + (\theta\tau)_{jk} + \epsilon_{ajk} ) \\ & ( \mu + \lambda_{ja} + \tau_l + \theta_j + (\theta\tau)_{jl} + \epsilon_{ajl} ) \end{aligned} \right]$$

$$\begin{aligned} \text{Cov} ( x_{j\alpha m}, x_{j\alpha l} ) &= E ( \lambda_{\alpha k} \lambda_{\alpha k} ) + E ( \lambda_{\alpha k} \epsilon_{jak} ) + E ( \lambda_{\alpha l} \epsilon_{jal} ) + \\ &+ E ( \epsilon_{jak} \epsilon_{jal} ) \end{aligned}$$

$$\text{Cov} ( x_{jak}, x_{jal} ) = \sigma_{\alpha}^2$$

por (A.1.1) , (A.1.3) e (A.1.4).

Para diferentes indivíduos ,  $\alpha \neq \beta$

$$\text{Cov} ( x_{j\alpha k} , x_{j\beta l} ) = 0$$

pela independência entre indivíduos.

Na notação matricial portanto,

$$V ( \underset{\sim}{x}_{j\alpha} ) = \sigma^2 \mathbb{I} + \sigma_\alpha^2 \mathbf{1} \mathbf{1}'$$

onde

$\mathbb{I}$  é a matriz identidade de dimensão  $p \times p$

$$\mathbf{1} = [ 1 \ 1 \ \dots \ 1 ]'$$

é o vetor dimensão  $p$ .

#### A.2 $\Sigma$ seriada

O surgimento da matriz de variância e covariância seriada da secção 2.2 , pode ser observada no modelo

$$\underset{\sim}{x}_{j\alpha} = \mathbb{A} \underset{\sim}{\mu}_j + \underset{\sim}{e}_{j\alpha}$$

onde,

$\underset{\sim}{\mu}_j$  = é o vetor de médias do  $j$ -ésimo grupo

$\underset{\sim}{e}_{j\alpha}$  = vetor dos erros

$\mathbb{A}$  é a matriz de especificação.

E por hipótese,

$$\epsilon_{j\alpha k} = \rho \epsilon_{j\alpha(k-1)} + \eta_{j\alpha k} \quad , \quad | \rho | < 1 \quad , \quad k = 1, 2, \dots, p \quad (\text{A.2.1})$$

$$\eta_{jak} \sim N ( 0, \sigma_u^2 ) \quad (A.2.2)$$

$$\text{Cov} ( \varepsilon_{jak}, \eta_{jak(k+1)} ) = 0 \quad (A.2.3)$$

$$\varepsilon_{jao} \sim N \left( 0, \frac{\sigma_u^2}{1 - \rho^2} \right) \quad (A.2.4)$$

Ou seja quando o modelo é autoregressivo de primeira ordem no vetor de erros que descreve o padrão de correlação entre observações de uma mesma unidade experimental.

A matriz de variância e covariância induzida por esse modelo através de substituições sucessivas é.

$$\text{Cov} ( \varepsilon_{jak}, \varepsilon_{jak'} ) = \begin{cases} \frac{\sigma_u^2}{1 - \rho^2} & k = k' \\ \frac{\rho^s \sigma_u^2}{1 - \rho^2} & k \neq k' \\ & k' = k+s \end{cases}$$

De fato, por (A.2.1)

$$\varepsilon_{j\alpha k} = \sum_{m=0}^{k-1} \rho^m \eta_{j\alpha(k-m)} + \rho^k \varepsilon_{j\alpha 0}$$

$$V(\varepsilon_{j\alpha k}) = V\left(\sum_{m=0}^{k-1} \rho^m \eta_{j\alpha(k-m)} + \rho^k \varepsilon_{j\alpha 0}\right)$$

$$V(x_{j\alpha k}) = \sum_{m=0}^{k-1} \rho^{2m} \sigma_u^2 + \frac{\rho^{2k} \sigma_u^2}{1 - \rho^2} = \frac{\sigma_u^2}{1 - \rho^2}$$

$$\text{Cov}(\varepsilon_{j\alpha k}, \varepsilon_{j\alpha(k-s)}) = E\left[\left(\sum_{m=0}^{k-1} \rho^m \eta_{j\alpha(k-m)} + \rho^k \varepsilon_{j\alpha 0}\right) \left(\sum_{m=0}^{k-s-1} \rho^m \eta_{j\alpha(k-s-m)} + \rho^{k-s} \varepsilon_{j\alpha 0}\right)\right]$$

$$\text{Cov}(\varepsilon_{j\alpha k}, \varepsilon_{j\alpha(k-s)}) = \rho^s \frac{\sigma_u^2}{1 - \rho^2}$$

por (A.2.2), (A.2.3) e (A.2.4)

Na notação matricial portanto,

$$V(x_{j\alpha}) = \frac{\sigma_u^2}{1 - \rho^2} C_\rho, \quad |\rho| < 1 \quad (\text{A.2.5})$$

onde,

$$C_p = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \dots & \rho^{p-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{p-1} & \rho^{p-2} & \dots & & 1 \end{bmatrix}$$

### A.3 $\Sigma$ seriada de Kenward

O surgimento da matriz de variância e covariância seriada, apresentada por Kenward (1985), pode ser observado no modelo

$$x_{j\alpha} = \tilde{A} \mu_j + \tilde{e}_{j\alpha}, \quad \alpha = 1, 2, \dots, n_j \\ j \in G$$

onde,

$$\varepsilon_{j\alpha k} = \lambda_{\alpha k} + \theta_j$$

$\lambda_{\alpha k}$  = efeito do  $\alpha$ -ésimo indivíduo na  $k$ -ésima ocasião

$\theta_j$  = efeito do  $j$ -ésimo grupo

$\varepsilon_{j\alpha k}$  = erro aleatório considerado como a soma de duas componentes independentes,  $\theta_j$  entre indivíduos,  $\lambda_{\alpha k}$  dentro do indivíduo, esta modelada por um processo autoregressivo de primeira ordem, isto é,

$$\begin{aligned} E(\theta_j) &= 0 \\ E(\theta_j^2) &= \sigma^2 \end{aligned} \tag{A.3.1}$$



$$\lambda_{\alpha k} = \begin{cases} b_1 & k = 1 \\ \rho \lambda_{\alpha(k-1)} + b_k & k = 2, 3, \dots, p \end{cases} \quad (\text{A.3.2})$$

$$E(\lambda_{\alpha k}) = 0$$

$$E(b_k) = 0$$

$$E(b_k b_l) = \begin{cases} \sigma_\alpha^2 & k = l \\ 0 & k \neq l \end{cases} \quad (\text{A.3.3})$$

tal como foi calculado no apêndice A.2, a matriz de variância e covariância de um processo autoregressivo de primeira ordem é,

$$V(\lambda_{\alpha k}) = \frac{\sigma_\alpha^2}{1 - \rho^2} \quad (\text{A.3.4})$$

$$\text{Cov}(\lambda_{\alpha k}, \lambda_{\alpha(k-e)}) = \rho^e \frac{\sigma_\alpha^2}{1 - \rho^2}$$

então,

$$V(\varepsilon_{j\alpha k}) = V(\lambda_{\alpha k} + \theta_j) = \frac{\sigma_\alpha^2}{1 - \rho^2} + \sigma_j^2$$

$$\begin{aligned} \text{Cov}(\varepsilon_{j\alpha k}, \varepsilon_{j\alpha(k-e)}) &= E\left[(\lambda_{\alpha k} + \theta_j)(\lambda_{\alpha(k-e)} + \theta_j)\right] \\ &= E\left[(\lambda_{\alpha k} \lambda_{\alpha(k-e)} + \theta_j^2)\right] \end{aligned}$$

$$= \rho^2 \frac{\sigma_\alpha^2}{1 - \rho^2} + \sigma^2$$

por (A.3.4) e (A.3.1).

Na notação matricial portanto,

$$VC \begin{pmatrix} x \\ \alpha \end{pmatrix} = \frac{\sigma_\alpha^2}{1 - \rho^2} \begin{pmatrix} \rho \\ \rho \end{pmatrix} + \sigma^2 \mathbf{1} \mathbf{1}' \quad (\text{A.3.5})$$

## APÊNDICE B

Cálculo das inversas das matrizes de variância e covariância dos dados originais.

### B.1 $\Sigma$ uniforme

A inversa da matriz de variância e covariância uniforme é obtida da expressão abaixo Mardia et al (1979) p.458.

$$(G + B C D)^{-1} = G^{-1} - G^{-1} B (C^{-1} + D G^{-1} B)^{-1} D G^{-1} \quad (B.1.1)$$

com,

$$G_{p \times p} = \sigma^2 I$$

$$B_{p \times n} = 1$$

$$C_{n \times n} = \sigma^2$$

$$D_{n \times p} = 1', \quad n = 1$$

então,

$$\Sigma^{-1} = \frac{1}{\sigma^2} I - \frac{1}{\sigma^2} I 1 \left( \frac{1}{\sigma^2} + 1' \frac{1}{\sigma^2} I 1 \right)^{-1} 1' \frac{1}{\sigma^2} I$$

$$\Sigma^{-1} = \frac{1}{\sigma^2} \left( I - \frac{\sigma^2}{\sigma^2 + p \sigma^2} 1 1' \right) \quad (B.1.2)$$

### B.2 $\Sigma$ seriada

A matriz de variância e covariância pode ser denotada por,

$$\Sigma = \frac{\sigma_u^2}{1 - \rho^2} C_\rho \quad , \quad | \rho | < 1 , \sigma_u^2 > 0$$

De Graybill (1983) p.201 , a inversa de

$$V = \sigma_u^2 C_\rho$$

$$V^{-1} = \frac{1}{\sigma_u^2 (1 - \rho^2)} F$$

onde

$$F = \begin{bmatrix} 1 & -\rho & 0 & \dots & 0 \\ -\rho & (1-\rho^2) & -\rho & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}$$

então,

$$\Sigma^{-1} = (1 - \rho^2) V^{-1}$$

ou seja,

$$\Sigma^{-1} = \frac{1}{\sigma_u^2} F \quad (B.2.1)$$

### B.3 $\Sigma$ seriada de Kenward

A matriz de variância e covariância de Kenward, na notação matricial,

$$\Sigma_e = \frac{\sigma_\alpha^2}{1 - \rho^2} C_\rho + \sigma^2 \mathbf{1} \mathbf{1}'$$

Por (B.1.1), tomando

$$G = \frac{\sigma_\alpha^2}{1 - \rho^2} C_\rho$$

$$B = \mathbf{1}$$

$$C = \sigma^2$$

$$D = \mathbf{1}'$$

segue

$$\Sigma_e^{-1} = \frac{1}{\sigma_\alpha^2} F - \frac{\sigma^2}{\sigma_\alpha^2} F \mathbf{1} \left[ \frac{1}{\sigma^2} + \mathbf{1}' \frac{1}{\sigma_\alpha^2} F \mathbf{1} \right]^{-1} \mathbf{1}' \frac{1}{\sigma_\alpha^2} F$$

$$\Sigma_e^{-1} = \frac{1}{\sigma_\alpha^2} F - w A \quad (B.3.1)$$

com  $F$  dada no apêndice B.2,

$$w = \frac{\sigma^2}{\sigma_\alpha^2 (\sigma_\alpha^2 + v \sigma^2)}, \quad v = 2(1-\rho) + (p-2)(1-\rho)^2$$

$$A = \begin{bmatrix} (1-\rho)^2 (1-\rho)^3 & \dots & (1-\rho)^2 \\ (1-\rho)^3 (1-\rho)^4 & \dots & (1-\rho)^3 \\ \dots & \dots & \dots \\ (1-\rho)^2 (1-\rho)^3 & \dots & (1-\rho) \end{bmatrix}$$

pois,

$$v = 1' F 1$$

$$F 1 = \begin{bmatrix} (1-\rho) & (1-\rho)^2 & \dots & (1-\rho)^2 (1-\rho) \end{bmatrix}$$

o

$$A = F 1 1' F$$



## APÊNDICE C

Matrizes de variância e covariância dos coeficientes ortonormais e suas inversas.

### C.1 $\Sigma$ Uniforme

A matriz de variância e covariância dos coeficientes ortonormais para a  $\Sigma$  uniforme.

$$\Sigma_u = P' \Sigma P$$

$$\Sigma_u = P' (\sigma^2 I + \sigma_\alpha^2 \mathbf{1} \mathbf{1}') P$$

$$\Sigma = \sigma^2 I + \sigma_\alpha^2 P' \mathbf{1} \mathbf{1}' P$$

$$\Sigma_u^{-1} = \begin{bmatrix} \sigma^2 + \sigma_\alpha^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & & 0 \\ 0 & 0 & \sigma^2 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix} \quad (C.1.1)$$

e a inversa dada por,

$$\Sigma_u^{-1} = \begin{bmatrix} \frac{1}{\sigma^2 + \sigma_\alpha^2} & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma^2} & 0 & & 0 \\ 0 & 0 & \frac{1}{\sigma^2} & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & & \frac{1}{\sigma^2} \end{bmatrix} \quad (\text{C.1.2})$$

### C.2 $\Sigma$ seriada

A matriz de variância e covariância dos coeficientes ortonormais

$$\Sigma_\beta = P' \Sigma P$$

$P$  é a matriz dos polinômios ortonormais.

A  $(i+1)$ -ésima linha de  $P' \Sigma$  é dada por.



$$P \cdot \Sigma_{(i+1)} = \text{cte} \begin{bmatrix} P_{i1} + \rho P_{i2} + \rho^2 P_{i3} + \dots + \rho^{p-1} P_{ip} \\ \rho P_{i1} + P_{i2} + \rho P_{i3} + \dots + \rho^{p-2} P_{ip} \\ \vdots \\ \rho^{p-1} P_{i1} + \rho^{p-2} P_{i2} + \rho^{p-3} P_{i3} + \dots + P_{ip} \end{bmatrix}$$

e o elemento  $(i+1, j+1)$  da matriz  $\Sigma_{\beta}$

$$\Sigma_{\beta} = \text{cte} \begin{bmatrix} P_{i1} P_{j1} + \rho P_{i2} P_{j1} + \rho^2 P_{i3} P_{j1} + \dots + \rho^{p-1} P_{ip} P_{j1} \\ + \rho P_{i1} P_{j2} + P_{i2} P_{j2} + \rho P_{i3} P_{j2} + \dots + \rho^{p-2} P_{ip} P_{j2} \\ + \rho^2 P_{i1} P_{j3} + \rho P_{i2} P_{j3} + P_{i3} P_{j3} + \dots + \rho^{p-3} P_{ip} P_{j3} \\ + \\ \vdots \\ + \rho^{p-1} P_{i1} P_{jp} + \rho^{p-2} P_{i2} P_{jp} + \rho^{p-3} P_{i3} P_{jp} + \dots + P_{ip} P_{jp} \end{bmatrix}$$

$$\Sigma_{\beta} = \text{cte} \begin{bmatrix} \sum_{k=1}^p P_{ik} P_{jk} + \sum_{k=2}^p \rho P_{ik} P_{j(k-1)} + \sum_{k=1}^{p-1} \rho P_{ik} P_{j(k+1)} + \\ \sum_{k=3}^p \rho^2 P_{ik} P_{j(k-2)} + \sum_{k=1}^{p-2} \rho^2 P_{ik} P_{j(k+2)} + \\ + \rho^{p-1} P_{ip} P_{j1} + \rho^{p-1} P_{i1} P_{jp} \end{bmatrix}$$

onde pela construção de P ,

$$\sum_{k=1}^p P_{ik} P_{jk} = 0$$

logo,

$$\begin{aligned} \Sigma_{\beta} &= \text{cte} \left[ \sum_{k=2}^p \rho^k P_{ik} P_{j(k-1)} + \sum_{k=1}^{p-1} \rho^k P_{ik} P_{j(k+1)} + \right. \\ &\quad \left. \sum_{k=2}^p \rho^{2k} P_{ik} P_{j(k-2)} + \sum_{k=1}^{p-2} \rho^{2k} P_{ik} P_{j(k+2)} + \right. \\ &\quad \left. + \rho^{p-1} P_{ip} P_{j1} + \rho^{p-1} P_{i1} P_{jp} \right] \quad (\text{C. 2.1}) \end{aligned}$$

e o elemento da diagonal, quando  $i=j$ ,

$$\begin{aligned} \Sigma_{\beta} &= \text{cte} \left[ \sum_{k=2}^p \rho^{2k} P_{ik} P_{ik} + \sum_{k=1}^{p-1} \rho^k P_{ik} P_{j(k+1)} + \right. \\ &\quad \left. \sum_{k=2}^p \rho^{2k} P_{ik} P_{ik} + \sum_{k=1}^{p-2} \rho^{2k} P_{ik} P_{i(k+2)} + \right. \\ &\quad \left. + \rho^{p-1} P_{ip} P_{i1} + \rho^{p-1} P_{i1} P_{ip} \right] \quad (\text{C. 2.2}) \end{aligned}$$

com,

$$\text{cte} = \frac{\sigma_u^2}{1 - \rho^2}$$

e a inversa é .

$$\Sigma_{\beta}^{-1} = P' \Sigma^{-1} P$$

onde  $P$  é a matriz dos polinômios ortonormais.

A  $(i+1)$ -ésima linha de  $P' \Sigma^{-1}$  por (B.2.1) é dada por

$$P' \Sigma_{(i+1)}^{-1} = \begin{bmatrix} P_{i1} - \rho P_{i2} \\ -\rho P_{i1} + (1+\rho^2) P_{i2} - \rho P_{i3} \\ \vdots \\ -\rho P_{i(k-1)} + (1+\rho^2) P_{ik} - \rho P_{i(k+1)} \\ \vdots \\ P_{ip} - \rho P_{i(p-1)} \end{bmatrix},$$

e o elemento  $(i+1, j+1)$  da matriz  $\sigma_u^2 \Sigma_{\beta}^{-1}$  é portanto,

$$\sigma_u^2 \Sigma_{(i+1, j+1)}^{-1} = (P_{i1} - \rho P_{i2}) P_{j1} + \left[ (1+\rho^2) P_{i2} - \rho P_{i1} - \rho P_{i3} \right] P_{j2} + \dots + (P_{ip} - \rho P_{i(p-1)}) P_{jp}$$

$$= \sum_{k=1}^p P_{ik} P_{jk} + \rho^2 \sum_{k=2}^{p-1} P_{ik} P_{jk} - \rho \sum_{k=2}^p P_{ik} P_{jk-1} - \rho \sum_{k=1}^{p-1} P_{ik} P_{jk+1}$$

pela construção de  $P$ ,

$$\sum_{k=1}^p P_{ik} P_{jk} = 0$$

logo,

$$\Sigma_{\beta}^{-1} = \frac{1}{\sigma_u^2} \left[ \rho^2 \sum_{k=2}^{p-1} P_{ik} P_{jk} - \rho \sum_{k=2}^p P_{ik} P_{jk-1} - \rho \sum_{k=1}^{p-1} P_{ik} P_{jk+1} \right] \quad (C. 2. 3)$$

e o elemento da diagonal, quando  $i=j$

$$\Sigma_{\beta}^{-1} = \frac{1}{\sigma_u^2} \left[ \sum_{k=1}^p P_{ik}^2 + \rho^2 \sum_{k=2}^{p-1} P_{ik}^2 - 2\rho \sum_{k=1}^{p-1} P_{ik} P_{ik+1} \right] \quad (C. 2. 4)$$

### C.3 $\Sigma$ seriada de Kenward

O cálculo da matriz de variância e covariância dos coeficientes ortonormais da  $\Sigma$  seriada apresentada por Kenward.

De (3.1.11) a matriz de variância e covariância dos coeficientes ortonormais para o caso da seriada apresentada por

Kenward (1985) e,

$$\begin{aligned} \Sigma_{\beta} &= P' \Sigma_{\epsilon} P \\ &= P' \left[ \frac{\sigma_{\alpha}^2}{1-\rho^2} C_{\rho} + \sigma^2 \mathbf{1} \mathbf{1}' \right] P \\ &= \frac{\sigma_{\alpha}^2}{1-\rho^2} P' C_{\rho} P + \sigma^2 P' \mathbf{1} \mathbf{1}' P \end{aligned}$$

⇒

$$\frac{\sigma_{\alpha}^2}{1-\rho^2} P' C_{\rho} P$$

calculada em (C.2.3) e (C.2.4), com  $\sigma_u^2$

onde pela construção de  $P$ , Fisher e Yates (1963), temos que,

$$P' \mathbf{1} \mathbf{1}' P = \begin{bmatrix} 1 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \end{bmatrix}$$

é óbvio que a variabilidade "entre" indivíduos interfere apenas com a variância do termo constante.

$$\Sigma_{\beta_0}^{-1} = P' \Sigma_u^{-1} P$$

do (B.3.1) , temos

$$\Sigma^{-1} = \frac{1}{\sigma_u^2} [F - w A]$$

então

$$\Sigma_{\beta_0}^{-1} = \frac{1}{\sigma_u^2} P' F P - w P' A P \quad (C.2.1)$$

e

$$\frac{1}{\sigma_u^2} P' F P$$

calculado em (C.2.3) e (C.2.4) , com  $\sigma_u^2$

Para obter a segunda parcela de (C.2.1), sendo

$$P' A_{(i+1)} = \begin{bmatrix} (1-\rho)^2 (P_{i1} + P_{ip}) + (1-\rho)^3 \sum_{k=2}^{p-1} P_{ik} \\ (1-\rho)^2 (P_{i1} + P_{ip}) + (1-\rho)^4 \sum_{k=2}^{p-1} P_{ik} \\ \vdots \\ (1-\rho)^2 (P_{i1} + P_{ip}) + (1-\rho)^3 \sum_{k=2}^{p-1} P_{ik} \end{bmatrix}$$

a i-esima linha de  $P' A$  , o elemento  $(i)$  do produto  $P' A P$  e dado

por,

$$\begin{aligned}
 P'AP_{(i+1,j+1)} = & (1-\rho)^2 \left( 1 - \sum_{k=2}^{p-1} P_{ik} \right) P_{j1} + (1-\rho)^3 \sum_{k=2}^{p-1} P_{ik} P_{j1} + \\
 & + (1-\rho)^3 \left( 1 - \sum_{k=2}^{p-1} P_{ik} \right) P_{j2} + (1-\rho)^4 \sum_{k=2}^{p-1} P_{ik} P_{j2} + \\
 & \vdots \\
 & + (1-\rho)^2 \left( 1 - \sum_{k=2}^{p-1} P_{ik} \right) P_{jp} + (1-\rho)^3 \sum_{k=2}^{p-1} P_{ik} P_{jp}
 \end{aligned}$$

$$\begin{aligned}
 P'AP_{(i+1,j+1)} = & \left[ \begin{array}{l} (1-\rho)^2 + \rho(1-\rho)^2 \sum_{l=2}^{p-1} P_{jl} - \rho(1-\rho)^2 \sum_{k=2}^{p-1} P_{ik} \\ \rho^2(1-\rho)^2 \sum_{l=2}^{p-1} P_{jl} \sum_{k=2}^{p-1} P_{ik} \end{array} \right] \quad (C.3.1)
 \end{aligned}$$

e o elemento da diagonal,

$$P'AP_{(i+1,i+1)} = \left[ \begin{array}{l} (1-\rho)^2 + 2\rho(1-\rho)^2 \sum_{k=2}^{p-1} P_{ik} - \rho^2(1-\rho)^2 \left[ \sum_{k=2}^{p-1} P_{ik} \right]^2 \end{array} \right] \quad (C.3.2)$$

## APÊNDICE D

### D.1 Geração dos perfis individuais.

A geração de vetores normais é feita utilizando o gerador RANNOR implementado no STATISTICAL ANALYSIS SYSTEM (SAS) no computador VAX/UNICAMP.

Na simulação foram geradas duas matrizes normais  $Z^{(1)}$  e  $Z^{(2)}$  de variáveis independentes, com média zero e matriz de variância e covariância identidade.

Para gerar variáveis, com vetor de médias diferente de zero e a matriz de variância e covariância com a estrutura de correlação foi utilizada a seguinte transformação.

$$\underset{\sim}{x}_{j\alpha} = \underset{\sim}{\Sigma}^{1/2} \underset{\sim}{z}_{j\alpha} + \underset{\sim}{\mu}_j$$

onde

$$\underset{\sim}{x}_{j\alpha} \sim N_p(\underset{\sim}{\mu}_j, \underset{\sim}{\Sigma})$$

$\underset{\sim}{\mu}_j$  é o vetor de médias do j-ésimo grupo

$\underset{\sim}{\Sigma}^{1/2}$  é raiz quadrada das matrizes de variância e covariância, obtida utilizando o fator de Cholesky, para cada caso descrito na tabela 4.3.1.



$$\Sigma^{1/2} = \text{ROOT}(\Sigma)$$

criando-se assim as matrizes de dados

$$X^{(j)} = (x_{\alpha j 1}, x_{\alpha j 2}, \dots, x_{\alpha j p}) \quad j = 1, 2$$

onde  $x_{\alpha j \alpha}$  é o vetor de observações contendo as  $p$  medidas da mesma variável do  $\alpha$ -ésimo indivíduo no  $j$ -ésimo grupo.

C Programa da transformação dos dados segundo o vetor de médias e a matriz de variância e covariância.

C Arquivo de dados de entrada das normais (0,1)

data dados;

infile 'b:nome do arquivo de dados para o primer grupo';

input z1 z2 z3 z4;

infile 'b:nome do arquivo de dados para o segundo grupo';

input z5 z6 z7 z8;

C Transformação dos dados proc iml;

use dados;

read all var(z1) into z1;

read all var(z2) into z2;

read all var(z3) into z3;

read all var(z4) into z4;

read all var(z5) into z5;

read all var(z6) into z6;

```

read all var(z7) into z7;
read all var(z8) into z8;

C      Entrada do vetor de médias para o primer grupo .
medi =(<          >);

C      Entrada do vetor de médias para o segundo grupo.
med2 =(<          >);

C      Entrada da matriz de variância e covariância segundo o
      caso considerado
var =(<          ,
          ,
          ,
          >);

C      Indicador do grupo.
ig1=1;
ig2=2;
grup1=repeat(ig1,10,1);
grup2=repeat(ig2,10,1);

z=z1||z2||z3||z4;
w=z5||z6||z7||z8;
per1=med1//med1//med1//med1//med1//med1//med1//med1//med1//med1;

C      Transformação dos dados utilizando o fator de Cholesky
      para obter  $\Sigma^{1/2}$ .
u=root(var);
v10=u'*u;
print v10;
per2=med2//med2//med2//med2//med2//med2//med2//med2//med2//med2;
x11=u'*z';
x12=u'*w';
x1=x11+per1';
x2=x12+per2';
x=x1' || grup1' || x2' || grup2';

```

C            Saída dos dados para os dos grupos.

```
create out from x;  
append from x;  
proc print data = out ;  
run;
```

## D.2 Gráfico dos perfis individuais

C Programa dos gráficos dos perfis para os dois grupos.

C Entrada dos arquivos de dados para os dois grupos.

```
cms filedef dados1 disk grap51 dat a1 ;
cms filedef dados2 disk gras51 dat a1 ;
data dados;
infile dados1;
input oc i1 i2 i3 i4 i5 i6 i7 i8 i9 i10;
infile dados2;
input oc j1 j2 j3 j4 j5 j6 j7 j8 j9 j10;
title 'curva media dos grupos por ocasiao';
goptions dev=grlink;
```

```
data labels1(keep= x y xsys ysys position);
set dados;
```

```
if oc=4 then do;
  xsys='2'; ysys='2';
  position='3';
  style='none';
  x=oc; y=i1; output;
  y=i2; output;
  y=i3; output;
  y=i4; output;
  y=i5; output;
  y=i6; output;
  y=i7; output;
  y=i8; output;
  y=i9; output;
  y=i10; output;
  y=j1; output;
  y=j2; output;
  y=j3; output;
  y=j3; output;
  y=j4; output;
  y=j5; output;
  y=j6; output;
  y=j7; output;
  y=j8; output;
  y=j9; output;
  y=j10; output;
```

```
end;
run;
```

```

proc gplot data =dados ;
  plot i1*oc=1   i2*oc=1 i3*oc=1 i4*oc=1 i5*oc=1
       i6*oc=1   i7*oc=1 i8*oc=1 i9*oc=1 i10*oc=1
       j1*oc=2   j2*oc=2 j3*oc=2 j4*oc=2 j5*oc=2
       j6*oc=2   j7*oc=2 j8*oc=2 j9*oc=2 j10*oc=2 /overlay
  vaxis =0 to i4 by 2
  haxis = 1 to 4 by 1
  annotate = labels1;
label  oc='ocasio'
       i1='perfis individuais';

symbol1 w=1 i=join l=1 ;
symbol2 w=1 i=join l=20 ;

title 'perfis individuais dos grupos por ocasio';
title  ' caso uniforme fixada variabilidade entre';
footnote ' correlacao 0.1';

run;

```

## BIBLIOGRAFIA

- BOCK, R. D. (1975). *Multivariate Statistical Methods in Behavioral Research*. Mc Graw-Hill New York.
- BOX, G. E. (1950). Problems in the analysis of growth and wear curves. *Biometrics* 6, 362 -389.
- COCHRAN, W.; BLISS, C. (1948). Discriminant functions with covariance. *The Annals of Mathematical Statistics* ,151-176.
- COLE, J. ; GRIZZLE, J. (1966) . Applications of multivariate analysis to repeated measurements experiments. *Biometrics* 22, 810-828.
- CSORGO, S. (1986). Testing for normality in arbitrary dimension *The Annals of Statistics* 14(2), 708-723.
- FISHER, R. A. ; YATES, F. (1963) *Statistical tables for Biological, Agricultural, and Medical Research*, Sixth Ed. , Hafner New York.
- GEISSER, S. (1963). Multivariate analysis of variance for a special covariance case. *Journal of the American Statistical Association* 58, 660-669.
- GRAYBILL, F.A. (1983) *Matrices with Applications in Statistical*. Wadsworth International Group Belmont , A Division of Wadsworth, Inc. California.



- GRIZZLE, J. ; ALLEN, D. (1969). Analysis of growth and dose response curves . *Biometrics* 25, 357-382.
- HEARNE, E.M. ; CLARK, G. M.; HATCH, J. P. (1983) . A test for serial correlation in univariate repeated-measures analysis. *Biometrics* 39, 237-243.
- HUYNH, H .; FELDT, L.S. (1970) Conditions under which mean squares ratios in repeated measurements designs have exact F-distributions, *Journal of the American Statistical Association* 65, 1582-1589.
- INFANTE, A. (1989). The difficulties of estimation of dispersion parameters in linear models - An Illustration. Submetido a publicação.
- KENWARD, M. G. (1985). The use of fitted higher-order polynomial coefficients as covariates in the analysis of growth curves. *Biometrics* 41, 19-28
- KENWARD, M. G. (1986) . The distribution of a generalized least squares estimator with covariance adjustment. *Journal of Multivariate Analysis* 20, 244-250.
- LESLIE, J. R.; SEPHERS, H.A.; FOTOPOULOS, S. (1986). Asymptotic distribution of the Shapiro-Wilks W for testing for normality . *The Annals of Statistics* 14, 1497-1506.
- MARDIA, K. ; KENT, J, ; BIBBY, J. (1979) . *Multivariate Analysis*. Academic Press. London.
- MASON, R.L. ; YOUN, J.C. (1985). Re-examining two tests



for bivariate normality . *Communications Statistical Theory and Methods* , A 14 (7) 1531-1545.

MANSOUR, H. ; NORDHEIM, E. V. ; RUTLEDGE, J. J. (1985). Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. *Biometrics* 41, 287-294.

PLEWIS, I. (1986) . Analysis data from longitudinal comparative studies In : *New Developments in Statistical for Psychology and the Social Sciences* . Lovles A. D. (Ed.) London, New York.

POTTHOF, R. ; ROY, S. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* 51, 313-326.

RAO, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves . *Biometrics* 52, 447-458.

RAO, C. R. (1966) . Covariance adjustment and related problems in multivariate analysis . In : *Multivariate Analysis II* . P. R. Krishnaiah (ed): Academic Press . New York, 87-103.

RAO, C. R (1967). Least squares theory using an estimated dispersion matrix and its application to measurements of signals . *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics* 1, 355-372.

SINGER, J. M. ; ANDRADE ; D. F. (1986) . Análise de dados longitudinais . VII - SINAPE , Simpósio Nacional de



Probabilidade e Estatística . Campinas , Sao Paulo.

WISHART, J. (1938). Growth rate determination in nutrition studies with the bacon pig and their analysis . *Biometrika* 30, 16-28.