ALGUNS MODELOS LINEARES

MULTIVARIADOS COM ESTRUTURA

DE COVARIÂNCIA

IBERÊ GUIMARĀES AGUIAR

DISSERTAÇÃO APRESENTADA

AO

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

DA

UNIVERSIDADE DE SÃO PAULO

PARA OBTENÇÃO DO GRAU DE MESTRE

EM

ESTATISTICA

ORIENTADOR:

Prof. Dr. Pedro Alberto Morettin

- SÃO PAULO, JULHO DE 1979 -

Em memória de meu pai

Para minha mãe, minha esposa e meus filhos.

AGRADECIMENTOS

Ao meu orientador e amigo, professor Pedro Alberto Morettin pela dedicada ajuda e pelo grande incentivo;

Ao professor, Adolpho Walter Pimazoni Canton, pelas sugestões apresentadas;

Aos meus professores

Carlos Alberto de Bragança Pereira, Flávio Wagner Rodrigues, Lindo Fava e Pedro Alberto Morettin,

pelos conhecimentos que me proporcionaram, permitindo a realização deste trabalho;

Ao professor Wilton de Oliveira Bussab, pela maneira cortês como atendeu às minhas consultas;

Aos colegas, professores, funcionários e ao pessoal da biblioteca do I.M.E, pelas atenções dispensadas.

PREFACIO

Este trabalho tem por objetivo, fazer comparações entre alguns modelos lineares, sob o aspecto da estrutura de covariancia, envolvendo dois tipos de variáveis: variáveis observadas que são denominadas de variáveis respostas ou indicadores observados; e as variáveis não observadas que são cha madas de fatores, variáveis latentes ou variáveis de interesse. Estas variáveis latentes são ricamente encontradas nas Ciências do Comportamento. Apresentamos os modelos mate máticos, determinamos os estimadores pelo processo de Máxima Verossimilhança e testamos hipóteses.

Obedece à seguinte sequência:

- 01) no primeiro capítulo, apresentamos a parte matemática utilizada;
- 02) no segundo capítulo, estudamos alguns modelos lineares multivariados;
- 03) no terceiro capítulo, expomos o modelo linear latente multivariado;
- O4) no quarto capítulo, fazemos o estudo do modelo geral de Yoreskog, como caso geral das anteriores;
- 05) e no quinto capítulo, apresentamos três exemplos.

INDICE

| CAPITULO 1: MATRIZES E DERIVADAS DE MATRIZES | |
|--|----|
| 1.1 - Algumas definições sobre Matrizes | 1 |
| 1.2 - Produto Direto e Estrela | |
| 1.3 - Algumas Matrizes Importantes | |
| 1.4 - Derivação de Função de Vetores | 5 |
| 1.5 - Derivada de uma matriz em relação a outra matriz | 7 |
| 1.6 - Derivadas de Funções de Matrizes | 8 |
| 1.7 - Da Distribuição Normal | 11 |
| 1.8 - Alguns Tipos de Convergência | 14 |
| CAPITULO 2: MODELOS LINEARES MULTIVARIADOS | |
| 2.1 - Análise Fatorial | 16 |
| 2.2 - Análise de Componentes Principais | 28 |
| 2.3 - Análise de Trajetórias | 32 |
| 2.4 - Análise de Curvas de Crescimento | 38 |
| CAPITULO 3: MODELO LINEAR LATENTE MULTIVARIADO | |
| 3.1 - Introdução | 49 |
| 3.2 - Apresentação do Modelo Linear | 49 |
| 3.3 - Modelo Linear Latente Multivariado | 50 |
| 3.4 - Restrições nos Parâmetros Estruturais | 51 |
| 3.5 - Estimação de Máxima Verossimilhança | 54 |

| | iv |
|---|-------|
| 3.6 - Consistência e Distribuição Assintótica | 57 |
| 3.7 - Obtenção da Matriz Informação Limite de Fisher | . 68 |
| 3.8 - Testes de Hipóteses | . 69 |
| CAPITULO 4: MODELO GERAL DE ESTRUTURA DE COVARIÂNCIAS | |
| 4.1 - Introdução | . 82 |
| 4.2 - Modelo Geral | 82 |
| 4.3 - Identificação do Modelo | 83 |
| 4.4 - Estimação | . 84 |
| 4.5 - Testes de Hipóteses sôbre E | 89 |
| 4.6 - Casos Particulares do Modelo Geral | 91 |
| CAPITULO 5: EXEMPLOS | |
| 5.1 - Exemplo Numérico | 102 |
| 5.2 - Um exemplo de Aplicação do Modelo Geral | 116 |
| 5.3 - Um exemplo em Análise Fatorial | .122 |
| <u>APĒNDICE</u> | |
| - Método de minimização de Fletcher e Powell | 125 |
| DEFEDÊNCIAS | . 128 |

CAPITULO 1

MATRIZES E DERIVADAS DE MATRIZES E ALGUMAS PROPRIEDADES DE CONVERGÊNCIA

1.1 - ALGUMAS DEFINIÇÕES SOBRE MATRIZES

1.1.1 - Uma matriz (mxn) e um conjunto de m.n números ordenados, dispostos em m linhas e n colunas: $\Lambda = \left[a_{ij}\right]$.

1.1.2 - A transposta da matriz A é a matriz A' de or dem nxm, tal que A' = $\begin{bmatrix} a_{ji} \end{bmatrix}$.

1.1.3 - O produto da matriz A(mxn) pela matriz B(nxp) é a matriz (mxp) dada por

$$A.B = \begin{bmatrix} a_1 & b_1 & a_1 & b_2 & \cdots & a_1 & b_p \\ a_2 & b_1 & a_2 & b_2 & \cdots & a_1 & b_p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_m & b_1 & a_m & b_2 & \cdots & a_m & b_p \end{bmatrix},$$

onde os a_j , j=1,2,... m são vetores linhas de A c os b_k , k=1,2...p são os vetores colunas de B e a_j b_k = $\sum_{r=1}^{n} a_{jr} b_{rk}$ (produto escalar do j-ésimo vetor linha de A pelo k-ésimo vetor coluna de B).

1.1.4 - O traço de uma matriz X(nxn) é a soma dos elementos da diagonal principal: $tr\left(X\right) = \sum_{i=1}^{n} x_{ii}$

- 1.1.5 Uma matriz Z (mxm) \tilde{e} dita ortogonal se $Z'Z' = Z Z' = I_m$.
- 1.1.6 Uma matriz quadrada Z (mxm) é dita simétrica se Z' = Z, ou seja, Z_{ij} = Z_{ji} ¥ i e j
- 1.1.7 Temos que, se A, B e C são matrizes de ordens apropriadas,
 - i) (A + B)' = A' + B'
 - ii) (A.B.C)' = C'.B'A'
- 1.1.8 A inversa da matriz A (nxn) é a matriz B (nxn) que satisfaz a condição A.B = B.A = I_m e escrevemos $B = A^{-1}$
- 1.1.9 Para A , B e C de ordens apropriadas, segue-se que $(A B C)^{-1} = C^{-1} B^{-1} A^{-1}$
- 1.1.10 Uma matriz A(pxp) e positiva definida (p.d)
 se X' A X > 0 \ \ X (pxl) \ \ \ 0
- 1.1.11 O determinante de A será indicado por |A| ou det A
- 1.1.12 Se Λ ē (p.d), escrevemos Λ > 0; então $|\Lambda|$ > 0 e todos os determinantes principais de ordem inferior são positivos.
 - 1.1.13 Se $\Lambda > 0$, então $\Lambda^{-1} > 0$

1.2 - PRODUTO DIRETO E ESTRELA

1.2.1 - O produto direto (de Kroneckor) de uma matriz A(mxn) pela matriz B(pxq) é uma matriz (m.p x n.q) representada por

Propriedades: Se A(pxp) e B(nxn), então

a)
$$(\underline{A} \boxtimes \underline{B})(\underline{C} \boxtimes \underline{D}) = \underline{A} \subseteq \underline{\otimes} \underline{B} \underline{D}, \underline{C}(pxp) \in \underline{D}(nxn)$$

b)
$$(\underline{A} \otimes \underline{B})^{-1} = \underline{A}^{-1} \otimes \underline{B}^{-1}$$

c)
$$|\underline{A} \otimes \underline{B}| = |\underline{A}|^n \cdot |\underline{B}|^p$$

d)
$$(\underline{A} \otimes \underline{B})' = \underline{A}' \otimes \underline{B}'$$

e)
$$\operatorname{tr}(A \otimes B) = \operatorname{tr} A \cdot \operatorname{tr} B$$

1.2.2 - O produto estrela

Sejam Λ e B duas matrizes, (pxq) e (pmxqn), respectivamente. A matriz B \tilde{e} particionada em p.q blocos B_{ij} (mxn). Então o produto estrela de Λ por B \tilde{e} $\Lambda * B = \sum_{ij} a_{ij} B_{ij}$. Observação: se Λ e B têm a mesma dimensão, então $\Lambda * B = \operatorname{tr} \left(\Lambda' B\right)$.

1.3 - ALGUMAS MATRIZES IMPORTANTES

1.3.1 - Matriz E_{ij}

A matriz E ij e a matriz (mxn) que contem o número 1 na posição da i-esima linha com a j-esima coluna e zeros nas de mais posições.

1.3.2 - A matriz E(m,n) A matriz E(m,n) e a matriz (m²xn²) definida por

$$\mathbf{E}_{(m,n)} = \begin{bmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} & \cdots & \mathbf{E}_{1n} \\ \mathbf{E}_{21} & \mathbf{E}_{22} & \cdots & \mathbf{E}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}_{m1} & \mathbf{E}_{m2} & \cdots & \mathbf{E}_{mn} \end{bmatrix}$$

1.3.3 - A matriz $I_{(m,n)}$

A matriz I(m,n) é uma matriz (mn x mn) definida por

$$\underline{I}_{(n,n)} = \begin{bmatrix} \underline{E}_{11}^{i} & \underline{E}_{12}^{i} & \cdots & \underline{E}_{1n}^{i} \\ \underline{E}_{21}^{i} & \underline{E}_{22}^{i} & \cdots & \underline{E}_{2n}^{i} \\ \vdots & \vdots & \vdots & \vdots \\ \underline{E}_{n1}^{i} & \underline{E}_{n2}^{i} & \underline{E}_{mn}^{i} \end{bmatrix}$$

Lema (1.1): Sejam A(mxp), B(pxq) e C(qxn) três matrizes. En

$$(\underline{A} \underline{B} \underline{C})^{\bullet} = \underline{B}^{\bullet} * \{ (\underline{C} \otimes \underline{I}_{n}) \underline{E}_{(n,m)} (\underline{A} \otimes \underline{I}_{m}) \}.$$

Lema (1.2): Sejam A(mxp), B(pxq) e C(qxn) três natrizes. En

$$(\Lambda \cdot B \cdot C) = B' * \{ (C \otimes I_n) I_{(m,n)} (\Lambda \otimes I_n) \}.$$

1.4 - DERIVAÇÃO DE FUNÇÕES DE VETORES

1.4.1 - Seja $f(\underline{x}) = f(x_1, x_2, \dots, x_p)$ uma função continua em relação às coordenadas do vetor $\underline{X}' = (x_1, x_2, \dots, x_p);$ suponha que as derivadas $\frac{\partial f(\underline{X})}{\partial x_i}$, $\frac{\partial^2 f(\underline{X})}{\partial x_i \partial X_j}$ existam numa cer ta região $R \subset \mathbb{R}^p$. Então:

$$\left[\begin{array}{c} \frac{\partial f(X)}{\partial X} \end{array}\right] = \left[\begin{array}{c} \frac{\partial f(X)}{\partial x_1}, & \frac{\partial f(X)}{\partial x_2} & \cdots & \frac{\partial f(X)}{\partial x_p} \end{array}\right]$$

Teorema (1.1): Sejam $X' = (x_1, x_2, \dots x_p), \underline{a}' = (a_1, a_2, \dots, a_p)$ $e f(X) = X'\underline{a}, X'\underline{a} = \sum_{k=1}^{p} a_k x_k; \text{ então } \frac{\partial f(X)}{\partial X} = \underline{a}.$

Teorema (1.2): Se $f(X) = X' \wedge X$, $X(px1) \in A$ a matriz (pxp) simétrica, então $\frac{\partial f(X)}{\partial X} = 2 \wedge X$.

1.4.2 - Matriz Hessiana

A matriz de derivadas de segunda ordem de uma função de p variáveis f(X) é chamada matriz Hessiana:

$$\underline{H} = \frac{\partial^2 f(\underline{X})}{\partial \underline{X}' \partial \underline{X}} = \left(\frac{\partial^2 f(\underline{X})}{\partial x_i \partial x_j} \right), \quad i, j = 1, 2, 3 \cdots p$$

1.4.3 - Consideremos, agora, derivadas de funções esca lares de matrizes; definimos

$$\frac{\partial f(X)}{\partial X} = \begin{bmatrix} \partial f(X) \\ \partial x_{ij} \end{bmatrix}, \text{ onde } f \in \text{uma função escalar da}$$

matriz X(mxn)

Teorema (1.3): Seja $\underline{Y} = \{y_{ij}\}$ uma matriz (nxn) e $f(\underline{Y}) = |\underline{Y}|$ então $\frac{\partial f(\underline{Y})}{\partial y_{ij}} = Y_{ij}$, onde Y_{ij} é o cofator do elemento y_{ij}

 $\frac{\text{Teorema (1.4)}}{\text{3trY}} = \underline{I}_n$ Seja Y = $\left(y_{ij}\right)$ uma matriz (nxn), então

Teorema (1.5): Se $\underline{Y} = \begin{bmatrix} y_{ij} \end{bmatrix}$ ē uma matriz (nxn), então $\frac{\partial \log(|\underline{Y}|)}{\partial \underline{Y}} = \underline{Y}^{-1}, \underline{Y} \text{ não singular.}$

Dem.: Aplicação do teorema (1.3).

1.5 - DERIVADA DE UMA MATRIZ EM RELAÇÃO A OUTRA MATRIZ

1.5.1 - Seja X = $\begin{bmatrix} x_{ij} \end{bmatrix}$ uma matriz mxn. Chamamos a matriz $\frac{\partial}{\partial X} = \begin{bmatrix} \frac{\partial}{\partial x_{ij}} \end{bmatrix}$ a matriz dos operadores derivadas parciais, de ordem mxn.

1.5.2 - Seja Y(pxq) uma matriz onde seus elementos y_{ij} estão em função da matriz X(mxn). Então definimos a derivada de Y em relação a X por $\frac{\partial Y}{\partial X} = Y \otimes \frac{\partial}{\partial X}$ (pm x qn), ou seja

$$\frac{\partial Y}{\partial X} = \begin{bmatrix}
\frac{\partial y_{11}}{\partial X} & \frac{\partial y_{12}}{\partial X} & \cdots & \frac{\partial y_{1q}}{\partial X} \\
\vdots & \vdots & \vdots & \vdots \\
\frac{\partial y_{p1}}{\partial X} & \frac{\partial y_{p2}}{\partial X} & \frac{\partial y_{pq}}{\partial X}
\end{bmatrix}$$

Teorema (1.6): (Regra da cadeia para produto estrela). Seja f uma função escalar de uma matriz Y(pxq) cujos elementos. são funções de uma matriz X(mxn). Então:

$$\frac{\partial f}{\partial X} = \frac{\partial f}{\partial Y} * \frac{\partial Y}{\partial X}$$

Dem.: A demonstração é consequência direta da regra da cadeia ordinária e da definição de produto estrela.

Teorema (1.7): Sejam U e W matrizes (mxn) tais que seus ele mentos u ij e w ks são funções da matriz X (pxq). Então

$$\frac{9 \ \tilde{X}}{9 \left(\tilde{\Pi} + \tilde{M} \right)} = \frac{9 \ \tilde{X}}{9 \ \tilde{\Pi}} + \frac{9 \ \tilde{X}}{9 \ \tilde{M}}$$

Dem.: Ver de Waal, 1975

Teorema (1.8): Sejam U(pxq) e W(qxp), duas matrizes cujos elementos uij e wks são funções da matriz X(mxn). Então

$$\frac{9 \cdot \bar{X}}{9 \left(\bar{n} \cdot \bar{M} \right)} = \frac{9 \bar{X}}{9 \bar{n}} \left(\bar{M} \otimes \bar{I}^{u} \right) + \left(\bar{n} \otimes \bar{I}^{u} \right) \frac{9 \bar{X}}{9 \bar{M}}$$

Dem.: Ver Vetter, 1970

Teorema (1.9): (Derivada de matriz inversa) Seja Y(pxp) ura matriz não singular que é função de X(mxn). Então

$$\frac{3\bar{X}}{3\bar{X}_{-1}} = -\left(\bar{X}_{-1} \otimes \bar{I}^{m}\right) \frac{3\bar{X}}{3\bar{X}} \left(\bar{X}_{-1} \otimes \bar{I}^{n}\right)$$

Dem.: Aplicando o teorema (1.8) em \underline{Y}^{-1} $\underline{Y} = \underline{I}_p$ e usando a propriedade b de 1.2.1, obtemos o resultado desejado.

1.6 - DERIVADAS DE FUNÇÕES DE MATRIZES

Seja
$$V(pxp)$$
, $T(pxp)$ e $F = log(|V|) + tr(V^{-1}T)$ (1.1)

Λ) Supomos que Λ(pxq) e a matriz diagonal Ψ(pxp) sejam matrizes de parâmetros.

Teorema (1.10): Se $V = \Lambda \Lambda' + V \text{ com } V \text{ diagonal e } T$, matriz de valores constantes, então:

1)
$$\frac{\partial F}{\partial \Lambda} = 2V^{-1}(V - I) V^{-1} \Lambda$$

2)
$$\frac{\partial F}{\partial \Psi} = \text{diag} \left(\underline{V}^{-1} (\underline{V} - \underline{I}) \underline{V}^{-1} \right)$$

Dem.: Ver Lawley e Maxwell, 1971

B) Com a mesma expressão (1.1) supomos que V(pxp), T(pxp), E (qxr) sejam matrizes paramétricas e A(rxn) uma matriz esca lar, onde

$$V = \Lambda \Phi \Lambda' + \Psi \qquad (1.2)$$

$$\underline{T} = \frac{1}{n} \left(\underline{X} - \underline{\beta} \underline{A} \right) \left(\underline{X} - \underline{\beta} \underline{A} \right)$$
 (1.3)

$$\beta = \Lambda \Xi$$
 (1.4)

Teorema (1.11): As derivadas da função (1.1) relativamente aos parâmetros β e V são dadas por

1)
$$\frac{\partial F}{\partial \beta} = -\frac{2}{n} \quad \underline{V}^{-1}(\underline{X} - \underline{\beta} \underline{\Lambda}) \Lambda'$$
 (1.5)

2)
$$\frac{\partial F}{\partial V} = V^{-1}(V - T) V^{-1}$$
 (1.6)

Dem.: A demonstração deste teorema é feita com auxílio dos teoremas (1.6), (1.8) e dos lemas (1.1) e (1.2)

C) Sejam, agora, A (nxg), P (hxp) matrizes numéricas;

Ξ(gxh), B(pxq), Λ(qxr), Φ(rxr)e as matrizes diagonais Ψ(qxq)
 e (pxp) matrizes de parâmetros desconhecidos.

Supomos que

$$\underline{V} = \underline{B} \left[\underbrace{\Lambda} \Phi \underline{\Lambda}^{*} + \underline{\Psi} \right] \underline{B} + \underbrace{\Theta}$$
 (1.7)

$$\underline{\mathbf{T}} = \frac{1}{n} \left(\begin{array}{ccc} \mathbf{X} - \mathbf{A} & \mathbf{E} & \mathbf{P} \end{array} \right)^{\mathsf{T}} \left(\begin{array}{ccc} \mathbf{X} - \mathbf{A} & \mathbf{E} & \mathbf{P} \end{array} \right) \tag{1.8}$$

Seja F uma função das matrizes A , P , E , B , A O , H e X , tendo a forma (1.1). Então

Teorema (1.12):
$$\frac{\partial F}{\partial E} = \frac{-2}{n} A' \left(X - A E P \right) V^{-1} P$$

Dem.: Aplicação dos teoremas (1.6), (1.8) e do lema (1.1)

Teorema (1.13):
$$\frac{\partial F}{\partial B} = 2\Omega B(\Lambda \Phi \Lambda' + \Psi),$$

$$\Omega = V^{-1}(V - T)V^{-1}$$

Dem.: A demonstração deste teorema é feita usando-se os teoremas (1.5), (1.6), (1.7), (1.8) e os lemas (1.1) e (1.2)

Teorema (1.14):
$$\frac{\partial F}{\partial \Lambda} = 2 B' \Omega B \Lambda \Phi$$
,
$$\Omega = \underline{V}^{-1} \left(\underline{V} - \underline{T} \right) \underline{V}^{-1}$$

Dem.: A demonstração deste teorema é feita usando-se os teoremas (1.5), (1.6), (1.7), (1.8) e o <u>lema (1.1)</u>.

Teorema (1.15):
$$\frac{\partial \vec{F}}{\partial \phi} = A B' \Omega B \Lambda - \text{diag}(\Lambda'B'\Omega B \Lambda)$$
, onde $\Omega = V^{-1}(V - T)V^{-1}$

<u>Dem.</u>: A demonstração é feita utilizando-se os teoremas (1.5), (1.6), (1.7), (1.8) e o lema (1.1)

Teorema (1.16):
$$\frac{\partial F}{\partial \Psi} = 2 B' \Omega B \Psi$$
, onde
$$\Omega = V^{-1} \left(V - T \right) V^{-1}$$

<u>Dem</u>.: A demonstração é feita através dos teoremas (1.5), (1.6), (1.7), (1.8) e o lema (1.1)

Teorema (1.17):
$$\frac{\partial F}{\partial H} = 2 \Omega H$$
, onde
$$\Omega = V^{-1} \left(V - T \right) V^{-1}$$

Dem.: Este teorema é demonstrado usando-se os teoremas (1.5), (1.6), (1.7), (1.8) e o lema (1.1)

1.7 - DA DISTRIBUIÇÃO NORMAL

1.7.1 - Distribuição normal univariada

Seja Y uma variável aleatória que tem distribuição nor mal com média μ e variância σ^2 . Então a função densidade de Y é

$$f_{\gamma}(y) = \frac{1}{\sqrt{2\pi} \sigma} e^{\frac{(y-\mu)^2}{2\sigma^2}} = \frac{(y-\mu)^{\frac{1}{\sigma^2}} (y-\mu)^{\frac{1}{\sigma^2}}}{\sqrt{2\pi} \sigma} e^{\frac{(y-\mu)^{\frac{1}{\sigma^2}} (y-\mu)^{\frac{1}{\sigma^2}}}{2\sigma^2}}$$

onde
$$(y-\mu)' = (y-\mu)$$

1.7.2 - Distribuição normal multivariada

Seja Z um.vetor aleatório de dimensão p que tem distribuição normal multivariada, com vetor média μ e matriz de covariância $\Gamma = (\sigma_{ks})$, κ , $\kappa = 1, 2, \ldots$, positiva definida que denotamos por $N(\mu, \Gamma)$. A função densidade de Γ é dada por

tamos por
$$N[\underline{\mu}, \underline{\Sigma}]$$
. A função densidade de $-\frac{1}{2}(z-\underline{\mu})'\underline{\Sigma}^{-1}(z-\underline{\mu})$

$$f(z) = \frac{1}{(2\pi)^{p/2}|\underline{\Sigma}|^{1/2}}e$$

1.7.3 - Algumas propriedades da distribuição normal multivariada

- a) Se o vetor aleatório Z tem distribuição normal multivaria da, então a distribuição marginal de qualquer conjunto de componentes de Z tem distribuição normal multivariada (ver Anderson, 1958, pag. 24)
- b) Qualquer combinação linear de variáveis aleatórias normal mente distribuídas tem distribuição normal (ver Anderson 1958, pag. 37)
- c) Se toda combinação linear das componentes do vetor Z tem distribuição normal, então Z é normalmente distribuído (ver Anderson, 1958, pag. 37)

1.7.4 - Distribuição de uma matriz aleatória X

Seja X uma matriz alcatória (Nxp), onde as linhas de X são independentes e identicamentes distribuídas, com distri-

buição normal multivariada $H_p\left(\underline{\mu}, \underline{\Sigma}\right)$. A matriz \underline{X} pode ser expressa por $\underline{X}' = \left(\underline{X}_1, \underline{X}_2 \cdots \underline{X}_p\right)$, onde \underline{X}_k , $k=1,2,\cdots p$ são vetores coluna.

Seja

Os vetores linhas de X têm distribuição normal multivariada, então por (a), cada elemento de X tem distribuição normal univariada. Assim todos as coordenadas do vetor Y têm distribuição normal univariada. Mas por (b), gamantimos que qualquer combinação linear das componentes de Y tem distribuição normal univariada. Logo, por (c), concluímos que Y tem distribuição normal univariada.

E facil verificar que a matriz de covariancias entre os vetores colunas X_k e X_s é dada por $\sum_{X_k} X_s = \sigma_{ks} I_{NN}$, onde σ_{ks} é o elemento de \sum localizado na k-ésima linha com a s-ésima coluna. Isto significa dizer que a matriz de comariancias de X_s é dada por

$$\overline{\Sigma}_{Y} = \overline{\Sigma} \otimes \overline{I}_{N}$$
.

Dizemos que X tem distribuição $H_{\text{Nxp}}\left(\underline{\mu}_{X}, \sum \otimes \underline{\mathbb{I}}_{N}\right)$ onde $\underline{\mu}_{X}$ e $\left[\underline{\sum} \otimes \underline{I}_{N}\right]$ é a matriz média e a matriz de covariâncias de X, respectivamente.

1.8 - ALGUNS TIPOS DE CONVERGÊNCIA

Seja (X_n) , n=1,2 ... uma sequência de variáveis aleatórias e seja X uma variável aleatória. As variáveis aleatórias da sequência (X_n) e X são definidas no espaço de probabilidades (Ω, A, P) .

1.8.1 - Dizemos que a sequência (X_n) converge quase sempre (q.s) para a variável aleatória X quando $n \longrightarrow \infty$ (ou com probabilidade um), se $X_n(\omega) \xrightarrow[n \to \infty]{} X(\omega)$ $\forall \omega \in \Omega$ exceto, possivelmente, para um subconjunto N de Ω tal que p(N) = 0. Isto é, $\forall \varepsilon > 0$ e $\forall \omega \in N^C \supseteq N(\varepsilon, \omega) > 0$ tal que que

$$|X_{n}(\omega) - X(\omega)| < \varepsilon \quad \forall \quad n \geq N(\varepsilon, \omega).$$

Este tipo de convergência é também chamado "convergência for te".

1.8.2 - A sequência (X_n) converge em <u>probabilidade</u> para a variável aleatória X, e escrevemos $X_n \xrightarrow{p} X$, se $\forall \epsilon > 0, P(|X_n - X| > \epsilon) \xrightarrow{n \to \infty} 0.$

1.8.3 - Seja F_n a função distribuição de X_n para cada $n=1, 2, \ldots,$ e seja F a função distribuição de varia vel aleatória X.

Diz-se que a sequência (Xn) converge em distribuição para X

quando $n \longrightarrow \infty$, e escrevemos $X_n \xrightarrow{d} X$, se $F_n(x) \xrightarrow{n \longrightarrow \infty} F(x)$ $\forall x \in \mathbb{R}$ para o qual a F e continua. Isto e, se $\forall \varepsilon > 0$ e $\forall x \in \mathbb{R} \text{ onde F e continua} \exists N(\varepsilon, x) \text{ tal que}$

$$|F_n(x) - F(x)| < \varepsilon \quad \forall \quad n \ge N(\varepsilon, x)$$

Este tipo de convergência é também conhecido como "convergência fraca".

1.8.4 - A sequência de variáveis aleatórias converge em média quadrática para a variável aleatória X(m,q), e escrevemos $X_n \xrightarrow{m,q} X$, se $E\left(|X_n-X|^2\right) \xrightarrow{n \to \infty} 0$. Isto é, $\forall \epsilon > 0$, $\exists N(\epsilon)$ tal que $E\left(|X_n-X|^2\right) < \epsilon$, $\forall n \geq N(\epsilon)$ 1.8.5 - $X_n \xrightarrow{m,q} X \xrightarrow{m,q} X \xrightarrow{n \to \infty} X$

$$1.8.6 - X_n \xrightarrow{q.s} X \Longrightarrow X_n \xrightarrow{p} X$$

CAPITULO 2

ALGUNS MODELOS LINEARES MULTIVARIADOS

2.1 - INTRODUÇÃO

Neste capítulo apresentaremos o estudo resumido dos se guintes modelos lineares multivariados: análise fatorial, análise de componentes principais, análise de trajetórias e a nálise de curvas de crescimento. Os dois primeiros modelos e os dois últimos (sob certas condições) tornam-se casos particulares do modelo geral que apresentaremos no 4º capítulo.

2.2 - ANÁLISE FATORIAL

2.2.1 - A técnica de Análise Fatorial é amplamente utilizada em Psicologia, Ciências Sociais, Economia, etc. O pioneiro da Análise Fatorial foi Spearman (1904) utilizando um único fator. Em seguida, Thurstone (1931, 1947) utilizou multiplos fatores comuns. Estes fatores são variáveis desconhecidas, onde as variáveis observadas são expressas como funções dêles.

A primeira abordagem estatística do problema de estima cão de parâmetros foi feita por Lawley (1940-43). Esta técnica foi aprimorada por Yöreskog e Lawley (1968).

2.2.2 - O modelo fatorial

O modelo de Análise Fatorial é dado por

$$\underline{X} = \underline{\Lambda} \ \underline{Y} + \underline{Z} + \underline{\mu} \tag{2.1}$$

onde X(px1) é um vetor estocástico de variáveis respostas, μ é um vetor de médias, $\Lambda = (\lambda_{ij})$ i=1,2..., j=1,2,..., q é uma matriz paramétrica de posto completo q\lambda_{ij} são chamados cargas fatoriais, Y(qx1) é um vetor estocás

tico de fatores comuns ou variaveis latentes, con $E\left(\frac{Y}{Y}\right) = 0$, $V\left(\frac{Y}{Y}\right) = \Phi$ positiva definida (p.d) e Z(px1) é um vetor de erros aleatórios con $E\left(\frac{Z}{Z}\right) = 0$, $V\left(\frac{Z}{Z}\right) = \text{diag}\left(\frac{\Psi_{i}}{Y}\right)$ i = 1, 2..., p, p.d e $\text{cov}\left(\frac{Y}{Y},\frac{Z}{Z}\right) = 0$.

Aqui, $V\left(\frac{X}{X}\right)$ indica a matriz de covariâncias do vetor X.

Como consequência do modelo temos que $E\left(\frac{X}{X}\right) = \mu$ e $V\left(\frac{X}{X}\right) = 0$

$$\sum = \Lambda \Phi \Lambda^{\circ} + \Psi \qquad (2.2)$$

Os elementos da diagonal de $\sum - \Psi = \Lambda \Phi \Lambda'$ são denomidas comunalidades, enquanto os elementos de Ψ . são chamados especificades (ou variâncias específicas).

= [, onde a matriz de covariância [(pxp) é dada por

O objetivo essencial do modelo é expressar as variáveis respostas como combinações dos fatores comuns (variáveis latentes) mais os fatores específicos. Os fatores explicam a estrutura de covariância das respostas uma vez que $V\left(Z\right) = V\left(X/Y\right) = Y$. As hipóteses tomadas para (2.1) foram corretamente colocadas, pois, se $E\left(Y\right) = E$, então $E\left(X\right) = \mu + \Lambda$ E e claramente μ e E são identificados juntos. Para este caso bastaria fazermos $Y^* = Y - E$ e $\mu^* = \mu + \Lambda$ E transformando o caso geral em (2.1) sem nenhuma perda de generalidade.

Muitos autores adotam o modêlo onde $E\left(X\right) = 0$ e $E\left(X\right) = 0$, ou seja, escrevendo $X^* = X - \mu$, e outros além disso obrigam os fatores comuns serem não correlacionados com

variâncias unitárias.

Através de restrições impostas a Λ e Φ, pode-se identificar o modelo. Duas soluções podem ser apresentadas: Análise Fatorial Irrestrita ("exploratory") e Análise Fatorial Restrita ("confirmatory").

2.2.3 - Análise Fatorial Irrestrita

A - Para este ramo da Análise Fatorial, $\frac{1}{2}$ q(q+1) restrições são impostas, com a imposição de que os fatores comuns sejam não correlacionados e com variância unitária . Queremos dizer que (2.2) torna-se

$$\underline{\Sigma} = \underline{\Lambda} \underline{\Lambda}^{4} + \underline{\Psi} \qquad (2.3)$$

B - Rotação de Fatores

Em alguns casos é necessário uma melhor interpretação dos fatores. Para isto, o sistema de referências deve ser mudado, ou seja, uma transformação linear dos fatores deve ser feita sem modificar a dimensão do espaço destes fatores comuns.

Seja M uma matriz ortogonal qualquer de ordem q; então substituindo Λ por $\Lambda^* = \Lambda$ M em (2.3) verifica-se que $\Sigma = \Lambda^* \Lambda^{*'} + \Psi$. Significa dizer, que houve uma rotação rigida dos fatores: $Y^* = M Y$.

Técnicas de rotação foram criadas, dentre as quais podemos citar: método gráfico de rotação dos fatores, método varimax, método quartimax e outros (Lawley e Maxwell, cap 6, 1971). Ver Fachel, 1976.

C - ESTIMAÇÃO

Um problema que aparece naturalmente, é aquêle de determinar os estimadores das cargas λ_{ij} e das variâncias específicas Ψ_i .

Muitos métodos para solucionar este problema têm surgido, como o clássico método de Estimação de Máxima Verossimilhança, Solução do Fator Principal, Método Centroide, Método de Redução de Posto, Método do Grupo Centroide, Método de Ortogonalização, Método Jacobi, etc...

O método de Máxima Verossimilhança tem recebido atenção por parte de Yöreskog (1967, 1968), Lawley (1971), Rodri gues (1975) e outros. A solução do Fator Principal pode ser encontrada em Afifi e Azen (1972). Para uma descrição dos demais, ver Horst, 1965.

Quem primeiro deu solução para o cálculo de estimadores pelo método de Máximo Verossimilhança foi Lawley, 1940; um desenvolvimento completo é encontrado no livro de Lawley e Maxwell, 1963.

Neste item iremos apresentar um resumo do trabalho de Yöreskog, 1967.

A equação fundamental é dada por

$$Y = \Lambda X + Z \tag{2.4}$$

com matriz [dada por (2.3).

Aqui, admitiremos que Y tenha distribuição normal multivaria da com matriz dispersão 2.

Sejam Y₁, Y, ..., Y_N N observações de Y. Então, a média amostral é

$$\overline{\underline{Y}} = \frac{1}{N} \sum_{k=1}^{N} \underline{Y}_k , \qquad (2.5)$$

e a matriz dispersão amostral é

$$\underline{S} = \frac{1}{N-1} \sum_{k=1}^{N} (\underline{Y}_k - \overline{\underline{Y}}) (\underline{Y}_k - \overline{\underline{Y}})' = \frac{1}{N-1} \left(\sum_{k=1}^{N} \underline{Y}_k \underline{Y}_k' - \underline{N} \ \underline{\overline{Y}} \ \underline{\overline{Y}}' \right), \quad (2.6)$$

que tem uma distribuição de Wishart (ver Anderson, 1958).

Omitindo uma função das observações, o logaritimo de função de verossimilhança é dada por

$$\log L = -\frac{1}{2} (N-1) \left[\log(|\Sigma|) + \text{tr}(S \Sigma^{-1}) \right], \quad (2.7)$$

mas $\sum = \Lambda \Lambda' + \frac{\pi}{2}$ o que significa dizer que (2.7) está em função de Λ e Ψ .

Nosso objetivo é determinar os estimadores de máxima verossimilhança das verdadeiras Λ e Ψ , que denotaremos por $\tilde{\Lambda}$ e $\tilde{\Psi}$, respectivamente.

Para a obtenção de A e Y basta maximizarmos (2.7) ou minimizarmos a função

$$F_{\mathbf{q}}\left(\underline{\Lambda}, \underline{\Psi}\right) = \log(|\underline{\Sigma}|) + \operatorname{tr}(\underline{S}, \underline{\Sigma}^{-1}) - \log(|\underline{S}|) - p. \qquad (2.8)$$

O primeiro passo é determinar um mínimo condicional de

Fq para uma dada y e o segundo passo, o minimo global.

Demonstra-se que (ver Lawley e Maxwell, 1971)

$$\frac{a}{\partial \Lambda} F_{q} = 2 \sum^{-1} \left(\sum - \sum \right) \sum^{-1} \Lambda \qquad (2.9)$$

Então, para

$$\sum_{n=0}^{\infty} (\sum_{n=0}^{\infty} - \sum_{n=0}^{\infty}) \sum_{n=0}^{\infty} \Lambda = 0 , \qquad (2.10)$$

 $\tilde{\Lambda}$ é uma solução de (2.10), para dada Ψ . A equação (2.9) do Livro de Lawley e Maxwell 1971, da a expressão de $\tilde{\Gamma}^{-1}$ em função de $\tilde{\Lambda}$ e Ψ .

Partindo de (2.10) e usando os desenvolvimentos de \sum e $\sum_{i=1}^{n-1}$, através de operações matriciais, obteremos a expressão

$$\left(\underbrace{\Psi}^{-\frac{1}{2}} \underbrace{S} \underbrace{\Psi}^{-\frac{1}{2}} \right) \underbrace{\Psi}^{-\frac{1}{2}} \underbrace{\Lambda} = \left(\underbrace{\Psi}^{-\frac{1}{2}} \underbrace{\Lambda} \right) \left(\underbrace{I}_{q} + \underbrace{\Lambda}^{\bullet} \underbrace{\Psi}^{-1} \underbrace{\Lambda} \right) \tag{2.11}$$

E conveniente que Λ' $\Psi^{-1}\Lambda$ seja uma matriz diagonal. Assim, de (2.11) vemos que os vetores colunas de $\Psi^{-\frac{1}{2}}\Lambda$ e os elementos da diagonal de $\left(\frac{1}{2}q + \Lambda' \Psi^{-1}\Lambda\right)$ são vetores e valores proprios da matriz Ψ S Ψ respectivamente.

A matriz $\Psi^{-\frac{1}{2}}$ S $\Psi^{-\frac{1}{2}}$ tem p linhas e p colunas, ou seja, esta matriz tem p vetores proprios e seus valores proprios associados.

Como $\Psi^{-\frac{1}{2}}$ Λ tem q colunas, então, q valores e vetores proprios de $\Psi^{-\frac{1}{2}}$ S $\Psi^{-\frac{1}{2}}$ são necessários.

Sejam $\tilde{0}_1 \geq \tilde{0}_2 \geq \cdots \leq \tilde{0}_p$ as raizes características de $\underline{v}^{-\frac{1}{2}} \leq \underline{v}^{-\frac{1}{2}} = \underline{\tilde{u}}_1 \cdot \underline{\tilde{z}}_2 \cdots \underline{\tilde{u}}_q$ os q vetores característicos

correspondentes as q maiores raízes características. O simbolo ~ é usado acima para indicar que se está condicionando à dada ♥ .

Seja $\widehat{\mathbb{H}} = \operatorname{diag}(\widehat{\Theta}_{i})$ i=1,2,...q e $\widehat{\Omega}$ a matriz cujos veto res colunas são $\widehat{\omega}_{1}$, $\widehat{\omega}_{2}$, ... $\widehat{\omega}_{q}$, os quais podem ser escolhidos ortonormais.

Então,

$$\tilde{\Omega}' \tilde{\Omega} = \underline{I}_{q} , \qquad (2.12)$$

e pelo que foi dito,

$$\underline{\Psi}^{-\frac{1}{2}}\tilde{\Lambda} = (\widehat{\Theta} - \underline{I}_{q})^{\frac{1}{2}} \qquad (2.13)$$

Multiplicando à esquerda de (2.13) por $\Psi^{\frac{1}{2}}$ temos

$$\tilde{\Lambda} = \Psi^{\frac{1}{2}} \tilde{\Omega} \left(\tilde{H} - I_{\alpha} \right)^{\frac{1}{2}}, \qquad (2.14)$$

que é o estimador de máxima verossimilhança condicionado de

Em seguida, através de técnicas de determinantes é mos trado que $F_q\left(\tilde{\Lambda}, \Psi\right)$ toma a forma

$$F_{\mathbf{q}}\left(\tilde{\Lambda}, \Psi\right) = f_{\mathbf{q}}(\Psi) = -\log\left(\prod_{i=q+1}^{p} 0_{i}\right) + \sum_{i=q+1}^{p} 0_{i} - (p-q) \quad (2.15)$$

que é o mínimo condicional de Fq.

Utilizando a função $(x - \log x)$ justifica-se que o conjunto das últimas (p-q) raízes características, com ordem fixada anteriormente, é aquêle que dá o menor valor de $f_q(\Psi)$,

entre todos os outros conjuntos constituídos de (p-q) raízes características, retiradas do conjunto das raízes características de Ψ S Ψ .

O próximo passo é o de minimizar f_q(Ψ). É importante observar que

$$\frac{\partial}{\partial \Psi} f_{\mathbf{q}}(\underline{\Psi}) = \frac{\partial}{\partial \Psi} F_{\mathbf{q}}(\underline{\tilde{\Lambda}}, \underline{\Psi}) \qquad (2.16)$$

A derivada de F_q com respeito a Ψ tem a forma

$$\frac{\partial}{\partial \Psi} F_{q} \left[\underline{\Lambda}, \underline{\Psi} \right] = \operatorname{diag} \left[\underline{\Sigma}^{-1} (\underline{\Sigma} - \underline{S}) \underline{\Sigma}^{-1} \right] \qquad (2.17)$$

Em $(\tilde{\Lambda}, \Psi)$, (2.17) torna-se

$$\frac{\partial}{\partial \Psi} f_{\mathbf{q}}(\underline{\Psi}) = \operatorname{diag}\left(\underline{\Psi}^{-1}(\underline{\tilde{\Lambda}} \underline{\tilde{\Lambda}}' + \underline{\Psi} - \underline{\hat{S}})\underline{\Psi}^{-1}\right), \qquad (2.18)$$

logo, o valor que anula (2.18) $\tilde{e} \tilde{\Psi} = \text{diag} \left(\tilde{S} - \tilde{\Lambda} \tilde{\Lambda}' \right)$ (2.19)

Os estimadores de máxima verossimilhança de $\tilde{\Lambda}$ e Ψ são dados em (2.14) e (2.19).

A obtenção dos estimadores será feito pelo metodo de Fletcher e Powell (1963), aplicado à função $f_q(\Psi)$.

D - TESTE DE HIPOTESE

Seja Ω o conjunto de todas as matrizes simétricas, positivas definidas com p linhas e p colunas, e denotemos ω o subconjunto de Ω , onde seus elementos satísfazem (2.3) para um dado valor q_0 de q, onde q o número de fatores comuns.

Consideremos as hipoteses

$$H_{0}: \sum_{\alpha} \in \omega \qquad (2.20)$$

$$H_{1}: \sum_{\alpha} \in (\Omega - \omega)$$

como teste de boa adaptação do modelo. O próximo passo é a apresentação da estatística teste que é dada através do princípio da razão de verossimilhança.

Seja

$$\log L_{\Omega} = \max_{\Omega} \log L = -\frac{1}{2} N \Big(\log(|S|) + p \Big),$$
 (2.21)

onde L e a função de verossimilhança (ver Anderson 1958, cap. 3).

Sob a hipotese nula Ho, temos

$$\log L_{\omega} = \max_{\omega} \log L = -\frac{1}{2} N \left(\log(|\tilde{\Sigma}|) + \operatorname{tr}(\tilde{\Sigma} \tilde{\Sigma}^{-1}) \right), \quad (2.22)$$

A razão de verossimilhana é dada por

$$\lambda = \frac{L_{\omega}}{L_{\Omega}}$$
 (2.23)

a qual não possui uma distribuição exata. Entretanto a esta tística

$$U_q = -2 \log \lambda = nN \Big(\log(|\tilde{\Sigma}|) - \log(|S|) + tr(S|\tilde{\Sigma}^{-1}) \Big), (2.24)$$
para uma amostra suficientemente grande, tem distribuição χ_v^2 , dado que H_0 seja verdadeira. O gran de liberdade v é o núme ro de parâmetros em Ω menos o número de parâmetros sobre H_0 , isto é,

$$v = \frac{1}{2} \left[(p-q)^2 - (p+q) \right].$$
 (2.25)

A hipotese nula II_0 com exatamente q fatores é rejeitada ao nível α se

$$U_{q} \geq x_{v;\alpha}^{2}, \qquad (2.26)$$

e não rejeitada no caso contrário. Ver Yöreskog, 1967.

2.2.4 - ANALISE FATORIAL RESTRITA

A - Na analise fatorial irrestrita, as seguintes condições são impostas:

a)
$$\phi = I_q$$

b) a matriz $\Lambda' \Psi^{-1} \Lambda \in \text{diagonal e seus elementos}$ $\left[\Lambda' \Psi^{-1} \Lambda\right]_{ii} \geq \left[\Lambda' \Psi^{-1} \Lambda\right]_{jj}, j=i+1, i=1,2\cdots q.$

Mas nenhuma restrição sôbre os fatores são feitas.

Na análise fatorial restrita, os fatores podem ser independentes ou não. As restrições que surgem são sôbre 1, e e v no modelo (2.2).

As soluções restritas e irrestritas podem ser únicas ou não. Uma solução é única se toda transformação linear dos fatores que mantém os parâmetros fixos não alterados , também mantém os parâmetros livres não alterados. A unicida de de solução tem sido abordada por vários autores. Ver Yöreskog 1969.

B - A FUNÇÃO F E A TÉCNICA DE MINIMIZAÇÃO

Admitamos que o vetor variavel X tenha uma distribui

ção normal multivariada. Como em (2.1.3), tomamos uma amostra de tamanho N (N > p) e denotaremos por S a matriz dispersão amostral, que tem a mesma formula de (2.6).

Então, omitindo a função das observações temos para logarítimo da função de verossimilhança L:

$$\log L = -\frac{1}{2} N \left(\log(|\underline{\Sigma}|) + \operatorname{tr}(\underline{S} \underline{\Sigma}^{-1}) \right)$$
 (2.27)

e em seguida consideremos a função

$$F\left(\Lambda, \Phi, \Psi\right) = \log(|\Sigma|) + \operatorname{tr}(S\Sigma^{-1}) - \log(|S|) - p (2.28).$$

Ao inves de maximizarmos (2.27), minimizaremos (2.28). Evidentemente a minimização será feita em relação aos parâmetros livres.

gumento, o vetor o e o gradiente de F(o) possuem h coordenadas.

Uma vez que as derivadas de primeira ordem foram expressas analiticamente, então a técnica de mimimização de Fletcher e Powell pode ser aplicada, (ver Yöræskog, 1969, Lawley e Maxwel, 1971). Para este método existe um programa para computador feito por Yöreskog e Gruvaœus, 1967, na linguagem FORTRAN IV. Uma vez que o mínimo da função F tenha sido encontrado, então os valores de Λ, Φ e Ψ que minimizaram a F, são chamados estimadores de máxima verossimilhanaça e são representados por Λ, Φ e Ψ. Assim o E. M. V de Σ, ê dado por

$$\hat{\Sigma} = \hat{\Lambda} \hat{\Phi} \hat{\Lambda} + \hat{\Psi} \qquad (2.29)$$

C - TESTE DE ADAPTAÇÃO

Da mesma maneira como foi tratado em (2.1.3), definire mos Ω como o conjunto das matrizes simétricas, positivas definidas e ω como subconjunto de Ω , cujos elementos satisfa zem (2.2). O teste empregado para testar Π_0 é o da razão de verossimilhança.

A estatística para o teste é

$$-2 \log \lambda = N \left(\log(|\underline{\Sigma}|) + \operatorname{tr}(\underline{S} \underline{\Sigma}^{-1}) - \log(|\underline{S}|) - p \right), (2.30)$$

onde - 2 log λ sob H_o segue aproximadamente um χ_v^2 , tal que $v = p^2 - \frac{1}{2} (p + q)(p + q + 1) + m$

2.3 - ANALISE DE COMPONENTES PRINCIPAIS

O estudo da Análise de Componentes Principais começou com K. Pearson (1901). Dado um conjunto de vetores observação, Pearson interessou-se em determinar a reta que minimiza a soma de quadrados de distâncias perpendiculares daqueles pontos à reta. Outros estatísticos como Hotelling, Anderson e Lawley também se interessaran pelo estudo da Análise de Componentes Principais sendo que os dois últimos apresentaram resultados recentes das propriedades amostrais das componentes principais, que estão apresentados no capítulo 7 de Morrison, 1967.

Supomos que temos p variaveis X₁, X₂...X_p e que estamos interessados em verificar as relações existentes entre elas. Por exemplo: um antropólogo pode tomar diversas medidas de comprimento e largura de um conjunto escolhido de indivíduos, como comprimento da orelha, largura da orelha, comprimento facial, largura facial e assim por diante. O antropólogo pode estar interessado em analisar como estes indivíduos diferem relativamente a estas características. Então, através de combinações de medidas êle pode explicar essas diferenças. As componentes principais formam um novo conjunto de combina

ções lineares de medidas. Seja X' = $\left(x_1, x_2, \dots, x_p\right)$, p - vetor aleatório que possue distribuição multivariada (não necessariamente normal), com vetor média $\mu' = \left(\mu_1, \mu_2, \dots, \mu_p\right)$ e matriz (pxp) de covariância $\sum = (\sigma_{ij})$, i, j = 1,2,...,p.

O método das componentes principais é aquêle que determina p combinações lineares.

 $Y_1 = \sum_{j=1}^{p} a_{1j} X_j$, $Y_2 = \sum_{j=1}^{p} a_{2j} X_j$,..., $Y_p = \sum_{j=1}^{p} a_{pj} X_j$, (2.30) tais que

 $\begin{aligned} &\text{Cov}\Big(Y_{\mathbf{i}},Y_{\mathbf{j}}\Big) = 0 \quad \text{para} \quad \mathbf{i},\mathbf{j} = 1,2...,p\,, \quad \mathbf{i} \neq \mathbf{j}\,, \\ &V\Big(Y_{\mathbf{i}}\Big) \geq V\Big(Y_{\mathbf{j}}\Big) \geq \cdots \geq V\Big(Y_{\mathbf{p}}\Big), \quad \text{com} \quad ||a_{\mathbf{i}}||^2 = a_{\mathbf{i}} a_{\mathbf{i}} = 1, \quad \mathbf{i} = 1, \\ &2,...,p \quad \text{onde} \quad a_{\mathbf{i}} = \Big(a_{\mathbf{i}1}, a_{\mathbf{i}2},...,a_{\mathbf{i}p}\Big). \quad \text{A tecnica para a determinação das componentes é puramente matemática.} \end{aligned}$

Seja X um vetor aleatório tal que X' = $\begin{bmatrix} X_1, X_2, ..., X_p \end{bmatrix}$, $E\begin{bmatrix} X \end{bmatrix}$ = μ e $V\begin{bmatrix} X \end{bmatrix}$ = \sum (pxp) de posto p, e que os elementos de μ e \sum sejam finitos. As praízes características de \sum são $r_1 > r_2 > \cdots > r_p$.

Supomos que o vetor aleatório X tenha distribuição multivariada, e que desta população uma amostra de N vetores aleatórios são extraídos. Com estes dados a matriz S, de covariân cia amostral fica determinada. Admitindo que S seja desconhecida, então a matriz S o substituirá neste desenvolvimento (ver Morrison, 1067 e Afifi & Azen, 1974).

Seja
$$\alpha_1' = \left(\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}\right)$$
 o vetor tal que
$$Y_1 = \sum_{j=1}^{p} \alpha_{1j} X_j = \alpha_1' X \qquad (2.31)$$
 ondé $||\alpha_1||^2 = \alpha_1' \alpha_1 = 1$ e $X' = \left(X_1, X_2, \dots, X_p\right)$.

Nosso objetivo é determinar um vetor solução entre todos os α_1 tal que $S_{y_1}^2$ (variância amostral de y_1) seja máxima, sob a restrição $||\alpha_1||^2 = 1$. Para este fim, usamos multiplicador de Lagrange. Definamos a equação

$$f_1(\underline{\alpha}_1) = \left(\underline{\alpha}_1' \ \underline{S} \ \underline{\alpha}_1 + \gamma_1(1 - ||\underline{\alpha}_1||^2)\right), \quad (2.32)$$

onde $S_{y_1}^2 = \alpha_1^* S_{x_1} = \alpha_1$ é a variância amostral de y_1 . Então

$$\frac{\partial}{\partial \underline{\alpha}_1} f_1(\underline{\alpha}_1) = 2 \left[\underline{S} - \gamma_1 \underline{I}_p \right] \underline{\alpha}_1 = 0 \quad (2.33)$$

Para obtermos uma solução não nula, devemos fazer $\det\left(\underline{S} - \gamma_1 \underline{I}_p\right) = 0 \text{ que tem p raizes } \lambda_1, \lambda_2, \dots, \lambda_p, \text{ os quais são raizes características de } \underline{S}. \quad 0 \text{ valor } \ell_1 = \max\left\{\lambda_i, i=1, 2\cdots p\right\}$ é o que maximiza $S_{\gamma_1}^2$. O vetor proprio associado a ℓ_1 é representado por a_1 .

Em seguida, propomos o 2º estágio do processo, ou seja, determinar um vetor solução a para que a combinação

$$Y_2 = \sum_{j=1}^{p} \alpha_{2j} X_j = \alpha_2^* X$$
, (2.34)

tenha variância maxima entre todos os α_2 subordinados as restrições, $||\alpha_2||^2 = 1$, $\alpha_1 \alpha_2 = 0$ e $x_{y1} \ge S_{y2}^2$. Então, definin

do a função

$$f_2(\alpha_2) = \alpha_2 S \alpha_2 + \gamma_2 \left[1 - ||\alpha_2||^2\right] + \mu \alpha_1 \alpha_2, \quad (2.35)$$

dai temos
$$\frac{\partial}{\partial \alpha_2}$$
 $f_2(\alpha_2) = 2\left(S - \gamma_2 I_p\right)\alpha_2 + \mu \alpha_1 = 0$ (2.36)

Para obtenção da solução não nula, façamos $\det \begin{bmatrix} S - \gamma_2 I_p \end{bmatrix} = 0$. Da mesma maneira como foi explicado anteriormente, concluímos que $\ell_2 = \max \left\{ \begin{cases} \lambda_i; \ i=1,2\cdots p \end{cases} - \left\{ \ell_1 \right\} \right\}$ é a segunda maior raíz característica de S e que, seu vetor proprio associado a solução procurada. Desta maneira, o processo continua até a p-ésima componente principal.

Definição

A k-ésima componente principal é $Y_k = \sum_{i=1}^{p} a_{ki} X_i = a_k X$ onde a_k é o vetor característico de S correspondente a k-ésima maior raíz característica ℓ_k , a qual é a variância amostral de Y_k .

A mesma técnica é empregada quando [é conhecida.

A matriz de variâncias e covariâncias é dada por:

$$\sum_{r} = r^{\dagger} \Phi r, \qquad (2.39)$$

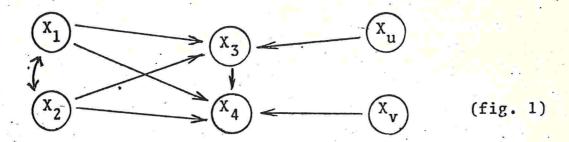
onde ϕ = diag $\left(\tau_{i}\right)$, i = 1,2 ··· p, tal que τ_{i} é a i-ésima maior raíz característica de Σ . A expressão (2.39) pode ainda ser apresentada da seguinte maneira

$$\Sigma = \Lambda^{\circ} \Lambda , \qquad (2.40)$$

onde $\Lambda = \frac{1}{2} \Gamma$. Assim, verifica-se que o modelo de componentes principais, é também um modelo fatorial, com p fatores comuns.

2.4 - ANALISE DE TRAJETORIAS

Em 1918 Sewall Wright desenvolveu o trabalho, "On nature of size factors", no qual introduziu a técnica de análise de trajetórias. Na análise de trajetórias (caminhos) temos dois tipos de variáveis; dependentes (observadas) e variáveis independentes ("ultimate"). As variáveis independentes tes são subdivididas em duas partes, variáveis independentes observadas é variáveis independentes latentes (não observadas). Uma maneira de representar tais variáveis é em forma de diagrama. Por exemplo



onde X₁, X₂, X₃ e X₄ são variáveis observadas, enquanto as variáveis X_u e X_v são variáveis independentes latentes (não observadas). Qualquer variável apontada por uma seta é expressa como função da variável da qual a seta partiu. Na figura 1, observamos que X₁ e X₂ são dependentes entre si, que X₃ fica completamente determinada por X₁, X₂ e X_u, e que X₄ é completamente determinada por X₁, X₂, X₃ e X_v.

O objetivo da análise de trajetória é especificar equações lineares que são equivalentes a um diagrama. As seguintes equações lineares

$$X_3 = b_{31}X_1 + b_{32}X_2 + b_{3u}X_u$$
 (2.41)

$$X_4 = b_{41}X_1 + b_{42}X_2 + b_{43}X_3 + b_{4v}X_v$$
 (2.42)

são equações lineares que representam o diagrama da (fig. 1).

Os coeficientes b são chamados coeficientes de trajetórias

e as equações lineares são as equações estruturais.

Suponha que as variáveis aleatórias X_k , k=1,2,3,4 sejam padronizadas, ou seja, $E\left(X_k\right)=0$ e $\sigma_{X_k}^2=1$.

Tomemos r_{ij} como o coeficiente de correlação entre X_i e X_j. Admitamos que estas variáveis aleatórias sejam vetores de valores observados com dimensão n. Suponhamos também que as variáveis aleatórias X_u e X_v sejam padronizadas e que cada uma delas represente um vetor de dimensão n de valores não observados. As variáveis aleatórias X_u e X_v são também chamadas <u>variáveis aleatórias residuais</u>. Multiplicando ambos os membros de (2.41) por X₁, temos

$$\underline{x}_{1}^{\prime} \underline{x}_{3} = b_{31} \underline{x}_{1}^{\prime} \underline{x}_{1} + b_{32} \underline{x}_{1}^{\prime} \underline{x}_{2} + b_{3u} \underline{x}_{1}^{\prime} \underline{x}_{u},$$
(2.43)

que dividida por n , teremos

$$r_{13} = b_{31} + b_{32} r_{12}$$
 (2.44)

Usando o mesmo processo de (2.43), encontramos

$$r_{23} = b_{31} r_{12} + b_{32}$$
, (2.45)

$$r_{3u} = b_{3u}$$
 (2.46)

$$1 = b_{31} r_{13} + b_{32} r_{23} + b_{3u} r_{3u}$$
 (2.47)

As equações (2.44) e (2.45) determinam os valores de b₃₁ e b₃₂, jã as (2.46) e (2.47) determinam o valor de b_{3u}.

Associado a cada diagrama, temos um conjunto de equações de regressão linear. Assim, cada equação pode aparecer na forma

$$X_0 = \alpha_1 + \sum_{i=1}^{m} \beta_{0i} X_i + \varepsilon \text{ ou na forma}$$
 (2.48)

$$Z_{o} = \sum_{i=1}^{m} P_{oi} Z_{i} + P_{ou} Z_{u}, \text{ onde}$$
 (2.49)

$$Z_{i} = \frac{X_{i} - \mu_{i}}{\sigma_{i}}$$
 $i = 0, 1, 2, ..., m$ (2.50)

$$\varepsilon = P_{\text{ou}} Z_{\text{u}}$$
 e finalmente (2.51)

$$P_{0i} = \frac{\beta_{0i} \sigma_{i}}{\sigma_{0}}$$
, $i = 1, 2, ... m$ (2.52)

Os coeficientes P_{oi} de (2.48) são chamados de coeficientes de trajetória. A variável dependente Z_o é demominada variável efeito, enquanto as variáveis independentes são chamadas variáveis causa. Os coeficientes P_{oi} medem o efeito direto de uma variável sôbre outra. Um estimador para o P_{oi} é dado por $\hat{P}_{oi} = \frac{S_i}{S_o}$ $\hat{\beta}_{oi}$ $i=1,2,\ldots$ m, onde S_o^2 e S_i^2 são as variân cias amostrais de X_o e X_i , $i=1,2,\ldots$ m, respectivamente. O teorema fundamental da análise de trajetória diz que o coeficiente de correlação de Pearson entre X_i e X_i é dado por

$$\mathbf{r}_{ij} = \frac{1}{n} \sum_{k=1}^{n} \mathbf{z}_{ik} \mathbf{z}_{jk}, \text{ onde n } \mathbf{e} \text{ o}$$
 (2.53)

número observado. Referente a determinado diagrama,

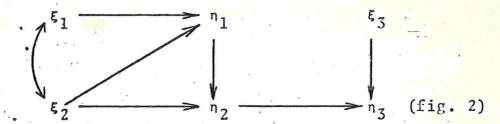
$$\mathbf{r}_{ij} = \begin{cases} \sum_{q} \hat{\mathbf{p}}_{iq} & \text{para } i \neq j \\ q & \text{omde q percorre } \underline{to} \end{cases}$$

$$\mathbf{para } i = j$$

das as variaveis associadas as possíveis trajetórias entre X_i e X_j. Ver Achear (1976).

Vejamos, agora, alguns modelos

Exemplo 2.1



As equações estruturais são dadas por

$$\eta_{1} = \gamma_{11} \xi_{1} + \gamma_{12} \xi_{2} + 0 \xi_{3}$$

$$\eta_{2} = \alpha_{21} \eta_{1} + \gamma_{22} \xi_{2} + 0 \xi_{3}$$

$$\eta_{3} = \alpha_{32} \eta_{2} + \gamma_{33} \xi_{3} ,$$
(2.54)

as quais podem ficar sob a forma

$$\eta_{1} = \gamma_{11} \xi_{1} + \gamma_{12} \xi_{2} + C \xi_{3}$$

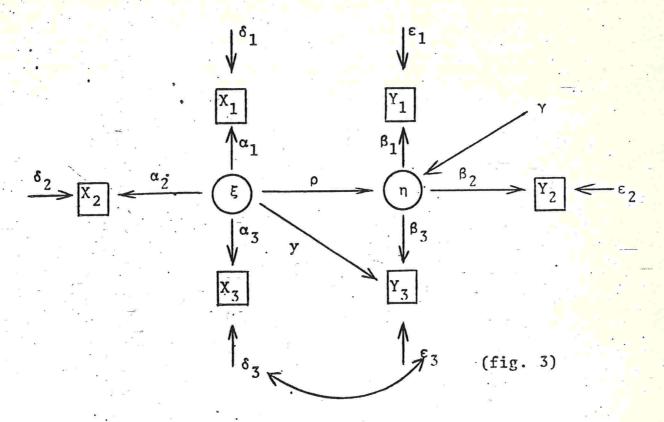
$$\eta_{2} = \alpha_{21} \gamma_{11} \xi_{1} + (\alpha_{21} \gamma_{12} + \alpha_{22}) \xi_{2} + C \xi_{3}$$

$$\eta_{3} = \alpha_{32} \alpha_{21} \gamma_{11} \xi_{1} + \alpha_{32} (\alpha_{21} \gamma_{12} + \alpha_{22}) \xi_{2} + \gamma_{33} \xi_{3}$$

$$(2.55)$$

O exemplo que segue, é um modelo fictício apresentado por Yöreskog (1970a)

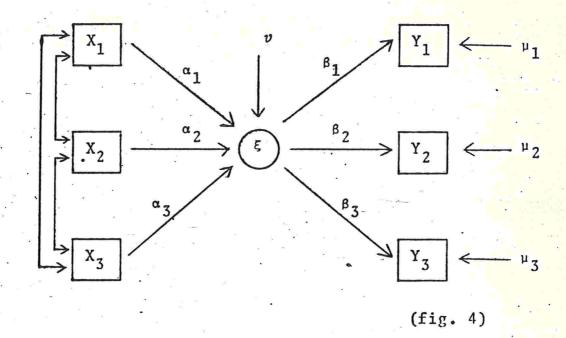
Exemplo 2.2



Aqui,

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \alpha_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \alpha_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & \beta_3 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi \\ \eta \\ \delta_1 \\ \delta_2 \\ \delta_3 \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$
(2.56)

Exemplo 2.3



$$\xi = \alpha X + v \quad e \quad \underline{Y} = \beta \xi + \underline{\mu}$$
 (2.57)

sabendo-se que $\underline{\alpha}' = \left(\alpha_1, \alpha_2, \alpha_3\right), \ \underline{X}' = \left(X_1, X_2, X_3\right), \ \underline{Y}' = \left(Y_1, Y_2, Y_3\right)$ e $\underline{\mu}' = \left(\mu_1, \mu_2, \mu_3\right)$. Os modelos (2.2) e (2.3) estão apresntados em Yöreskog, 1972.

2.5 - ANÁLISE DE CURVAS DE CRESCIMENTO

2.5.1 - Introdução

Suponhamos que se tenha r grupos de indivíduos, onde o k-ésimo grupo tenha N_k indivíduos e que o número total de indivíduos seja $N = \sum_{k=1}^{r} N_k$. Suponhamos ainda, que no k-ésimo grupo, k=1,2...r, em cada unidade experimental, a mesma ca-

racterística é medida em p diferentes ocasiões durante cer to intervalo de tempo. Desejamos determinar para cada grupo, a curva que melhor se adapte as observações, comparar as curvas, testar hipótese sobre os parâmetros envolvidos, achar intervalos de confiança para estes parâmetros e determinar os seus estimadores. O modelo apropriado para estudo de dados deste tipo ê chamado "análise de cruvas de crescimento".

2.5.2 - O modelo de curva de crescimento

Seja X a matriz (Nxp) das observações, A una natriz (Nxr) conhecida, B uma matriz (qxp) conhecida, g a matriz (rxq) dos parâmetros desconhecidos e E a matriz (Nxp) de erros cujos vetores linha são independentes e identicamente distribuídos (i.i.d), com distribuição p-variada, matriz de covariância [e vetor média a matriz nula 0 . 0 modelo geral de curva de crescimento é dado por

$$X = A \xi B + E$$
(2.58)

Nxp Nxr rxq qxp Nxp

onde A é a matriz de planejamento e B é a matriz cujo elemento localizado na k-ésima linha com a e-ésima coluna é t_e^{k-1} .

2.5.3 - Estimação

Supomos sem perda de generalidade que B(qxp), $\Lambda(Nxr)$ sejam ambas de posto completo. Admitamos que os vetores li-

nha de E [Nxp.] sejam i.i.d com distribuição p-variada N_p [0 Σ]. O modelo (2.58) pode também ser apresentado na forma

$$X' = B' \quad \xi' \quad A' + E'$$
 $pxN \quad pxq \quad qxr \quad rxN \quad pxN$
(2.59)

Khatri (1966) desenvolveu um trabalho onde mostra que o esti mador de máxima-verossimilhança de ξ' e dado por

$$\hat{\xi}' = \begin{pmatrix} B & D^{-1} & B' \end{pmatrix} \begin{pmatrix} B & D^{-1} & X' & A & (A' & A)^{-1} \\ qxr & qxp & pxp & pxq & qxp & pxp & pxN & Nxr & rxN & Nxr \end{pmatrix} (2.60)$$

onde

$$\underline{D} = \underline{X}' \left(\underline{I}_{N} - \underline{A} \left(\underline{A}' \underline{A} \right)^{-1} \underline{A}' \right) \underline{X} \quad \text{(ver Morrison, 1976, pxp pxN NxN Nxr rxN Nxr rxN Nxr Nxp Cap. 5).}$$

2.5.4 - Testes de Hipóteses Lineares Referentes a ¿

Como na análise de variância, as hipóteses lineares so

bre ¿ podem ser dadas por

$$H_0: \bar{F} = \underline{\xi}' \quad \underline{C}' = \underline{0} \quad \text{contra}$$
 (2.61)
 $H_1: \bar{F} = \underline{\xi}' \quad \underline{C}' \neq \underline{0}$,

onde F é uma matriz conhecida (cxq) de posto c ≤ q, enquanto C' é uma matriz conhecida (rxg) de posto g ≤ r. A estatística teste é expressa em função das matrizes

$$\underline{E} = \underline{F} \left(\underline{B} \underline{D}^{-1} \underline{B}' \right)^{-1} \underline{F}'$$

$$\underline{Cxc} \quad \underline{Cxq} \quad \underline{qxp} \quad \underline{pxq} \quad \underline{qxc}$$
onde

$$\underline{R} = (\underline{A}' \underline{A})^{-1} + (\underline{A}' \underline{A})^{-1} \underline{A}' \underline{X} \underline{D}^{-1} \underline{X}' \underline{A} (\underline{A}'\underline{A})^{-1}$$

$$\underline{rxr} \quad \underline{rxN} \quad Nxr \quad \underline{rxr} \quad \underline{rxN} \quad Nxp \quad pxp \quad pxN \quad Nxr \quad \underline{rxr}$$

$$- \underline{\hat{\xi}} (\underline{B} \quad \underline{D}^{-1} \quad \underline{B}')^{-1} \quad \underline{\hat{\xi}}'$$

$$\underline{rxq} \quad qxq \quad qxr$$

Para testar H_0 , podemos usar o critério da maior raiz de $H E^{-1}$, $e = \frac{1}{1 + \lambda_{max}}$, com parâmetros s = min(c,g), $m = \frac{|c-g|-1}{2}$ $e = n = \frac{N-r-p+q-c-1}{2}$, aplicando a tabela de Heck ou a de Pillai (ver Morrison, 1976).

2.5.5 - Alguns Exemplos

A - Potthoff e Roy, 1964 apresentaram um modelo de análise de Curvas de Crescimento (2.58), que é uma generalização do usual modelo de ANOVA, sendo que X tem vetores linha independentes, idênticamente distribuídos, com distribuíção normal p-variada, média 0 e matriz de covariância . Neste trabalho, apresentaram também quatro exemplos aplicados a este modelo, mas citaremos aqui somente dois exemplos. O primeiro exemplo envolve somente um grupo de animais, enquanto o segundo exemplo envolve m grupos de animais.

a₁) Seja um grupo com N animais, onde todos os animais estão

sujeitos a mesma medida, e que a medida de cada animal seja observada em p tempos diferentes, t_1, t_2, \cdots, t_p . As p observações são dependentes e possuem distribuição normal p-variada com media O e matriz de covariância desconhecida \sum_{0} . A curva de crescimento associada obedece ao modelo (2.58), onde A(Nx1), $\xi(1xp)$ e P(px1), são dada por

A' =
$$\begin{bmatrix} 1, 1, \cdots & 1 \end{bmatrix}$$
, $\xi = \begin{bmatrix} \xi_0, \xi_1 & \cdots & \xi_{p-1} \end{bmatrix}$ e P' = $\begin{bmatrix} 1, t, \cdots t^{p-1} \end{bmatrix}$, assim o valor esperado da variável medida do k-ésimo animal $\xi \in \mathbb{E} \begin{bmatrix} X_k \end{bmatrix} = \begin{bmatrix} p-1 \\ j=0 \end{bmatrix}$ ξ_i t^i

a₂) Sejam m grupos de animais. O k-ésimo grupo possui N_k animais e os m grupos possuem um total de $N = \sum_{k=1}^{m} N_k$ animais. Cada grupo está sujeito a diferentes tratamentos, e os animais de cada grupo são medidos em p pontos de tempo, num determinado intervalo de tempo. O vetor das p medidas de cada animal de cada grupo, tem uma distribuição normal p-variada com média 0 e matriz de covariância \sum_{0} . Estes vetores são independentes e idênticamente distribuídos. O modelo de cur vas de crescimento associado ao estudo destes m grupos de animais, tem a mesma expressão (2.58), sendo que as matrizes A(Nxm), $\xi(mxp)$ e P(pxq) são dadas por

$$\underline{\xi} = \begin{cases} \xi_{\text{oo}} & \xi_{\text{ol}} & \vdots & \xi_{\text{op-1}} \\ \xi_{\text{lo}} & \xi_{\text{11}} & \vdots & \xi_{\text{1p-1}} \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{\text{mo}} & \xi_{\text{ml}} & \vdots & \xi_{\text{mp-1}} \end{cases}$$

$$P =
 \begin{cases}
 1 & 1 & \cdots & 1 \\
 t & t & \cdots & t \\
 t^2 & t^2 & \cdots & t^2
 \end{cases}$$

$$\begin{bmatrix}
 t & t & \cdots & t \\
 t^2 & t^2 & \cdots & t^2
 \end{bmatrix}$$

$$\begin{bmatrix}
 t & t & t & t & t \\
 t & t & t & t & t \\
 \vdots & \vdots & \vdots & \vdots & \vdots \\
 t^{p-1} & t^{p-1} & t^{p-1}
 \end{bmatrix}$$

A curva de crescimento associada ao k-esimo grupo e dada por

$$E\left(X_{kj}\right) = \sum_{i=0}^{p-1} \xi_{ki} t^{i}$$

Vas de crescimento. Esse trabalho envolve r grupos de individuos (ou animais), onde o k-ésimo grupo tem N_k indivíduos (ou animais), e que $N = \sum_{k=1}^{r} N_k$ é o número total de indivíduos (ou animais). Em cada grupo a mesma característica é medida em cada indivíduo (ou animal) em p tempos. O vetor das observações X_{ij} $i=1,2\cdots N_j$, do i-ésimo indivíduo (ou animal) do j-ésimo grupo, tem o valor esperado μ_j que é dada por

O modelo de curva de crescimento é aquele dado pela expressão (2.59), onde a matriz A (de planejamento) é constituída de uns e zeros.

O objetivo focalizado por Grizzle e Allem messe trabalho foi estudar os estimadores dos parâmetros e testar hipóteses no modelo de curva de crescimento (2.59)

$$\underline{X}' = \underline{B}' \underline{\xi}' \underline{A}' + \underline{E}'$$
pxN pxq qxr rxN pxN

Eles exemplificaram esse modelo, comparando quatro tratamentos, envolvendo estudos cardíacos em quatro grupos de cães. As matrizes B' e A' adaptadas nesse exemplo são

$$\mathbf{B}^{\circ} =
\begin{bmatrix}
1 & -3 & 5 & -1 \\
1 & -2 & 0 & 1 \\
1 & -1 & -3 & 1 \\
1 & 0 & -4 & 0 \\
1 & 1 & -3 & -1 \\
1 & 2 & 0 & -1 \\
1 & 3 & 5 & 1
\end{bmatrix}$$

$$\Lambda^{\circ} =
\begin{cases}
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\
0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1
\end{cases}$$

Os nove uns na primeira linha correspondem ao primeiro grupo, os dez uns da segunda linha ao segundo grupo etc...

Exemplo 2.5 - Na Universidade da Carolina do Norte (E.U.A), foi feita uma pesquisa em dois grupos de crianças, onde os pesquisadores colheram dados em 11 meninas e 16 meninos, com idades de 8, 10, 12 e 14 anos. Estes dados são as distâncias medidas entre o centro de pituitária e a fissura ptério - maxilar, em mm.

O problema é verificar o comportamento das variáveis ao longo destas idades e comparar as curvas associadas a cada grupo. (Exemplo apresentado por Potthoff e Roy, 1964. Ver Morrison, 1976 e Singer, 1977).

O modelo de curva de crescimento é aquele dado por (2.58), ou seja,

$$X = A \quad \xi \quad B \quad + E \quad , \text{ onde}$$
27x4 27x2 2x2 2x4 27x4

$$A = \begin{bmatrix}
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
1 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

$$A = \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}$$

Significa dizer que a curva de crescimento associada ao s-ésimo grupo, s = 1, 2 é dada por

$$E\left(X_{slj}\right) = \xi_{so} + \xi_{s1} t \qquad (2.65)$$

Um outro modelo que pode ser usado é

$$E\left(X_{s\ell j}\right) = \xi_{so}^* + \xi_{s1}^*(t-\overline{t}) ; \qquad (2.66)$$

através desta transformação, temos uma nova matriz B* que é formada por coeficientes de polinômios ortogonais. Aqui,

$$\mathbf{B}^{*} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ & & & \\ -3 & -1 & 1 & 3 \end{pmatrix} \mathbf{e} \quad \boldsymbol{\xi}^{*} = \begin{pmatrix} \boldsymbol{\xi}_{10}^{*} & \boldsymbol{\xi}_{11}^{*} \\ \boldsymbol{\xi}_{20}^{*} & \boldsymbol{\xi}_{21}^{*} \end{pmatrix}$$

Com estas novas informações, e usando a expressão (2.60), obtemos

1° - a)
$$\xi^* = \begin{cases} 22,6659 & 24,9373 \\ 0,4762 & 0,8255 \end{cases}$$
 e

b) as médias previstas (2.66), segundo quadro abaixo

| IDADE | MEDIAS PREVISTAS | | MEDIAS | |
|-------|------------------|---------|---------|---------|
| | MENINAS | MENINOS | MENINAS | MENINOS |
| . 08 | . 21,24 | 22,46 | 21,18 | 22,88 |
| 10 | 22,19 | 24,11 | 22,23 | 23,81 |
| 12 | 23,14 | 25,76 | 23,09 | 25,72 |
| 14 | 24,09 | 27,41 | 24,09 | 27,47 |

2º - a) Testando a hipótese

$$H_{0}: \left[0,0\right] = \underbrace{F}_{\xi} \, \underline{\xi}^{*} \, \underline{C}^{*} = \left[\xi_{11}^{*} \quad \xi_{21}^{*}\right]$$

$$contra$$

$$H_{1}: \left[\xi_{11}^{*} \quad \xi_{21}^{*}\right] \neq \left[0,0\right], \text{ onde}$$

$$\underbrace{F}_{\xi} = \left[0,1\right] \quad e \quad \underline{C}^{*} = \underline{I}_{2}.$$

verificamos que H_o é não rejeitada ao nível α razoável (ver Morrison, 1976 e Singer, 1977).

b) Testando a hipótese

$$H_0: 0 = F_{\xi}' * C' = \xi_{11}^* - \xi_{21}^*,$$
contra

$$H_1: \xi_{11}^* - \xi_{21}^* \neq 0$$
, onde

 $F = \begin{pmatrix} 0,1 \end{pmatrix}$ e $C = \begin{pmatrix} 1 & -1 \end{pmatrix}$, verificamos que H_o e rejeitada ao nível $\alpha = 0.025$ (ver Morrison, 1976 e Singer, 1977).

CAPITULO 3

MODELO LINEAR LATENTE MULTIVARIADO

3.1 - INTRODUÇÃO

Rodrigues, 1975, propôs uma generalização do modelo de análise fatorial, onde os fatores satisfazem um modelo linear multivariado. O modelo que resulta desta combinação, recebeu o nome de "modelo linear latente multivariado", o qual também pode ser chamado de modelo linear-fatorial composto. O modelo apresenta dois casos, não restrito e restrito.

3.2 - APRESENTAÇÃO DO MODELO LINEAR

Seja o modelo linear

$$\underline{Y} = \underline{\Xi} \underline{A} + \underline{\varepsilon} \tag{3.1}$$

onde Y é uma matriz estocástica (qxn) de n observações e q variáveis, Ξ matriz (qxr) de parâmetros regressão desconhecidas, Λ é uma matriz (rxn) de planejamento, de posto linha completo r<n e ε uma matriz estocástica de erros com $E\left[\varepsilon\right]=0$ e $Var\left[\varepsilon\right]=\frac{\Phi}{2}\otimes I_n$, onde Φ é uma matriz p.d de variâncias e covariâncias desconhecidas. Quando estimações e testes forem feitos, tomaremos ε $\sim N_{qxn}\left[0\right]$, Φ Φ I_n . Supomos que Y não seja observável. Por outro lado, uma matriz de dados X(pxn) é observada, com p>q e que esteja ligada a Y através de um modelo de análise fatorial

$$X = \Lambda Y + Z \tag{3.2}$$

onde Λ é a matriz (pxq) das cargas desconhecidas, de posto coluna completo q<p. Z é uma matriz (pxn) de erros, tal que $E\left(\frac{Z}{Z}\right) = 0$ e $Var\left(\frac{Z}{Z}\right) = \frac{\Psi}{2} \otimes \frac{I}{1}$, onde $\frac{\Psi}{2} = \operatorname{diag}(\Psi_1), i=1,2,...,p$ é uma matriz diagonal cujos elementos são desconhecidos, e $Cov\left(\frac{Y}{2},\frac{Z}{2}\right) = 0$. Quando estimações e testes forem requeridos tomaremos $\frac{Z}{2} \sim N_{pxn}\left(\frac{O}{2},\frac{\Psi}{2}\otimes \frac{I}{2}\right)$.

Em (3.2), Y é uma matriz que representa as variáveis latentes ou fatores, enquanto a matriz X representa as variáveis veis respostas.

Estamos interessados em estimar os parâmetros Ξ e Φ do modelo linear (3.1) e testar hipóteses com respeito a Ξ atra vés da matriz de dados X. Por outro lado, as estimações dos parâmetros Λ e Ψ do modelo fatorial também são requeridas.

3.3 - MODELO LINEAR LATENTE MULTIVARIADO

A combinação do modelo (3.1) com o (3.2), resulta o modelo em forma geral dado por

$$X = \beta A + U$$
 (3.3)
 $pxn pxr rxn pxn$

onde \underline{U} é uma matriz estocástica com $E\left(\underline{U}\right) = \underline{0}$ e $Var\left(\underline{U}\right) = \underline{V} \otimes \underline{I}_{n}$,

$$\beta = \Lambda \quad \Xi \quad e \qquad (3.4)$$

$$\rho xr \quad \rho xq \quad q xr$$

$$\underline{V} = \underline{\Lambda} \quad \Phi \quad \underline{\Lambda}^{\circ} + \underline{\Psi} \qquad (3.5)$$

$$\underline{pxp} \quad \underline{pxq} \quad qxq \quad qxp \quad pxp$$

Sob as suposições que ε e \overline{Z} são normais multivariadas, $\underline{U} \sim N_{pxn} \left(\underline{O}, \underline{V} \otimes \underline{I}_{n} \right)$.

A equação (3.3) representa um modelo linear multivaria do, cujas variáveis independentes são não observáveis. Por este motivo é chamado "modelo linear latente multivariado". Os parâmetros Ξ , Λ , Φ e Ψ são chamados parâmetros estruturais, enquanto, β e V são denomiandos de forma reduzida.

3.4 - RESTRIÇÕES NOS PARÂMETROS ESTRUTURAIS

Seja L uma matriz não singular de ordem q, e sejam $\underline{E}^* = \underline{L} \underline{E}$, $\underline{\Lambda}^* = \underline{\Lambda} \underline{L}^{-1}$ e $\underline{\Phi}^* = \underline{L} \underline{\Phi} \underline{L}^{\prime}$. (3.6) Sem modificar a estrutura do modelo (3.3), as matrizes \underline{E} , $\underline{\Lambda}$ e $\underline{\Phi}$ podem ser substituídas por \underline{E}^* , $\underline{\Lambda}^*$ e $\underline{\Phi}^*$. Esta indetermina ção do modelo corresponde à transformação linear dos fatores $\underline{Y}^* = \underline{L} \underline{Y}$. Uma vez que \underline{L} tem \underline{q}^2 elementos, então no mínimo \underline{q}^2 restrições são impostas a \underline{E} , $\underline{\Lambda}$ e $\underline{\Phi}$. No caso não restrito, $\underline{1}$ q(q-1) restrições serão impostas, fazendo $\underline{\Phi} = \underline{I}_q$, significa dizer que a expressão (3.5) toma a forma

$$\underline{V} = \underline{\Lambda} \qquad \underline{\Lambda}' + \underline{\Psi} , \qquad (3.7)$$

$$pxp pxq qxp pxp$$

Esta imposição pode ser feita sen perda de generalidade pois, se p # Ig, através da transformação Y* = L Y, onde L = = $\frac{1}{2}$ verifica-se que (3.7) permanece a mesma. Para q>1, qualquer que seja a matriz ortogonal M(qxq) (que tem $\frac{1}{2}q(q-1)$ elementos livres), as transformações em E e A $\Xi^* = M'\Xi$ e $\Lambda^* = M \Lambda$ são válidas, pois, as expressões (3.3), (3.4) e (3.7), permanecem as mesmas. Como na análise fatorial, podemos desejar que Λ'Ψ⁻¹Λ seja diagonal e que seus elementos estejam em ordem decrescente de valor, com o objeti vo de determinar uma matriz M(qxq) ortogonal, cujos vetores colunas são vetores proprios ortogonais de M'V-1M, envolvendo assim, um problema de redução canônica do modelo fatorial. Usando o processo (algoritimo) de Gram-Schmidt, podemos terminar a matriz ortogonal M(qxq), quando igualamos a os primeiros (j-1) elementos da j-esima coluna de Λ , j = 1, 2, ..., q.

No caso restrito temos restrições sobre Φ fazendo diag $\left[\Phi\right] = \frac{1}{2}$, sobre Λ um número mínimo de Φ (q-1) restrições são impostas tomando (q-1) elementos nulos em cada coluna. Em termos gerais, Λ Φ e Ψ deverão seguir os seguintes casos:

- 1°) parâmetros conhecidos, fixados
- 2°) restrições feitas através de igualdades de parâmetros no modelo;
- 3°) parâmetros não restritos, livres.

O modelo fatorial (3.2) não contêm parâmetro de locação µ, entretanto, para algumas aplicações, este parâmetro é requisitado. Para colocação deste parâmetro, modificações na matriz de planejamento devem aparecer (ver Rodrigues, 1975). Alguns modelos podem ser vistos como casos particulares do modelo apresentado. Em particular, temos o modelo de componentes de variância, dado por

$$x_{ij} = \alpha + \beta_j + s_i + z_{ij}$$
, (3.8)

onde x_{ij} é a medida designada ao i-ésimo tratamento pelo j-ésimo nível do fator B, β_j é o j-ésimo efeito fixo do fator B, $j=1,2\cdots p$, com $\sum_{j=1}^{n}\beta_j=0$, β_j é o efeito aleatório devido ao i-ésimo tratamento, α é a media geral, α_s^2 é a variância do efeito β_s^2 é a variância de β_s^2 é a variância do efeito β_s^2 é a variância de β_s^2 é a variâ

O modelo (3.8) pode ser apresentado na forma

$$x_{ij} = \mu_j + \lambda y_i + z_{ij}$$
 (3.9)

onde $\mu_j = \alpha + \beta_j$, $\lambda = \sigma_s$, $y_i = \frac{s_i}{\sigma_s}$ e $Var(z_{ij}) = \Psi = \sigma_e^2$. Expressando (3.9) em forma matricial, temos

$$\underline{X} = \underline{\mu} + \lambda \underline{1}_{p} \underline{Y} + \underline{Z} , \qquad (3.10)$$

com $E\left(y\right) = 0$, $Var\left(y\right) = 1$, $E\left(Z\right) = 0$, $Var\left(Z\right) = \Psi \cdot \underline{I}_{p}$, $Cov\left(y \cdot Z\right) = 0$, $E\left(X\right) = \mu$ e $Var\left(X\right) = V = \lambda \cdot \underline{I}_{p} \cdot \underline{I}_{p}' + \Psi \cdot \underline{I}_{p}$, onde $\underline{I}_{p}' = \begin{bmatrix} 1, 1, \dots, 1 \end{bmatrix}$. Assim (3.10) appresentado como caso particular de (3.2) com um único fator, onde $\underline{\Lambda} = \lambda^{2}$ e $\underline{\Psi} = \Psi \cdot \underline{I}_{p}$.

3.5 - ESTIMAÇÃO DE MÁXIMA VEROSSIMILHANÇA

Agora apresentamos a função de verossimilhança, que é dada por

$$\log L = -\frac{1}{2} \operatorname{np} \log 2\pi - \frac{1}{2} \operatorname{n} \log(|\underline{V}|) - \frac{1}{2} \operatorname{tr} \left(\underline{V}^{-1} \left(\underline{X} - \underline{\beta} \underline{A} \right) \left(\underline{X} - \underline{\beta} \underline{A} \right)' \right)$$
(3.11)

Para facilitar mais o trabalho, uma função

$$\underline{T} = \frac{1}{n} \left((\underline{X} - \underline{\beta} \underline{A}) (\underline{X} - \underline{\beta} \underline{A})' \right), \qquad (3.12)$$

é dada para que a função F seja definida da seguinte forma:

$$F = \log(|\underline{V}|) + tr(\underline{V}^{-1} \underline{T})$$
 (3.13)

Como V é dada por (3.5) e β por (3.4), então $\underline{T} = \underline{T}(\underline{\Xi})$ e $F = F(\underline{\Xi}, \underline{\Lambda}, \underline{\Phi}, \underline{\Psi})$.

Com o mesmo objetivo de análise fatorial, aqui determinaremos os valores de E, A, A e Y que minimizam F. Estes valo
res passarão ter a notação E, A, A e Y.

B - Estimação de E

Trataremos do caso onde Ξ é <u>não restrito</u>. Determinaremos Ξ que minimiza F para Λ , Φ e Ψ fixados.

Teorema 3.1 - Seja F definida por (3.13). Se Ξ é irrestrita, o valor de Ξ que minimiza F, condicionado aos valores fixados de Λ , Φ e Ψ , é dado por

$$\tilde{\Xi} = (\Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} X \Lambda (\Lambda \Lambda')^{-1}$$
(3.14)

Demonstração:

Pelo teorema (1.12), temos que

 $\frac{\partial F}{\partial \Xi} = -\frac{2}{n} \wedge V^{-1} \left(X - \beta \wedge A \right) \wedge A'; \text{ então, fazendo } \frac{\partial F}{\partial \Xi} = 0 \text{ temos}$ $\wedge V^{-1} \times A' = \wedge V^{-1} \beta \wedge A', \text{ ou seja.}$

$$\Lambda' \quad V^{-1} \quad X \quad \Lambda' = (\Lambda' \quad V^{-1} \quad \Lambda) \equiv (\Lambda \quad \Lambda'), \quad (3.15)$$

onde $\beta = \Lambda \Xi$.

Em (3.3) afirmamos que A tem posto linha completo e que Lem posto coluna completo. Logo, A A' e A' V⁻¹ A são inversiveis e o valor de E que e solução de (3.15) e dado por

$$\tilde{\Xi} = (\tilde{V}, \tilde{\Lambda}_{-1}, \tilde{V})_{-1}, \tilde{V}, \tilde{\Lambda}_{-1}, \tilde{X}, \tilde{V}_{-1}, \tilde{X}, \tilde{V}, (\tilde{V}, \tilde{V}_{-1})_{-1}.$$

C - Estimação de Λ, φ e Ψ

Apesar das expressões analíticas de $\frac{\partial F}{\partial \Lambda}$, $\frac{\partial F}{\partial \phi}$ e $\frac{\partial F}{\partial \mathcal{V}}$ existirem, não existem soluções analíticas onde estas derivadas se anulam. Em (B) encontramos a expressão analítica para Ξ que minimiza F, para Λ , ϕ e Ψ fixados. Substituindo Ξ em F, esta toma a seguinte forma:

$$\vec{F} = \vec{F}(\Lambda, \Phi, \Psi) = \min_{\Xi} F(\Xi, \Lambda, \Phi, \Psi) = F(\Xi, \Lambda, \Phi, \Psi), \qquad (3.16)$$

ou seja,
$$\tilde{F} = \log(|V|) + \operatorname{tr}(V^{-1}\tilde{I}),$$
 (3.17)

onde T e T calculado em E = E

Munidas das expressões analíticas, \tilde{F} , $\frac{\partial \tilde{F}}{\partial \Phi_S}$ e $\frac{\partial \tilde{F}}{\partial \Psi_{\alpha}}$, o processo de Fletcher e Powell (1963), é aplicado.

Seja y um vetor linha com $(p+q)^2 + pq$ componentes que são constituídas pelos elementos de Λ , os elementos da parte triângular superior (inclusive os da diagonal) de Φ e os elementos da diagonal de Ψ . O vetor y tem a seguinte forma

$$\mathbf{y} = (\lambda_{11}, \dots, \lambda_{p1}, \dots, \lambda_{pq}, \lambda_{11}, \lambda_{12}, \dots, \lambda_{pq}, \lambda_{1q}, \dots, \lambda_{pq}, \lambda_{pq},$$

Os parâmetros livres de Λ , Φ e Ψ formam um vetor

$$\underline{\theta} = (\theta_1, \theta_2, \cdots, \theta_m). \tag{3.19}$$

Assim F deve ser encarada como uma função de 0, isto é, F = f(0). As derivadas de f(0), são

$$g_{i} = \frac{\partial f(\underline{\theta})}{\partial \theta_{i}} = \sum_{j} \alpha_{ij} \frac{\partial \tilde{F}}{\partial y_{j}}, \qquad (3.20)$$

onde $\alpha_{ij} = 1$ se $\theta_i = y_j$ e $\alpha_{ij} = 0$ em caso contrário.

Para iniciar o processo, a matriz E pode ser a identidade ou G^{-1} , onde $G = E\left(\frac{\partial^2 f}{\partial \theta \partial \theta'}\right)$. Este processo de minimização de F é calculado na região onde $\Psi_i > \epsilon$, $i=1,2,\ldots$, $p \in \epsilon > 0$.

3.6 - CONSISTÊNCIA E DISTRIBUIÇÃO ASSINTÓTICA DE $\overline{\beta}_n$ e \overline{V}_n

As propriedades de convergência em probabilidade em distribuição são largamente empregadas nos estudos dos estimadores para grandes amostras. Um conceito também bastante utilizado é o de informação de Fisher I(0).

Considere um modelo linear multivariado

$$X_{\mathbf{n}} = \beta A_{\mathbf{n}} + U_{\mathbf{n}}, \qquad (3.21)$$

onde An é de posto linha completo r<n e os vetores colunas de Un são independentes e idênticamente distribuídas (i.i.d) com média O e matriz de covariância V.

Definição: Os estimadores de mínimos quadrados de β e V são definidos por

$$\overline{\underline{g}}_{n} = \underline{X}_{n} \underline{A}_{n}^{\dagger} (\underline{A}_{n} \underline{A}_{n}^{\dagger})^{-1} e$$

$$\overline{\underline{V}}_{n} = \frac{1}{n} \underline{X}_{n} (\underline{I}_{n} - \underline{A}_{n}^{\dagger} (\underline{A}_{n} \underline{A}_{n}^{\dagger})^{-1} \underline{A}_{n}) \underline{X}_{n}^{\dagger} .$$
(3.22)

 $\overline{\beta}_n$ e \overline{V}_n poderão ser estimadores de maxima verossimilhança

$$(M \cdot V)$$
, se $\underline{U}_n \sim N_{pxn} \left(\underbrace{0}_{n}, \underbrace{V} \otimes \underline{I}_{n} \right)$.

Vamos estabelecer a seguinte Hipótese,

Hipótese 3.1 - Seja (Λ_n) uma sequência de matrizes (rxn) com valores fixos conhecidos. Definimos $Q_n = \frac{1}{n} \Lambda_n \Lambda_n'$. Admitamos que

$$\lim_{n\to\infty} Q_n = Q \tag{3.23}$$

tal que, Q existe e seja finita.

A consistência de $\overline{\beta}_n$ e \overline{V}_n pode ser esclarecida pelo seguinte Lema.

Lema 3.1: Seja (X_n) uma sequência de matrizes que satisfazem ao modelo linear (3.21), onde A_n $(n = r+1, r+2\cdots)$ é a matriz de planejamento. Se a sequência (A_n) satisfaz a hipótese 3.1, então quando $n \longrightarrow \infty$

a)
$$\frac{\overline{\beta}}{n} \xrightarrow{P} \underline{\beta}$$

b)
$$\overline{V}_n \xrightarrow{p} \underline{V}$$

Demonstração:

a) De (3.21) temos que
$$\underline{X}_n = \underline{\beta} \ \underline{\Lambda}_n + \underline{U}_n$$
, onde $E\left(\underline{U}_n\right) = \underline{O}$.
Então, $E\left(\underline{X}_n\right) = \underline{\beta} \ \underline{\Lambda}_n \longrightarrow E\left(\overline{\underline{\beta}}_n\right) = E\left(\underline{X}_n \ \underline{\Lambda}_n' \ (\underline{\Lambda}_n \ \underline{\Lambda}_n')^{-1}\right)$

$$\longrightarrow E\left(\overline{\underline{\beta}}_n\right) = \left(E\left(\underline{X}_n\right)\underline{\Lambda}_n' (\underline{\Lambda}_n \underline{\Lambda}_n')^{-1}\right) = \underline{\beta} (\underline{\Lambda}_n \underline{\Lambda}_n') (\underline{\Lambda}_n \underline{\Lambda}_n')^{-1} = \underline{\beta}.$$

Sabenos que
$$\operatorname{Var}\left(\overline{\beta}_{n}\right) = \underline{V} \otimes \left(\underline{\Lambda}_{n}\underline{\Lambda}_{n}^{\prime}\right)^{-1} = \left(\frac{1}{n}\underline{V}\right) \otimes \underline{Q}_{n}^{-1} \qquad (3.24)$$

(ver Anderson. 1958, pag 182).

Pela hipótese 3.1 lim
$$Q_n = Q \implies \lim_{n \to \infty} Var \left(\frac{\overline{\beta}}{\beta_n} \right) = Q \quad (3.25)$$

Como $\overline{\beta}_n$ é un estimador não viciado de β e como $\mathrm{Var}\left(\overline{\beta}_n\right)$ satisfaz (3.25), $\Longrightarrow \overline{\beta}_n$ converge em nédia quadrática para β , logo por 1.85, segue-se que $\overline{\beta}_n$ converge em probabilidade para β .

b)
$$\overline{V}_{n} = \frac{1}{n} \left(X_{n} X_{n}' - \overline{\beta}_{n} (A_{n} A_{n}') \overline{\beta}_{n} \right) e$$

$$\underline{T}_{n} = \frac{1}{n} \left(X_{n} - \beta A_{n} \right) (X_{n} - \beta A_{n})' \Longrightarrow$$

$$\overline{V}_{n} = \underline{T}_{n} - (\overline{\beta}_{n} - \beta) Q_{n} (\beta_{n} - \beta_{n})' \qquad (3.26)$$

A matriz T_n pode ser representada na forma

 $\underline{T}_{n} = \frac{1}{n} \underline{U}_{n} \underline{U}_{n}^{*} = \frac{1}{n} \sum_{k=1}^{n} \underline{u}_{k} \underline{u}_{k}^{*}, \text{ onde } \underline{u}_{k} \underline{u}_{k}^{*} \text{ são matrizes i.i.d com}$ $E\left(\underline{u}_{k} \underline{u}_{k}^{*}\right) = \underline{V}, \quad V \quad k = 1, 2 \cdots n$

Como $\underline{Y}_k = \underline{u}_k u_k^*$, $k=1,2,\cdots$ n são i.i.d e $E\left(\underline{Y}_k\right) = \underline{V} = (\sigma_{ij})$ com $\sigma_{ij} < \infty$, então pelo Teorema de Khinchine,

$$\underline{T}_{n} = \frac{1}{n} \sum_{k=1}^{n} Y_{k} \xrightarrow{p} \underline{V}, \text{ quando } n \rightarrow \infty$$
 (3.27)

(ver Rao 1973, pag 112).

Por outro 1ado, $\overline{\beta}_n \xrightarrow{P} \underline{\beta}$, $\underline{Q}_n \xrightarrow{} \underline{Q}$ quando $n \xrightarrow{} \underline{\circ}$ e $(\overline{\beta}_n - \underline{\beta})$ $\underline{Q}_n(\overline{\beta}_n - \underline{\beta})$ expressa uma função quadrática de $\overline{\beta}_n$ que é uma função continua de $\overline{\beta}_n$, então

$$(\underline{\beta}_n - \underline{\beta}) \ \underline{Q}_n(\underline{\beta}_n - \underline{\beta}) \xrightarrow{P} \underline{Q}, \text{ quando } n \rightarrow \infty$$
 (3.28)
(Ver Rao 1973, (xiii) pag 124).

Assim, da expressão de \overline{V}_n de (3.26) e dos resultados (3.27) e (3.28), concluímos que

$$\overline{V}_n \xrightarrow{P} V$$
 , quando $n \longrightarrow \infty$

Lema 3.2. Seja (\underline{X}_n) uma sequência de matrizes satisfazendo o modelo (3.21), onde \underline{A}_n é a matriz de planejamento e $\underline{U}_n \sim N_{\text{pxn}} \left[\underline{O} , \underline{V} \otimes \underline{I}_n \right]$, $(n = r + 1, r + 2 \cdots)$. Supomos que (\underline{A}_n) satisfaça a Hipôtese 3.1; então,

a)
$$\sqrt{n} \left(\overline{\beta} - \underline{\beta} \right) \xrightarrow{d} N_{pxr} \left(\underline{0}, \underline{v} \otimes \underline{0}^{-1} \right)$$
, quando $n \longrightarrow \infty$;

b) independentemente de $\overline{\beta}$

$$\sqrt{n} \left(\overline{\underline{V}}_n - \underline{V} \right) \xrightarrow{d} N \left(\underline{0}, (v_{ik} \ v_{je}^{+v} ie^{v}_{jk}) \right), \text{ quando } n \longrightarrow \infty,$$
onde $(v_{ik} \ v_{je}^{+v} ie^{v}_{jk})$ é a covariância assintótica do (i,j)
-ésimo e (k,e) -ésimo elemento de $\sqrt{n} \left(\overline{\underline{V}}_n - \underline{V} \right).$

Demonstração:

a) podemos afirmar que, se $\underline{U}_{n} \sim N_{pxn} \left(\begin{array}{c} \underline{0} \end{array}, \underline{V} \otimes \underline{I}_{n} \right)$ então $\overline{\beta}_{n} \sim N_{pxr} \left(\begin{array}{c} \underline{\beta} \end{array}, \underline{V} \otimes (\underline{A}_{n} \ \underline{A}_{n}')^{-1} \right) \left(\begin{array}{c} \text{ver Anderson 1965}, \ \underline{Te} \end{array} \right)$ orema 8.2.2, pag. 183 . $\overline{\beta}_{n} \sim N_{pxr} \left(\begin{array}{c} \underline{\beta} \end{array}, \underline{V} \otimes (\underline{A}_{n} \ \underline{A}_{n}')^{-1} \right) \longrightarrow \left(\begin{array}{c} \overline{\beta}_{n} - \underline{\beta} \end{array} \right) \sim$ $N_{pxr} \left(\begin{array}{c} \underline{0} \end{array}, \underline{V} \otimes (\underline{A}_{n} \ \underline{A}_{n})^{-1} \right) \longrightarrow Z_{n} = \sqrt{n} \left(\begin{array}{c} \underline{\beta}_{n} - \underline{\beta} \end{array} \right) \sim$ $N_{pxr} \left(\begin{array}{c} \underline{0} \end{array}, \underline{V} \otimes (\underline{A}_{n} \ \underline{A}_{n})^{-1} \right) \longrightarrow Z_{n} = \sqrt{n} \left(\begin{array}{c} \underline{\beta}_{n} - \underline{\beta} \end{array} \right) \sim$ $\sim N_{pxr} \left(\begin{array}{c} \underline{0} \end{array}, \underline{V} \otimes (\underline{A}_{n} \ \underline{A}_{n})^{-1} \right) \longrightarrow \overline{V}_{Zn} \left(\begin{array}{c} \underline{t} \end{array} \right) = e^{-\frac{1}{2}} \underline{t}' \left(\begin{array}{c} \underline{V} \otimes \underline{Q}_{n}^{-1} \end{array} \right) t$

Pela hipótese 3.1, $\lim_{n\to\infty} Q_m = Q \implies \lim_{n\to\infty} \Psi_{Zn}(t) =$

$$= e^{-\frac{1}{2} \underline{t}' \left(\underline{v} \otimes \underline{Q}^{-1} \right) \underline{t}} \longrightarrow \sqrt{n} \left(\underline{\beta}_{n} - \underline{\beta} \right) \xrightarrow{d} N_{pxr} \left(\underline{Q}, \underline{v} \otimes \underline{Q}^{-1} \right).$$

b) Temos que

$$\sqrt{n} (\overline{V}_n - \overline{V}) = \sqrt{n} \left(\underline{T}_n - \underline{V} \right) - \sqrt{n} (\overline{\beta}_n - \underline{\beta}) \underline{Q}_n (\overline{\beta}_n - \underline{\beta})$$
 (3.29)

Vimos que $\underline{T}_n = \frac{1}{n} \sum_{k=1}^{n} \underline{u}_k \underline{u}_k$ com $E\left(\underline{u}_k \underline{u}_k^*\right) = \underline{V}$ e covariância dada por

$$E\left(\left(u_{i\alpha}u_{j\alpha}-v_{ij}\right)\left(u_{k\alpha}u_{e\alpha}-v_{ke}\right)\right)=v_{ik}v_{je}+v_{ie}v_{jk}$$
 (ver

Anderson 1958, pag 39).

Pelo Teorema de Lindberg - Levy (T.L.C),

$$\sqrt{n} \left(\underline{T}_{n} - \underline{V} \right) \xrightarrow{d} N \left[\underline{0}, \left(\left(v_{ik} \ v_{je} + v_{ie} \ v_{jk} \right) \right) \right]. \tag{3.30}$$

Seja $\underline{S}_n = n^{\frac{1}{4}} (\underline{\overline{g}}_n - \underline{\beta})$ onde $Var \left(\underline{S}_n \right) = \frac{1}{\sqrt{n}} \left(\underline{V} \otimes \underline{Q}_n^{-1} \right)$, tal que

 $\lim_{n\to\infty} \operatorname{Var}\left(S_n\right) = 0 \Longrightarrow S_n \xrightarrow{P} 0, \text{ quando } n\to\infty. \quad \text{Mas}$

$$n^{\frac{1}{2}}(\overline{\beta}_{n} - \underline{\beta})Q_{n}(\overline{\beta}_{n} - \underline{\beta})' = \left(n^{\frac{1}{4}}(\overline{\beta}_{n} - \beta)\right)Q_{n}\left(n^{\frac{1}{4}}(\overline{\beta}_{n} - \underline{\beta})'\right) =$$

= $\underline{S}_n \ \underline{Q} \ \underline{S}_n'$ expressa uma forma quadrática de \underline{S}_n que é uma função continua de \underline{S}_n . Logo, $\sqrt{n} \ (\underline{\beta}_n - \underline{\beta}) \ \underline{Q}_n \ (\underline{\beta}_n - \underline{\beta})' \xrightarrow{P} \underline{0}$,

quando $n \rightarrow \infty \longrightarrow \sqrt{n} (\underline{\beta}_n - \underline{\beta}) \underline{Q}_n (\overline{\beta}_n - \underline{\beta}) \xrightarrow{d} \underline{Q}$, quando (3.31)

 $n \rightarrow \infty$. Pelas conclusões (3.30) e (3.31), podemos afirmar que $\sqrt{n} \left(\overline{V}_n - V \right) \xrightarrow{d} N \left(0, \left(\left(v_{ik} v_{je} + v_{ie} v_{jk} \right) \right) \right)$.

O lema 3.2 mostrou que os estimadores $\overline{\beta}_n$ e \overline{V}_n são assintóti-

camente normais

B - Consistência de θ_n em um modelo linear estrutural Supomos que X_n seja uma matriz que satisfaz o modelo (3.21), sendo que θ e V são agora funções deriváveis e contínuas de um vetor parâmetro θ . Em (3.13) definimos a função

$$F = log(|V|) + tr(V^{-1}T_n)$$
 (3.32)

e por (3.26) vimos que

$$\underline{T}_{n} = \overline{V}_{n} + (\overline{\beta}_{n} - \beta) Q_{n} (\overline{\beta}_{n} - \beta)$$
(3.33)

Substituindo (3.32) em (3.33) obtemos

$$F = \log(|\underline{v}|) + \operatorname{tr}\left\{\underline{v}^{-1}\left[\underline{\overline{v}}_{n} + (\underline{\overline{\beta}}_{n} - \underline{\beta})\underline{Q}_{n}(\overline{\beta}_{n} - \underline{\beta})'\right]\right\}$$
(3.34)

Através da minimização de (3.32) obtemos um estimador de 0,

$$\hat{\theta}_n \cdot \hat{\theta}_n$$
 será estimador de M.V. de $\hat{\theta}$ se $\hat{U}_n \circ N_{pxn} \mid \hat{Q}, \hat{V} \otimes \hat{I}_n \mid$.

Teorema 3.2. Seja \underline{X}_n satisfazendo (3.21) com matriz de planejamento \underline{A}_n (n=r+1, r+2,...) e sejam $\underline{\beta}$ e \underline{V} funções continuas e deriváveis de um parâmetro estrutural $\underline{0}$, e seja $\underline{\hat{0}}_n$ o valor que minimiza (3.32). Se a sequência (\underline{A}_n) satisfaz a hipôtese 3.1 e se $\underline{0}$ e identificado, então

$$\hat{\underline{\theta}}_n \xrightarrow{P} \underline{\theta}$$
, quando $n \to \infty$ (3.35)

Demonstração: (ver Rodrigues 1975, pag's 66-68).

C - O Limite da Matriz Informação de Fisher do Modelo
Linear Estrutural.

Sejam $X_1, X_2, \cdots X_n$ vetores aleatórios independentes tais que X_{α} tem densidade $f_{\alpha}(x; \theta)$, $\alpha = 1, 2, \dots, n$ dependendo de θ .

Definição: A matriz informação de Fisher associada a α-ésima observação é dada por

$$\underline{I}_{\alpha}(\underline{\theta}) = E \left[\begin{array}{cc} \frac{\partial \log f_{\alpha}}{\partial \underline{\theta}} & \frac{\partial \log f_{\alpha}}{\partial \underline{\theta}'} \end{array} \right]. \tag{3.35}$$

Sob certas condições impostas nas densidades $f_{\alpha}(x;0)$ e no espaço dos parâmetros, a matriz informação de Fisher pode ser escrita na forma

$$\underline{I}_{\alpha}(\theta) = E\left(\frac{\partial^{2} \log f_{\alpha}}{\partial \theta \partial \theta}\right)$$
 (3.36)

Definição: A matriz informação de Fisher limite associada a sequência (X_k) de vetores aleatórios independentes é dada por

$$\underline{\underline{I}} \quad (\underline{\theta}) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \underline{\underline{I}}_{k}(\underline{\theta}) , \qquad (3.37)$$

em relação a θ .

Lema 3.3. Seja F uma função definida como em (3.13). Suponhamos que os parâmetros β e V da forma reduzida tenham deri vadas parciais de primeira e segunda ordem relativamente a 9 e que sejam contínuas. Então,

$$\frac{\partial F}{\partial \theta_{i}} = -2 \operatorname{tr} \left[\frac{1}{n} \underbrace{A_{n} (X_{n} - \underline{\beta} A_{n})' \underline{V}^{-1}}_{\partial \theta_{i}} \underbrace{\frac{\partial \underline{\beta}}{\partial \theta_{i}}} \right] + \operatorname{tr} \left[\underbrace{\underline{V}^{-1} (\underline{V} - \underline{T}_{n})}_{\partial \theta_{i}} \underline{V}^{-1} \underbrace{\frac{\partial \underline{V}}{\partial \theta_{i}}}_{\partial \theta_{i}} \right] (3.38)$$

$$e$$

$$\frac{\partial^{2} F}{\partial \theta_{i} \partial \theta_{j}} = 2 \operatorname{tr} \left[\frac{1}{n} \underbrace{A_{n}}_{A_{n}} \underbrace{A_{n}'}_{\partial \theta_{i}} \underbrace{\frac{\partial \underline{\beta}'}{\partial \theta_{i}}}_{\partial \theta_{i}} \underline{V}^{-1} \underbrace{\frac{\partial \underline{\beta}}{\partial \theta_{j}}}_{\partial \theta_{j}} \right] +$$

$$+ 2 \operatorname{tr} \left[\underline{V}^{-1} \underbrace{\frac{\partial \underline{\beta}}{\partial \theta_{i}}}_{\partial \theta_{i}} \cdot \frac{1}{n} \underbrace{A_{n} (X_{n} - \underline{\beta} A_{n})' \underline{V}^{-1} \underbrace{\frac{\partial \underline{V}}{\partial \theta_{j}}}_{\partial \theta_{j}} \right] - 2 \operatorname{tr} \left[\underbrace{\frac{1}{n}}_{A_{n}} \underbrace{A_{n} (X_{n} - \underline{\beta} A_{n})' \underline{V}^{-1} \underbrace{\frac{\partial^{2} \underline{\beta}}{\partial \theta_{i} \partial \theta_{j}}}_{\partial \theta_{j}} \right] - \operatorname{tr} \left[\underbrace{\underline{V}^{-1}}_{\partial \theta_{i}} \underbrace{\underline{V}^{-1}}_{\partial \theta_{i}} \underbrace{\underline{V}^{-1}}_{\partial \theta_{j}} \underbrace{$$

Demonstração: a demonstração é feita usando os resultados de matrizes do primeiro capítulo (ver Rodrigues 1975,pag's 71-73)

Lema 3.4. Seja X_n satisfazendo (3.21) com matriz de planejamento A_n. A sequência (A_n) satisfaz a Hipótese 3.1 e

U_n · N_{pxn} (O, V ⊗ I_n), (n=r+1,r+2,...). Os parâmetros da for ma reduzida β e V têm derivadas de primeira e segunda ordem contínuas em relação ao parâmetro estrutural Θ. Então

$$\underline{\underline{I}}_{n}(\underline{\theta}) \cdot \xrightarrow{p} \underline{\underline{I}}(\underline{\theta})$$
 quando $n \to \infty$,

isto \tilde{e} , $\underline{I}_n(\theta)$ \tilde{e} um estimador consistente de $\underline{I}(\theta)$.

<u>Demonstração</u>: Pelo Teorema (3.2), $\frac{\hat{\theta}}{n} \xrightarrow{P} \underline{\theta}$, quando $n \xrightarrow{\infty}$, e pelo fato de $\frac{\partial \underline{\beta}}{\partial \theta_i}$, $\frac{\partial \underline{V}}{\partial \theta_i}$ e \underline{V}^{-1} serem funções continuas

de θ , temos que

$$\frac{\partial \ \underline{\beta}}{\partial \ \theta_{\mathbf{i}}}(\ \underline{\hat{\theta}}_{\mathbf{n}}) \xrightarrow{P} \frac{\partial \ \underline{\beta}}{\partial \ \theta_{\mathbf{i}}}(\ \underline{\theta}\), \text{ quando } n \to \infty,$$

$$\frac{\partial \ \underline{V}}{\partial \ \theta_{\mathbf{i}}}(\ \underline{\hat{\theta}}_{\mathbf{n}}) \xrightarrow{P} \frac{\partial \ \underline{V}}{\partial \ \theta_{\mathbf{i}}}(\ \underline{\theta}\), \text{ quando } n \to \infty,$$

$$\underline{V}^{-1}(\ \underline{\hat{\theta}}_{\mathbf{n}}) \xrightarrow{P} \underline{V}^{-1}(\ \underline{\theta}\), \text{ quando } n \to \infty.$$

Como os elementos de $\underline{I}(\hat{\theta}_n)$ são funções continuas de $\frac{\partial \hat{\beta}}{\partial \theta_i}$ $\frac{\partial \hat{\beta}}{\partial \theta_i}$ e \underline{V}^{-1} em $\hat{\theta}_n$, então podemos afirmar que $\underline{I}(\hat{\theta}_n) \xrightarrow{P} \underline{I}(\hat{\theta})$, quando $n \to \infty$.

D - Normalidade Assintótica de en no Modelo Linear Estrutural

Estamos de posse de um estimador $\hat{\theta}_n$ de M.V de θ que foi determinado através da equação

$$\frac{\partial \log L(\overline{\beta}_{n}, \overline{V}_{n}, \hat{\theta}_{n})}{\partial \underline{\theta}} = \underline{0}, \text{ ou pela equação}$$

$$\frac{\partial F(\overline{\beta}_{n}, \overline{V}_{n}, \hat{\theta}_{n})}{\partial \underline{\theta}} = \underline{0}.$$

Supomos $\underline{U}_n \sim N_{pxn} \left(\underline{O}, \underline{V} \otimes \underline{I}_n \right)$.

Lema 3.5. Seja F a função definida por (3.32). Sejam ß e V os parâmetros na forma reduzida que têm derivadas de primeira e segunda ordem contínuas em relação a 0. Se (An) satisfaz a Hipótese 3.1, então

$$\frac{1}{2} \xrightarrow{\partial \theta \partial \theta} \xrightarrow{P} \underline{I}(\theta)$$
, quando $n \longrightarrow \infty$

Demonstração:

(ver Rodrigues 1975, pag's 78-79)

Corolario 3.1. Com as condições do Lema 3.5, se $\hat{\theta}_n$ é um vetor tal que $\hat{\theta}_n \xrightarrow{P} \theta$ quando $n \to \infty$, então $\frac{1}{2} \frac{\partial^2 F(\hat{\theta}_n)}{\partial \theta \partial \theta'} \xrightarrow{P} \underline{I}(\theta) , \text{ quando } n \to \infty .$

Teorema 3.3. Seja X_n satisfazendo (3.21) com a matriz de pla nejamento A_n e $U_n \sim N_{pxn}(0, V \otimes I_n)$, $(n=(r+1), (r+2)\cdots)$. Sejam β e V os parâmetros na forma reduzida que têm derivadas de primeira e segunda ordem contínuas em relação ao parâmetro estrutural θ . Se (A_n) satisfaz a Hipótese 3.1, então quando $n \to \infty$, \sqrt{n} $(\hat{\theta}_n - \theta) \to N(0, I^{-1}(\theta))$, onde $\hat{\theta}_n$ $\hat{\theta}_n$ e o estimador de máxima verossimilhança de θ e $I(\theta)$ $\hat{\theta}$ a matriz informação de Fisher.

Demonstração:

(ver Rodrigues 1975, pag's 80-82)

E - Teoria de Grandes Amostras para o Modelo Linear Latente Podemos encarar este item como um caso particular dos anteriores, ou seja, representar o vetor θ como o vetor contendo todos os parâmetros livres de Ξ , Λ , Φ e Ψ e que seu estima dor de M·V $\hat{\theta}_n$ seja aquele determinado pelo processo de Fletcher e Powell (1963)

- a) $\hat{\theta}_n \xrightarrow{P} \theta$, isto $\hat{\theta}_n$ $\hat{\theta}$ um estimador consistente de θ .
- b) $\sqrt{n} \left[\hat{\underline{\theta}}_n \underline{\theta} \right] \xrightarrow{d} N \left[\hat{\underline{0}} , \underline{I}^{-1} (\underline{\theta}) \right]$, onde $\underline{I} (\underline{\theta}) \in a$ matriz informação limite de Fisher.
- c) $\underline{\underline{I}}(\hat{\underline{\theta}}_n) \xrightarrow{\underline{P}} \underline{\underline{I}}(\underline{\theta})$, $\underline{\underline{I}}(\hat{\underline{\theta}}_n)$ $\underline{\underline{e}}$ um estimador consistente de $\underline{\underline{I}}(\underline{\theta})$

Demonstração:

O vetor θ e constituído dos parâmetros livres Ξ , Λ , Φ e Ψ . Como β e V são funções de Ξ , Λ , Φ e Ψ e como têm derivadas de primeira e segunda ordem contínuas \Longrightarrow β e V são funções de θ e têm derivadas de primeira e segunda ordem contínuas em relação a θ , então

- a) Pelo Teorema 3.2, segue-se que $\hat{\theta}_n \xrightarrow{P} \hat{\theta}$, quando $n \rightarrow \infty$
- b) Pelo Teorema 3.3, segue-se que $\sqrt{n} \left[\hat{\underline{\theta}} \underline{\underline{\theta}} \right] \xrightarrow{d} N \left[\underline{\underline{0}} , \underline{\underline{I}}^{-1} (\underline{\underline{\theta}}) \right]$, quando $n \longrightarrow \infty$

c) pelo Lema 3.4, segue-se que

$$\underline{\underline{I}}(\hat{\underline{\theta}}_n) \xrightarrow{P} \underline{\underline{I}}(\underline{\underline{\theta}})$$
, quando $n \longrightarrow \infty$.

3.7 - OBTENÇÃO DA MATRIZ INFORMAÇÃO LIMITE DE FISHER

3.7.1 - No Modelo Linear Estrutural

A obtenção dos elementos da matriz informação de Fisher para o modelo linear estrutural é garantida pela seguinte proposição: Seja X_n satisfazendo (3.21) com matriz de planejamento A_n e $U_n \sim N_{pxn} \left[\begin{array}{c} 0 \end{array}$, $V \otimes I_n \right]$ (n=r+1,r+2···), e suponha que os parâmetros da forma reduzida β e V, ambos em função do parâmetro estrutural θ , sejam continuas em relação a θ e tenham derivadas de primeira e segunda ordem continuas em relação a θ . Se a sequência de matrizes de planejamento (Λ_n) satisfaz a Hipôtese 3.1, então os elementos da matriz imformação limite de Fisher são dadas por

$$I(\theta_{i},\theta_{j}) = tr\left[Q \frac{\partial \underline{\beta'}}{\partial \theta_{i}} \underline{V}^{-1} \frac{\partial \underline{\beta}}{\partial \theta_{j}}\right] + \frac{1}{2} tr\left[\underline{V}^{-1} \frac{\partial \underline{V}}{\partial \theta_{i}} \underline{V}^{-1} \frac{\partial \underline{V}}{\partial \theta_{j}}\right]$$
(3.39)

Demonstração: (ver Rodrigues 1975, pag. 74)

3.7.2 - No Modelo Linear Latente

 matrizes de planejamento (A_n) satisfaz a Hipótese 3.1, então os elementos da matriz informação limite de Fisher com respeito a Ξ , Λ , Φ e Ψ são

1)
$$I(\xi_{ij}, \xi_{ke}) = (\Lambda' V^{-1} \Lambda)_{ik} (Q)_{je}$$
 (3.40)

2)
$$I(\xi_{ij},\lambda_{ke}) = (V^{-1}\Lambda)_{ki} (Q E')_{je}$$
,

3)
$$I(\lambda_{ij}, \lambda_{ke}) = (\underline{V}^{-1})_{ik} (\underline{E} \underline{Q} \underline{E}')_{je} + (\underline{V}^{-1}\underline{\Lambda} \underline{\Phi})_{ie} \cdot (\underline{V}^{-1}\underline{\Lambda} \underline{\Phi})_{ke} + (\underline{V}^{-1})_{ik} (\underline{\Phi} \underline{\Lambda}'\underline{V}^{-1}\underline{\Lambda} \underline{\Phi})_{je}$$

4)
$$I(\lambda_{ij}, \phi_{ke}) = \frac{1}{2}(2-\delta_{ke}) \left((\underline{V}^{-1}\underline{\Lambda})_{ik} (\underline{\Lambda}, \underline{V}^{-1}\underline{\Lambda}, \underline{\phi})_{ej} + (\underline{V}^{-1}\underline{\Lambda})_{ie} (\underline{\Lambda}, \underline{V}^{-1}\underline{\Lambda}, \underline{\phi})_{kj} \right)$$

5)
$$I(\lambda_{ij}, \Psi_k) = (\underline{V}^{-1})_{ik} (\underline{V}^{-1}\underline{\Lambda} \Phi)_{kj}$$

6)
$$I(\phi_{ij},\phi_{ke}) = \frac{1}{4}(2-\delta_{ij})(2-\delta_{ke}) \left((\underline{\Lambda}'\underline{V}^{-1}\underline{\Lambda})_{ik} (\underline{\Lambda}'\underline{V}^{-1}\Lambda)_{je} + (\underline{\Lambda}'\underline{V}^{-1}\underline{\Lambda})_{ie} (\underline{\Lambda}'\underline{V}^{-1}\underline{\Lambda})_{jk} \right)$$

7)
$$I(\phi_{ij}, \Psi_k) = \frac{1}{2}(2-\delta_{ij})(\underline{V}^{-1}\underline{\Lambda})_{ki}(\underline{V}^{-1}\underline{\Lambda})_{kj} e$$

8)
$$I(\Psi_{i}, \Psi_{k}) = \frac{1}{2}(V^{-1})_{ik}^{2}$$

Os elementos restantes são nulos, por exemplo:

$$I(\xi_{ij}, \phi_{ke}) = I(\xi_{ij}, \Psi_k) = \cdots = 0$$

3.8 - TESTES DE HIPOTESES

Os testes de hipótese que desenvolveremos aqui são os seguintes:

a) teste de adaptação do modelo;

- b) teste de hipótese sôbre os parâmetros Λ, φ e Ψ;
- c) teste de hipótese sobre o parâmetro E .

O primeiro teste é o teste de adaptação do modelo linear estrutural, onde β e V são funções do parâmetro estrutural θ . O segundo teste é feito em relação a Λ , Φ e Ψ , ou seja, em função do proprio parâmetro estrutural θ . O terceiro teste é feito envolvendo combinações lineares dos elementos de Ξ , da forma $C \subseteq \beta = 0$, sendo que o modelo aplicado é o modelo linear latente. Seja X_n a matriz dada por

$$X_n = \beta A_n + U_n \tag{3.41}$$

onde \underline{A}_n é de posto linha completo r< n, $\underline{\beta} = \underline{\Lambda} \, \underline{\Xi}$, \underline{V} positiva definida (p.d) e $\underline{U}_n \sim N_{pxn} \left(\underline{O}, \underline{V} \otimes \underline{I}_n \right)$

3.8.1 - Teste de Adaptação do Modelo

Seja A = $\{\text{todas as matrizes } V(pxp), \text{ simetricas e positivas definidas}\}$, e seja

 $B = \{ todas \ as \ matrizes \ \underline{\beta}(pxr) \}$. Definimos:

a)
$$\Omega = B \times A = \{(\underline{\beta}, \underline{V}); \underline{\beta} \in B, \underline{V} \in A\}$$

b)
$$\omega = \{(\underline{\beta}, \underline{V}); \beta \in B, \underline{V} \in A \in \underline{\beta} \in \text{função de } \underline{\theta}\}$$

Testar a adaptação do modelo linear estrutural é o mesmo que testar a Hipótese

$$H_0: (\beta, V) \in \omega$$
, contra (3.42)

$$H_1: (\beta, V) \in \Omega - \omega$$

onde β e V são funções do parâmetro estrutural θ . observe que β tem p·r parâmetros, V tem $\frac{1}{2}$ p(p+1) e θ tem m parâmetros. Então ℓ = p·r + $\frac{1}{2}$ p(p+1) é o número de parâmetros de β e V. Logo o número de restrições impostas por H₀ é (ℓ -m) com m < ℓ . Desejamos agora determinar a estatística teste. Para a realização deste evento, os seguintes fatos são dados:

$$\underline{\underline{\beta}}_{n} = \underline{X}_{n} \underline{A}_{n}^{\prime} (\underline{A}_{n} \underline{A}_{n}^{\prime})^{-1} \quad e$$

$$\frac{\dot{\overline{y}}_n}{\overline{y}_n} = \frac{1}{n} \underline{x}_n \left[\underline{I} - \underline{A}_n' (\underline{A}_n \underline{A}_n')^{-1} \underline{A}_n \right] \underline{x}_n',$$

que são estimadores de máxima verossimilhança não restritos de $\underline{\beta}$ e \underline{V} , onde $\underline{U}_n \sim N_{\mathrm{pxr}} \Big(\underline{0}, \underline{V} \otimes \underline{I}_n \Big)$. Seja $\overline{\underline{\gamma}}_n$ um vetor (ℓ x1) constituído de todos os elementos distintos de $\overline{\beta}_n$ e \overline{V}_n . Logo

$$\max_{\Omega} \log L(\underline{\beta},\underline{V}) = C - \frac{1}{2} F(\overline{\gamma}_n) = C - \frac{1}{2} n \left(\log(|\overline{V}|) + p \right)$$
 (3.43)

$B - Em \omega$

Seja $\hat{\theta}$ o vetor (mx1) que $\hat{\epsilon}$ estimador de máxima verossimilhança de $\hat{\theta}$ • $\hat{\beta}$ e \hat{V}_n são estimadores de máxima verossimilhança de $\hat{\beta}$ e \hat{V} sob a restrição de \hat{H}_o . Logo

$$\hat{\underline{T}}_{n} = \frac{1}{n} (\underline{X}_{n} - \hat{\underline{\beta}}_{n} \underline{A}_{n}) (\underline{X}_{n} - \hat{\underline{\beta}}_{n} \underline{A}_{n}) e$$

$$\max_{\omega} \log L (\underline{\beta}, \underline{V}) = C - \frac{1}{2} F(\hat{\underline{\theta}}_{n}) =$$

$$= C - \frac{1}{2} n \left[\log(|\hat{\underline{V}}| + \operatorname{tr}(\hat{\underline{V}}^{-1} \hat{\underline{T}}) \right]$$
(3.44)

C - Estatistica

 $\lambda_n = \frac{L_\omega}{L_\Omega}$, onde L_ω é a a função de verossimilhança $L(\beta, V)$ sob H_0 e L_Ω sob Ω . Segue-se que

$$-2 \log \lambda_{n} = -2 \left[\log L_{\omega} - \log L_{\Omega} \right] =$$

$$= -2 \left[\max_{\omega} \log L - \max_{\Omega} \log L \right] =$$

$$= n \left[\log(|\widehat{\mathbb{V}}|) + \operatorname{tr}(\widehat{\mathbb{V}}^{-1}|\widehat{\mathbb{T}}) - \log(|\overline{\mathbb{V}}_{n}|) - p \right].$$

Para obter a distribuição de - 2 $\log \lambda_n$, quando n é grande, usamos o seguinte teorema:

Teorema 3.4. Seja X_n satisfazendo (3.41) para '(n = r+1, r+2,...), e sejam ß e V parâmetros na forma reduzida, que são funções continuas em relação a θ e com derivadas parciais de primeira e segunda ordem continuas em relação a θ . Se a sequência de matrizes de planejamento (A_n) satisfaz a Hipótese 3.1, então quando n $\longrightarrow \infty$

$$T = -2 \log \lambda_n \xrightarrow{d} \chi_v^2, \qquad (3.46)$$

onde $v = \ell - m$ e -2 log λ_n é dada por (3.45).

Demonstração: (ver Rodrigues 1975, pag 97).

Concluímos pelo Teorema 3.4 que a região crítica do teste ao nível α é dada por

$$-2 \log \lambda_n > \chi_{v,1-\alpha}^2$$
, (3.47)

ou seja, rejeitamos H nesta região.

3.8.2 - Teste de Hinóteses en Relação a A o e Y .

O vetor $\theta' = (\theta_1, \theta_2 \cdots \theta_n)$ conten todos os parâmetros livres de Λ , Φ e Ψ . Então, testar hipóteses sobre restrições impostas a Λ , Φ e Ψ é o mesmo que testar hipóteses impondo restrições sobre o parâmetro estrutural θ .

Seja

$$II_{0}: \underline{g}(\underline{0}) = \underline{0}$$

$$II_{1}: \underline{g}(\underline{0}) \neq \underline{0}$$

$$(3.48)$$

onde g é uma função vetorial de dimensão s, isto é, $g'(0) = (g_1(0), g_2(0) \cdots g_s(0))$, con $0 \in \Omega \subset \mathbb{R}^m$. Supondo que as g_i , $i=1,2,\ldots,s$, são continuas em relação a $0 \in \mathbb{R}^m$ então as restrições $g_i(0) = 0$ \mathbb{V} $i=1,2,\ldots,s$, impostas por \mathbb{H}_0 , são equivalentes as restrições de colocar as $0 \in \mathbb{R}^m$ $0 \in \mathbb{R}$

 h_j , j=1,2,...,n, são continuas em relação a α e com derivadas parciais de primeira ordem continuas em relação a α , onde $\alpha' = (\alpha_1,\alpha_2,...,\alpha_d)$ (ver Rao 1965, pag's 149-150). Este resultado pode ser verificado, usando-se alguns teoremas de Análise do \mathbb{R}^n (ver Spivak 1965, pag 111).

Seja $\hat{\theta}_n$ o estimador de máxima verossimilhança não restrito de $\hat{\theta}$ e $\hat{\hat{\theta}}_n$ o estimador de máxima verossimilhança de $\hat{\theta}$ sob a restrição de \hat{H}_0 .

O teste da razão de verossimilhança é dado por

$$-2 \log \lambda_{n} = -2 \left(\max_{\omega} \log L - \max_{\Omega} \log L \right)$$

$$= n \left(F(\hat{\theta}_{n}) - F(\hat{\theta}_{n}) \right), \text{ onde}$$
(3.49)

 $\Omega = \left\{ \text{conjunto de todos os parâmetros livres de } \underline{\Lambda} \text{ , } \underline{\Phi} \text{ e } \underline{\Psi}, \text{ a-presentados na forma vetorial } \underline{\Theta}' = (\Theta_1 \cdots \Theta_m) \right\} \text{ e}$

 $\omega = \left\{ \underbrace{\theta} \in \Omega; \ g(\underline{\theta}) = 0 \right\}.$ A estatística teste (-2 log λ_n) tem distribuição limite χ_s^2 , quando $n \longrightarrow \infty$. A verificação deste resultado é vista através dos mesmos argumentos do Teorema 3.4, adaptados para este teste.

3.8.3 - Testes de Hipóteses Lineares sobre E .

O parâmetro E é o parâmetro do modelo linear geral dado em (3.1). Então o teste de Hipóteses sobre E é dado por,

$$H_0: C = \beta = 0$$
, contra (3.50)
 $H_1: C = \beta \neq 0$,

onde C (txq) e B (rxs) são matrizes de valores constantes conhecidas, e têm posto linha completo ($t \le q$) e posto coluna completo ($s \le r$), respectivamente.

A - Teste da Razão de Máxima Verossimilhança

Para o desenvolvimento deste item utilizaremos os seguintes valores:

- a) a função F definida em (3.13),
- b) os estimadores de máxima verossimilhança $\hat{\Xi}_n$, $\hat{\lambda}_n$, $\hat{\Phi}_n$ e $\hat{\Psi}_n$ de $\hat{\Xi}_n$, $\hat{\lambda}_n$, $\hat{\Phi}_n$ e $\hat{\Psi}_n$ respectivamente.

Teorema 3.5. Se Ξ está subordinada a hipótese H_o, então o valor Ξ que minimiza a função F condicionado aos valores fixados de Λ, φ e Ψ é dado por

$$\tilde{\Xi} = \tilde{\Xi} - (\tilde{\Lambda}'\tilde{V}^{-1}\tilde{\Lambda})^{-1} \tilde{C}' \left[\tilde{C}(\tilde{\Lambda}'\tilde{V}^{-1}\tilde{\Lambda})^{-1}\tilde{C}' \right]^{-1} \tilde{C}\tilde{\Xi}\tilde{B} \left[\tilde{B}'(\tilde{\Lambda}\tilde{\Lambda}')^{-1}\tilde{B} \right]^{-1} \beta'(\tilde{\Lambda}\tilde{\Lambda}')^{-1},$$
(3.51)

onde $\tilde{\Xi}$ é o valor que minimiza F condicional aos valores fixos Λ , Φ e Ψ , quando Ξ é não restrito.

Demonstração: Desejamos minimizar a função F em relação a Ξ com valores de Λ, Φ e Ψ fixados, subordinado a restrição C Ξ β = O. Este problema, é um problema de mínimo condicionado, e para a sua solução, usaremos a matriz multiplicador de Lagrange L(sxt), e minimizaremos a função

$$f(\underline{z},\underline{L}) = F(\underline{z}) + tr(\underline{L}' \underline{C} \underline{z} \underline{B}),$$
 (3.52)

com respeito a E e L .

$$\frac{\partial f}{\partial f} = -\frac{2}{3} \nabla, \tilde{\Lambda}_{-1} \left(\tilde{\chi} - \tilde{V} \tilde{\Xi} \tilde{V} \right) \tilde{V}_{+} + \tilde{C}_{1} \tilde{F}_{1} \tilde{B}_{1}$$
 (3.23)

e

$$\frac{\partial f}{\partial L} = C \Xi B . \tag{3.54}$$

Fazendo $\frac{\partial f}{\partial \Xi} = 0$, obtenos

$$(\underline{\Lambda}, \underline{V}^{-1}, \underline{\Lambda}) \stackrel{\underline{\tilde{\Xi}}}{\underline{\tilde{\Xi}}} (\underline{\tilde{\Lambda}}, \underline{\tilde{\Lambda}}) = \underline{\Lambda}, \underline{\tilde{\Lambda}}, \underline{\tilde{\Lambda}}, -\frac{n}{2}, \underline{\tilde{C}}, \underline{\tilde{L}}, \underline{\tilde{B}}, \underline{\tilde{L}}, \underline$$

$$\tilde{\Xi} = \left[\underline{\Lambda}'\underline{V}^{-1}\underline{\Lambda}\right]^{-1} \left[\underline{\Lambda}'\underline{V}^{-1}\underline{X}\ \underline{A}' - \frac{n}{2}\ \underline{C}'\ \underline{L}\ \underline{\beta}'\right] \ (\underline{A}\ \underline{A}')^{-1}, \tag{3.55}$$

mas o estimador de máxima verossimilhança de Ξ não restrito, sujeito a Λ , Φ e Ψ fixos, é dado por

$$\tilde{\Xi} = (\Lambda' V^{-1} \Lambda)^{-1} \Lambda' V^{-1} X \Lambda' (\Lambda \Lambda')^{-1}$$
 (3.56)

Substituindo (3.56) em (3.55), temos

$$\tilde{\Xi} = \tilde{\Xi} - \frac{n}{2} \left(\tilde{\Lambda}' \tilde{V}^{-1} \tilde{\Lambda} \right) C' \tilde{L} B' \left(\tilde{A} \tilde{A}' \right)^{-1}. \tag{3.57}$$

De (3.54) obtemos
$$C = \beta = 0$$
. (3.58)

$$\underline{C} \, \underline{\tilde{E}} \, \underline{B} = \frac{n}{2} \, \underline{C} \left(\underline{\Lambda} \, \underline{V}^{-1} \underline{\Lambda} \right)^{-1} \, \underline{C} \, \underline{L} \, \underline{B} \, \underline{C} \left(\underline{A} \, \underline{A}' \right)^{-1} \, \underline{B} \, , \qquad (3,59)$$

dai obtemos o valor de

$$\widetilde{\underline{L}} = \frac{2}{n} \left[C(\Lambda' V^{-1} \Lambda)^{-1} C' \right]^{-1} C \widetilde{\underline{E}} B \left[B'(A A')^{-1} B \right]^{-1}.$$
 (3.60)

Substituindo (3.60) em (3.57), obtemos

$$\tilde{\Xi} = \tilde{\Xi} - (\tilde{\Lambda}'\tilde{V}^{-1}\tilde{\Lambda})^{-1}C'\left(\tilde{C}(\tilde{\Lambda}'\tilde{V}^{-1}\tilde{\Lambda})^{-1}\tilde{C}'\right)^{-1}\tilde{C}\tilde{\Xi}\tilde{B}\left(\tilde{B}'(\tilde{A}\tilde{A}')^{-1}\tilde{B}\right)\tilde{B}'(\tilde{A}\tilde{A}')^{-1}$$

Desejamos usar aqui o teste da razão de máxima verossimilhan ça, mas para obter este teste, teremos que determinar o esti mador $\hat{\mathbb{E}}$ de máxima verossimilhança não condicional de \mathbb{E} , sob a restrição H_0 . Para isto, procederemos da seguinte maneira : minimizaremos a função $f(\tilde{\mathbb{E}}, \Lambda, \Phi, \Psi)$ em relação às matrizes paramétricas Λ , Φ e Ψ , usando a técnica de Fletcher Powell, em seguida, substituindo estes valores obtidos em $\tilde{\mathbb{E}}$, obtemos $\hat{\mathbb{E}}$. Então, o teste da razão de verossimilhança $\hat{\mathbb{E}}$

$$-2 \log \lambda_{n} = n \left[F\left(\tilde{\Xi}, \tilde{\hat{\Lambda}}, \tilde{\hat{\phi}}, \tilde{\hat{\psi}} \right) - F\left(\tilde{\Xi}, \tilde{\hat{\Lambda}}, \tilde{\hat{\phi}}, \tilde{\hat{\psi}} \right) \right]. \tag{3.61}$$

- 2 log
$$\lambda_n \xrightarrow{d} \chi_v^2$$
, quando $n \longrightarrow \infty$, (3.62) onde $v = t.s$

O teste da razão de verossimilhança envolve diversos cálculos de estimação de E, determinando E, E, Ê e Ê; por isso, uma outra opção para testar a Hipótese (3.50) foi encontrada, a estatística de Wald. O teste de Wald não utiliza os estimadores de máxima verossimilhança restritos, significa dizer, que c um teste mais vantajoso.

B - TESTE DE WALD

Para o desenvolvimento desta parte, necessitamos definir um

vetor formado pelos elementos de uma matriz.

Definição. Seja A uma matriz (mxn); definimos o vet A como sendo um vetor (m.nxl) contendo as colunas de A, colocadas uma sobre a outra da primeira a n-esima, respectivamente.

Como consequência da definição de vet A, pode ser mostrado que, se A, B e C são matrizes tais que o produto A B C existe, então

$$vet(\underline{A} \cdot \underline{B} \cdot \underline{C}) = (\underline{C}' \otimes \underline{A}) \text{ vet } \underline{B}$$
 (3.63)

[ver Neudecker, 1969, resultado(2.10)].

Seja

$$\xi = \text{vet } \Xi = e \qquad (3.64)$$

$$\hat{\xi} = \text{vet } \hat{\Xi} \quad , \tag{3.65}$$

onde $\Xi(qxr)$ é a matriz de parâmetros desconhecidos do modelo linear geral (3.1) e Ξ_n é o estimador de máxima verossimilhança não restrito de Ξ . Então o teste de hipóteses (3.50) pode ser expresso na forma,

$$H_0: (\underline{B}' \otimes \underline{C}) \underline{\xi} = \underline{0}$$
 (3.66)

$$H_1: (B' \otimes C) \notin \emptyset$$

Seja e o vetor de parâmetros estruturais, cujos primeiros q.r elementos são os elementos de g e os restantes, são parâ metros livres de A, e e y, como em (3.18).

A matriz informação limite de Fisher em função de e dada por

$$\underline{\underline{I}}(\underline{\theta}) = \begin{pmatrix} \underline{\underline{I}}_{11}(\underline{\theta}) & \underline{\underline{I}}_{12}(\underline{\theta}) \\ \underline{\underline{I}}_{21}(\underline{\theta}) & \underline{\underline{I}}_{22}(\underline{\theta}) \end{pmatrix}, \qquad (3.67)$$

onde sua inversa I-1(e) e dada por

$$\underline{\mathbf{I}}^{-1}(\underline{\theta}) = \begin{bmatrix} \underline{\mathbf{J}}_{11}(\underline{\theta}) & \underline{\mathbf{J}}_{12}(\underline{\theta}) \\ \underline{\mathbf{J}}_{21}(\underline{\theta}) & \underline{\mathbf{J}}_{22}(\underline{\theta}) \end{bmatrix}, \tag{3.68}$$

.com

$$\underline{J}_{11}(\underline{\theta}) = \left[\underline{I}_{11}(\underline{\theta}) - \underline{I}_{12}(\underline{\theta}) \ \underline{I}_{22}(\underline{\theta}) \cdot \underline{I}_{21}(\underline{\theta})\right]^{-1}, \quad (3.69)$$

(ver Anderson 1976, pag 68).

A estatística teste de Wald (1943) é dada por

$$W_{\mathbf{n}} = \eta \hat{\underline{\xi}}_{\mathbf{n}}^{\prime} (\underline{B} \otimes \underline{C}^{\prime}) \left((\underline{B}^{\prime} \otimes \underline{C}) \ \underline{J}_{11} (\hat{\underline{\theta}}) (\underline{B} \otimes \underline{C}^{\prime}) \right)^{-1} (\underline{B}^{\prime} \otimes \underline{C}) \hat{\underline{\xi}} , \qquad (3.70)$$

onde $\hat{\theta}$ e $J_{11}(\hat{\theta})$ são estimadores de máxima verossimilhança de θ e $J_{11}(\hat{\theta})$, respectivamente.

Para a determinação da distribuição limite, enunciaremos o seguinte teorema:

tros livres. Se a sequência (An) satisfaz a Hipótese 3.1 en

$$W_n \xrightarrow{d} \chi_v^2$$
, quando $n \to \infty$, (3.71)

onde v = t·s é igual ao número de restrições impostas por Ho.

Demonstração: (ver Rodrigues 1975, pag 108)

A região crítica do teste é

$$W_n > \chi_{v,1-\alpha}^2$$
, (3.72)

onde a é o nível de significância.

C - Sequência de Pitman

Suponhamos que as alternativas satisfaçam a

$$H_1^{(n)}: (\underline{B}, \underline{\otimes} \underline{C}) = \frac{1}{\sqrt{n}} \underline{n} , \qquad (3.73)$$

onde n é um vetor de valores fixos.

O teorema que segue da a distribuição limite da estatística de Wald W_n , sob a condição alternativa (3.73)

 então

$$W_n \xrightarrow{d} \chi_v^2(\delta)$$
, quando $n \to \infty$, (3.74)

onde $v = t \cdot s$ é o número de graus de liberdade c δ é o par \hat{a} metro de não centralidade, com

$$\delta = \underline{\eta}' \left((\underline{B}' \otimes \underline{C}) \sum_{i=1}^{n} (\underline{\theta}) (\underline{B} \otimes \underline{C}') \right)^{-1} \underline{\eta}$$
 (3.75)

<u>Demonstração</u>: (ver Rodrigues, 1975)

Observação: A região crítica do teste é dada por (3.72), então o poder do teste quando a alternativa H_1 for H_1 : (B' \otimes C) $\xi = \eta$, é dado por

$$\underline{\beta}_{\mathbf{H}_{1}}(\xi) = P\left(\chi_{v}^{2}(\delta) > \chi_{v,1-\alpha}^{2}\right)$$
 (3.76)

onde o parâmetro de não centralidade é

$$\delta = \mathbf{n} \cdot \underline{\mathbf{n}} \left((\underline{\mathbf{B}}' \otimes \underline{\mathbf{C}}) \sum_{\underline{\mathbf{C}}} (\underline{\theta}) (\underline{\mathbf{B}} \otimes \underline{\mathbf{C}}) \right)^{-1} \underline{\mathbf{n}}$$
 (3.77)

CAPITULO - 4

O MODELO GERAL DE ESTRUTURA DE COVARIÂNCIAS

4.1 - INTRODUÇÃO

A análise de estrutura de covariâncias é um método geral para analisar testes de medidas educacionais, psicológicas, etc. Esta análise é amplamente aplicada nas ciências do comportamento. Tem por finalidade detectar e avaliar fon tes latentes de variação e covariação nas medidas observadas. A técnica da análise de estrutura de covariância é a de produzir variáveis latentes (não observadas), as quais têm a função de julgar o parentesco entre as variáveis observadas. As variáveis latentes possuem todas as informações necessárias referentes às interrrelações lineares dos indicadores observados (variáveis observadas). O modelo que abordaremos aqui, é aquêle apresentado por Yöreskog, (1970-a) e (1972).

4.2 - O MODELO GERAL

Seja X o vetor (px1) de variaveis observadas que tem a forma

$$X = \underline{u} + \underline{B} \wedge \underline{\xi} + \underline{B} \underline{\zeta} + \underline{\varepsilon} , \qquad (4.1)$$

onde $\xi(rx1)$, $\xi(qx1)$ e $\xi(px1)$ são vetores latentes não correlacionados de variâncias e covariâncias ϕ , ψ^2 e Θ^2 ,

respectivamente, e μ = E $\left(\begin{array}{c} X \end{array}\right)$. A matriz de variancias e covariancias do vetor X é dada por Σ e esta matriz tem a expansão

$$\sum = \underline{B} \left(\underline{\Lambda} \Phi \underline{\Lambda}' + \underline{\Psi}^2 \right) \underline{B}' + \underline{\bigoplus}^2 , \qquad (4.2)$$

onde as matrizes $\underline{B}(pxq) = (\beta_{ik}) e \underline{\Lambda}(qxr) = (\lambda_{km})$, são matrizes de parâmetros desconhecidos, como também, a matriz simétrica $\underline{\Phi}(rxr) = (\psi_{mm})$ e as matrizes diagonais $\underline{\Psi}(qxq) = (\delta_{ke} \underline{\Psi}_{k})$ e $\underline{\Theta}(pxp) = (\delta_{ij} \theta_{i})$, onde $\delta_{ke} = 1$ se $k = e \delta_{ke} = 0$ em caso contrário.

4.3 - IDENTIFICAÇÃO DO MODELO

No modelo geral (4.2) se fizermos a substituição de \underline{B} , $\underline{\Lambda}$, $\underline{\Phi}$ e $\underline{\Psi}^2$ por \underline{B} \underline{T}^{-1} , \underline{T}_1 $\underline{\Lambda}$ \underline{T}_2^{-1} , \underline{T}_2 $\underline{\Phi}$ \underline{T}_2' e \underline{T}_1 $\underline{\Psi}^2$ \underline{T}_1' , respectivamente, mantendo $\underline{\bigoplus}^2$, verifica-se que a matriz $\underline{\Sigma}$ permanece a mesma. As matrizes \underline{T}_1 (qxq) e \underline{T}_2 (pxp) são matrizes não singulares quaisquer, tais que \underline{T}_1 $\underline{\Psi}^2$ \underline{T}_1' seja uma matriz diagonal. Pelo que foi exposto, observamos que existe uma gran de quantidade de indeterminação no modelo geral. Por isso, para obter um conjunto único de parâmetros e seu conjunto de estimadores associados, algumas restrições são impostas. Os tipos de restrições que podem surgir sobre os parâmetros do modelo (4.2) são: 1°) parâmetros fixados (parâmetros com valores conhecidos), 2°) parâmetros restritos desconhecidos

(os parâmetros são desconhecidos, mas alguns dêles são i-guais) e 3º) parâmetros livres (parâmetros não conhecidos e diferentes)

4.4 - ESTIMAÇÃO

Seja X uma matriz (Nxp) de N observações sobre p variáveis. Supomos que os vetores linhas de X sejam independentes, e que tenham distribuição normal multivariada com matriz de variâncias e covariâncias \(\sum_{\text{.}} \)

A matriz X tem média dada por

$$E\left(\begin{array}{c} X \\ \end{array}\right) = A = P$$
 (4.3)

onde A (Nxg) = $(a_{\alpha s})$, P (hxp) = (p_{ti}) são matrizes conhecidas, tais que posto (A) = $g \le N$ e posto (P) = $h \le p$, enquanto E é uma matriz paramétrica de dinensão (gxh).

A matriz de variâncias e covariâncias \sum tem a forma (4.2), ou seja, $\sum = B \left(\bigwedge \Phi \Lambda' + \Psi^2 \right) + \bigoplus^2$.

Sejam U (gxg), V (gxp) e W (pxp), matrizes provenientes das matrizes A e X , que são representadas por

$$\underline{U} = \frac{1}{N} \underline{A}' \underline{A} \tag{4.4}$$

$$\underline{Y} = \frac{1}{N} \underline{A}' \underline{X}$$
 (4.5)

$$\underline{W} = \frac{1}{N} \underline{X} \cdot \underline{X} \tag{4.6}$$

Definamos a função T(E) por

O logaritimo da função de verossimilhança é dado por

log L =
$$-\frac{1}{2}$$
 N log (2π) $-\frac{1}{2}$ N log $(|\Sigma|)$ $-\frac{1}{2}$ $\sum_{\alpha=1}^{N}$ $\sum_{i=1}^{p}$ $\sum_{j=1}^{p} (x_{\alpha i} - \mu_{\alpha i})^{ii} x$ x $(x_{\alpha j} - \mu_{\alpha j})$, (4.8) onde $\mu_{\alpha i}$ e σ^{ii} são elementos da matriz média $E\left(|X|\right)$ e da inversa da matriz de variâncias e covariâncias Σ^{-1} , respectivamente.

Omitindo o têrmo constante, a expressão de log L pode também ser escrita como

$$\log L = -\frac{1}{2} N \left(\log(|\underline{\Sigma}|) + \operatorname{tr}(\underline{T} \underline{\Sigma}^{-1}) \right)$$
 (4.9)

Para determinar os estimadores de máxima verossimilhança de B, A, O, Y e H basta determinar os valores destas matrizes que maximizam (4.9) ou que minimizam

$$F = \frac{1}{2} \left\{ \log(|\underline{\Sigma}|) + \operatorname{tr}(\underline{T} \underline{\Sigma}^{-1}) \right\}$$
 (4.10)

onde $\underline{T} = \underline{T}(\underline{\Xi})$ dado por (4.7) e $\underline{\Sigma}$ tem a decomposição (4.2).

A função \underline{F} é encarada como uma função de $\underline{\Xi}$, \underline{B} , $\underline{\Lambda}$, $\underline{\phi}$, $\underline{\Psi}$ e $\underline{\Theta}$, uma vez que $\underline{T} = \underline{T}(\underline{\Xi})$ e $\underline{\Sigma}$ são funções de \underline{B} , $\underline{\Lambda}$, $\underline{\phi}$, $\underline{\Psi}$ e $\underline{\Theta}$.

As derivadas da função F são dadas por

$$\frac{\partial F}{\partial \Xi} = -\frac{2}{N} \underline{A}' (\underline{X} - \underline{A} \underline{\Xi} \underline{P}) \underline{\Sigma}^{-1} \underline{P}'$$
 (4.11)

$$\frac{\partial \mathbf{F}}{\partial \mathbf{B}} = 2 \Omega \mathbf{B} \left(\Lambda \Phi \Lambda' + \Psi^2 \right) \tag{4.12}$$

$$\frac{\partial F}{\partial \Lambda} = 2 B' \Omega B \Lambda \Phi \qquad (4.13)$$

$$\frac{\partial F}{\partial \phi} = 2 \Lambda' B' \Omega B \Lambda - \operatorname{diag} \left(\Lambda' B' \Omega B \Lambda \right)$$
 (4.14)

$$\frac{\partial F}{\partial \Psi} = 2 \operatorname{diag} (B' \Omega B) \Psi \tag{4.15}$$

$$\frac{\partial F}{\partial H} = 2 \Omega H$$
 (4.16)

onde
$$\Omega = \sum_{i=1}^{n-1} \left(\sum_{i=1}^{n} - \underline{T}(\underline{\Xi}) \right) \underline{\Sigma}^{-1}$$
 (4.17)

Vamos considerar três casos especiais para a determinação dos estimadores de Ξ , B , Λ , Φ , Ψ e Θ .

A - Caso Geral

Nosso primeiro objetivo é determinar o estimador de Ξ para Σ fixado. Fazendo $\frac{\partial F}{\partial \Xi} = 0$ em (4.11), temos que

$$\frac{1}{N} A' \left(X - A = P \right) \underline{\Sigma}^{-1} P' = \underline{O} , \qquad (4.18)$$

dai temos a solução

$$\tilde{\Xi} = \tilde{\mathbf{u}}^{-1} \ \tilde{\mathbf{v}} \ \tilde{\Sigma}^{-1} \ \tilde{\mathbf{P}} \cdot \left[\tilde{\mathbf{P}} \ \tilde{\Sigma}^{-1} \ \tilde{\mathbf{P}} \cdot \right]^{-1} \tag{4.19}$$

onde $\tilde{\Xi}$ recebe a notação $\hat{\Xi}_{\Sigma}$.

Como na análise fatorial, nosso segundo objetivo é determinar os valores de \underline{B} , $\underline{\Lambda}$, $\underline{\Phi}$ $\underline{\Psi}$ e $\underline{\Theta}$ que minimizam a função $\underline{f}(\underline{B}', \underline{\Lambda}, \underline{\Phi}, \underline{\Psi}, \underline{\Theta}) = \min_{\underline{\Xi}} F(\underline{\Xi}, \underline{B}, \underline{\Lambda}, \underline{\Phi}, \underline{\Psi}, \underline{\Theta}) = \underline{F}(\underline{\widehat{\Xi}}_{\Sigma}, \underline{B}, \underline{\Lambda}, \underline{\Phi}, \underline{\Psi}, \underline{\Theta}).$ (4.20)

Para minimizar f, colocamos os parâmetros livres de B, Λ , Φ , Ψ e Θ , respectivamente, em forma de um vetor $\Theta' = \begin{bmatrix} \theta_1, \theta_2, \dots, \theta_m \end{bmatrix}$; o que significa dizer, que f agora é vista como função de θ_1 , θ_2 , \dots , θ_m . Por outro lado, as expressões analiticas das derivadas primeiras de f em relação a B, Λ , Φ , Ψ e Θ , existem e são iguais às derivadas primeiras de F em relação a B, Λ , Φ , Ψ e Θ , respectivamente, no valor $\hat{\Xi}_{\Sigma}$, (ver Yöreskog, (1970-a) e (1972)). Assim, estamos mu nidos das condições para aplicar o método de Fletcher e Powell, da mesma maneira como foi usado em análise fatorial. Para o método de Fletcher e Powell um programa de computador é dado por Gruvaeus e Yöreskog (1970)

B - 1°) As matrizes envolvidas no modelo geral são dadas por, Λ(Nxg), P(hxp), B(pxq), Λ(qxr), Φ(rxr), Ψ(qxq) e
(H)(pxp).

Para este caso, as seguintes restrições são feitas junto ao modelo geral: r = q = p, $B = I_p$, $\Lambda = I_p$, $\Psi = 0$ e H = 0. Além destas condições, tomamos Ξ não restrita. Desta forma,

 $F = F(\Xi, \phi)$, significa dizer, que inicialmente determinaremos o estimador de Ξ condicionado a um dado $\phi(\Sigma = \phi)$.

Para este modelo, as derivadas primeiras de F são

$$\frac{\partial F}{\partial E} = -\frac{2}{N} \underline{A}' \left(\underline{X} - \underline{A} \underline{E} \underline{P} \right) \underline{\Sigma}^{-1} \underline{P}' \qquad (4.21)$$

$$\frac{\partial F}{\partial \Phi} = 2\Omega - \operatorname{diag}\left(\Omega\right) \tag{4.22}$$

onde $\Sigma = \Phi$ e $\Omega = \Sigma^{-1} \left(\Sigma - T(\Xi) \right) \Sigma^{-1}$.

Fazendo $\frac{\partial F}{\partial \Xi} = 0$, obtemos o E.M.V de Ξ condicionado a dada Σ , $\Xi_{\Sigma} = U^{-1} \ V \ \Sigma^{-1} \ P' (P \ \Sigma^{-1} - P')^{-1}$. Por outro lado, o valor de Σ ($\Sigma = \Phi$) que anula (4.22) E $\Sigma = T$ (Ξ). Significa dizer, que Σ fica em função de Ξ . Desta maneira, o processo de Fletcher e Powell pode ser utilizado.

Usando o processo de Fletcher Powell, encontramos o E.M.V de E que é dado por

$$\hat{\mathbf{E}} = \mathbf{U}^{-1} \mathbf{V} \mathbf{S}^{-1} \mathbf{P}' (\mathbf{P} \mathbf{S}^{-1} \mathbf{P}')^{-1}$$
 (4.23)

onde
$$S = W - V' U^{-1} V$$
, (4.24)

e o E.M.V de ∑ é dado por

$$\hat{\Sigma} = S + \hat{Q}' \cup \hat{Q} , \qquad (4.25)$$

onde
$$\hat{Q} = U^{-1} V - \hat{E} P$$
. (4.26)

O segundo membro de (4.25) é igual a $T(\frac{2}{5})$, pois, substituindo (4.24), (4.26) e sua transposta em (4.25), temos que

$$\hat{\underline{\Sigma}} = \underline{\underline{W}} - \underline{\underline{V}}, \ \underline{\underline{U}}^{-1} \ \underline{\underline{V}} + \left(\underline{\underline{V}}, \ \underline{\underline{U}}^{-1} - \underline{\underline{P}}, \ \hat{\underline{\Xi}}, \right) \underline{\underline{U}} \left(\underline{\underline{U}}^{-1} \ \underline{\underline{V}} - \hat{\underline{\underline{\Xi}}} \ \underline{\underline{P}}\right) =$$

$$= \underline{\underline{W}} - \underline{\underline{P}}, \ \hat{\underline{\Xi}}, \ \underline{\underline{V}} - \underline{\underline{V}}, \ \hat{\underline{\Xi}}, \ \underline{\underline{P}} + \underline{\underline{P}}, \ \hat{\underline{\Xi}}, \ \underline{\underline{U}}, \ \hat{\underline{\Xi}}, \ \underline{\underline{P}} = \underline{\underline{T}} \left(\hat{\underline{\Xi}}\right).$$

2°) Assumindo as hipóteses de (1°) acima, e fazendo P uma ma triz quadrada não singular, verificamos que (4.23) toma a forma $\hat{\Xi} = U^{-1} V P^{-1}$. (4.27)

$$\hat{Q} = 0 , \qquad (4.28)$$

 $\hat{\Sigma} = S \tag{4.29}$

4.5 - TESTES DE HIPOTESES

A - Teste de hipótese sobre E .

Um teste que pode aparecer para E é o seguinte

$$H_0: C = D = 0 \tag{4.30}$$

contra

$$H_1: C = D \neq 0$$

Através do processo dos mínimos quadrados, podemos determi - nar as matrizes

$$S_e = D' (P S^{-1} P') D (devida ao êrro)$$
 (4.31)

$$S_h = (C \hat{E} D)' (C \hat{R} C')^{-1} (C \hat{E} D) \text{ (sob a hipotese)},$$
onde
$$\hat{R} = U^{-1} + \hat{Q} S^{-1} \hat{Q}'.$$

Sejam $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_t$ os valores proprios da matriz $S_h S_e^{-1}$.

As estatísticas testes que podem ser usadas são: a da maior raíz, da soma das raízes e razão de verossimilhança. O teste da maior raíz é devido a Roy (1953); o da soma das raízes é devido a Lawley (1938); o teste da razão de verossimilhança é uma extensão do λ -teste de Wilks (1932) e pode ser usado com a tabela de correção dada por Schatzoff (1966). Quando N cresce, o valor - $\left(N-g-(p-h)-\frac{1}{2}(t-s+1)\right)$ vezes a razão de verossimilhança, tem aproximadamente a distribuição χ^2 com st graus de liberdade.

B - Em (4.9) apresentamos o logarítimo da função de verossimilhanaça pela formula $\log L = -\frac{1}{2} \, N \Big[\log(|\Sigma|) + \operatorname{tr}(T \, \Sigma^{-1}) \Big] \, e$ a função F por $F = \frac{1}{2} \Big[\log(|\Sigma|) + \operatorname{tr}(T \, \Sigma^{-1}) \Big] \, , \, \text{onde} \, T = T \, (\Xi)$ está definido em (4.7). Seja F_0 o mínimo da função F sob qualquer hipótese H_0 , referente à estrutura paramétrica do modelo geral, e F_1 o mínimo da função F sob a alternativa H_1 . A técnica da razão de máxima verossimilhança é empregada aqui para testar a hipótese H_0 contra a alternativa H_1 . A razão de verossimilhança é dada por

$$\lambda = \frac{L_1}{L_0} , \qquad (4.33)$$

onde L_o é a função de verossimilhança sob H_o , e L_1 sob H_1 .

Assim,
$$-2 \log x = N(F_0 - F_1)$$
,
onde, $F_1 \le F_0$. (4.34)

A estatística $u = -2 \log \lambda$, para N bastante grande tem distribuição χ^2 com o número de graus de liberdade igual a diferênça dos números de parâmetros estimados sobre H_1 e H_0 . Os cálculos das soluções sob H_0 e H_1 são geralmente requisitadas. Tomenos como exemplo o segundo exemplo de B de(4.4), supondo que sob H_1 , a matriz P é quadrada não singular. Então sob H_1 , os E.M.V's de E e Σ são dados por (4.27) e (4.29). A estatística teste é dada por

$$u = N[F_0 - \log(|\hat{\Sigma}_1|) - p]$$
 (4.35)

com $d = gp + \frac{1}{2}p(p+1) - n$ graus de liberdade, onde mé o número de parâmetros independentes estimados sob H_0 .

4.6 - CASOS PARTICULARES DO MODELO GERAL

1º Analise Fatorial

$$\sum = \Lambda \Phi \Lambda' + \Psi^2, \qquad (4.36)$$

onde $\underline{\Psi}^2 = \underline{\underline{\underline{\underline{I}}}} \left(\underline{\underline{\underline{Z}}} \, \underline{\underline{\underline{Z}}} \right)$. Assim, fazendo $\underline{\underline{\underline{B}}} = \underline{\underline{\underline{I}}}_q e \bigoplus \underline{\underline{\underline{H}}} = \underline{\underline{\underline{0}}}$ no mode

lo geral (4:2), obtemos a expressão (4.36).

Significa dizer que, o modelo fatorial (4.36) é um caso particular do modelo geral de estrutura de covariâncias.

2º - Análise de Componentes Principais.

Vimos no capítulo 2, que a análise de componentes principais é um caso particular da análise fatorial.

O modelo de análise de componentes principais, está representado em (2.39) pela formula $X = \Gamma' Y$, onde X(qx1) e Y(qx1), com E(Y) = 0. A matriz de variâncias e covariâncias de X é expressa por $\Sigma = \Gamma' \Phi \Gamma$, onde $\Gamma(qxq)$ e $Var(Y) = \Phi(qxq)$. Fazendo $B = I_q$, H = 0, $\Psi = 0$ e $\Lambda = \Gamma'$, verificamos que o modelo (2.39) é um caso particular de (4.2)

3°) - Analise de Trajetoria

a) Começaremos com o exemplo (2.1) da secção (2.4), que e representado pelas equações

$$\eta_{1} = \gamma_{11} \xi_{1} + \gamma_{12} \xi_{2} + 0 \xi_{3} , \qquad (4.37)$$

$$\eta_{2} = \alpha_{21} \eta_{1} + \gamma_{22} \xi_{2} + 0 \xi_{3} , \qquad (4.37)$$

$$\eta_{3} = \alpha_{32} \eta_{2} + \gamma_{33} \xi_{3} .$$

Estas expressões podem também ser escritas na forma

$$\eta_{1} = \gamma_{11} \xi_{1} + \gamma_{12} \xi_{2} + 0 \xi_{3}, \qquad (4.38)$$

$$\eta_{2} = \alpha_{21} \gamma_{11} \xi_{1} + (\alpha_{21} \gamma_{12} + \gamma_{22}) \xi_{2} + 0 \xi_{3}, \qquad (4.38)$$

$$\eta_{3} = \alpha_{32} \alpha_{21} \gamma_{11} \xi_{1} + \alpha_{32} (\alpha_{21} \gamma_{12} + \gamma_{22}) \xi_{2} + \gamma_{33} \xi_{3}.$$

As formulas em (4.37) formam as equações estruturais, enquanto as formulas de (4.38) são equações na forma reduzida.

Temos seis parâmetros nas equações estruturais, Nas equações de forma reduzida, temos nove parâmetros, sendo dois nu los e os restantes em função dos seis parâmetros estruturais.

Definamos agora uma transformação biunívoca que leva $(\alpha_{21}^2, \alpha_{32}^2, \gamma_{11}^2, \gamma_{12}^2, \gamma_{22}^2, \gamma_{33}^2)$ em $(\mu, \nu, \delta, \epsilon, \Phi, \kappa)$ que é dada por

$$\mu = \alpha_{21},
v = \alpha_{32} \cdot \alpha_{21},
\delta = \gamma_{11},
\epsilon = \gamma_{12},
\phi = \gamma_{12} + (\gamma_{22}/\alpha_{21}),
\kappa = \gamma_{33}/\alpha_{32} \cdot \alpha_{21}.$$
(4.39)

Com a transformação paramétrica (4.39), as expressões de (4.38) tomam agora, o seguinte aspecto:

$$\eta_{1} = \delta \xi_{1} + \varepsilon \xi_{2} + 0 \xi_{3}
\eta_{2} = \mu \delta \xi_{1} + \mu \Phi \xi_{2} + 0 \xi_{3}
\eta_{3} = \nu \delta \xi_{1} + \nu \Phi \xi_{2} + \nu \kappa \xi_{3}$$
(4.40)

Supomos que η_1 , η_2 , η_3 , ξ_1 , ξ_2 e ξ_3 não sejam diretamente observadas; mas que $y_i = v_i + \eta_i + \varepsilon_i$ e $x_i = \mu_i + \xi_i + \delta_i$, i = 1, 2, 3, sejam observadas. Os parâmetros v_i e μ_i são

médias de y_i e x_i , respectivamente, enquanto ε_i e δ_i são os fatores erros de medidas, i = 1, 2, 3.

Definamos as matrizes B e Λ por

$$\Lambda =
\begin{bmatrix}
\delta & \epsilon & 0 \\
\delta & \Phi & 0 \\
\delta & \Phi & \kappa \\
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}$$

como consequência das definições de B e Λ temos

$$\underline{B} \cdot \underline{\Lambda} =
\begin{pmatrix}
\delta & \epsilon & 0 \\
\mu \delta & \mu \Phi & 0 \\
v \delta & v \Phi & v \kappa \\
1 & 0 & 0 \\
-0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}$$
(4.42)

Colocando os yi's com os xi's na forma vetorial, temos

$$\underline{\underline{Y}} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \underline{\underline{X}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \underline{\underline{v}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad \underline{\underline{\mu}} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

$$\underline{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}, \quad \underline{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}, \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix} e \underline{\delta} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix}$$

É fácil verificar que

Substituindo (4.44) em (4.43) temos

A matriz de variâncias e covariâncias de $\begin{pmatrix} Y \\ \overline{X} \end{pmatrix}$ tem a expansão

$$\Sigma = B \wedge \Phi \wedge B' + \bigoplus^{2}, \qquad (4.46)$$

onde o (3x3) é a matriz de variâncias e covariâncias de ¿ e

 $\overset{\bullet}{\mathbb{H}}^2$ (6x6) é a matriz de variâncias e covariâncias de $\begin{pmatrix} \varepsilon \\ \gamma \end{pmatrix}$. A matriz $\overset{\bullet}{\mathbb{H}}^2$ é diagonal.

Concluimos que a expressão (4.46) é o caso particular de (4.2), quando se faz $\Psi = 0$.

b) Agora, trataremos do exemplo (2.2) apresentado no segundo capítulo.

Denotaremos A por

$$\Lambda =
\begin{bmatrix}
\alpha_1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\alpha_2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
\alpha_3 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
y & \beta_1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & \beta_2 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & \beta_3 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

Logo, a expressão (2.56) também pode ser escrita na forma

$$X = \Lambda \xi$$

A matriz de variâncias e covariâncias de X tem a expressão

$$\sum = \Lambda \Phi \Lambda' \tag{4.48}$$

Este modelo foi discutido por Costner, 1969, onde a matriz de variâncias e covariâncias o possui a forma

Com os dados anteriores, verificamos que (4.48) é um caso particular de (4.2), quando se faz $B = I_6$, $\psi^2 = 0(6 \times 6)$ e Θ

c) 0 exemplo (2.3) $\tilde{\epsilon}$ discutido por Hauser ϵ Goldberg 1971. As expressões (2.57) apresentam a forma $\xi = \alpha' X + \nu$ $\epsilon Y = \beta \xi + \mu$.

Hauser e Goldberg adimitem que μ_1 , μ_2 e μ_3 sejam mutuamente não correlacionados e v=0. Denotaremos por \sum_{yy} e \sum_{xx} as matrizes de covariâncias de Y e X, respectivamente, enquanto \sum_{xy} é a matriz de covariâncias entre X e Y.

Seja o vetor $\begin{pmatrix} \frac{Y}{Z} \\ \frac{X}{Z} \end{pmatrix}$ com matriz de variâncias dada por

$$\underline{\underline{\Gamma}} = \begin{bmatrix} \underline{\underline{\Gamma}}_{xy} & \underline{\underline{\Gamma}}_{xx} \\ \underline{\underline{\Gamma}}_{xy} & \underline{\underline{\Gamma}}_{xx} \end{bmatrix}$$

Das expressões (2.57) tiramos que

$$\underline{\sum}_{yy} = \underline{\beta} \underline{\alpha}' \, \underline{\sum}_{xx} \underline{\alpha} \underline{\beta}' + \underline{\beta}\underline{\beta}' + \underline{\theta}^2 . \qquad (4.50)$$

$$\underline{\sum}_{xy} = \underline{\beta} \underline{\alpha}' \, \underline{\sum}_{xx}. \quad \text{tal que } \underline{\sum}_{xx} \text{ seja não restrita.}$$

Daí temos que

$$\Sigma = \begin{bmatrix}
\underline{\beta} \underline{\alpha}' & \underline{\Sigma}_{xx} & \underline{\alpha} \underline{\beta}' + \underline{\beta} \underline{\beta}' & \underline{0}^2 & \underline{\beta} \underline{\alpha}' \underline{\Sigma}_{xx} \\
(\underline{\beta} \underline{\alpha}' & \underline{\Sigma}_{xx})' & \underline{\Sigma}_{xx}
\end{bmatrix} .$$
(4.51)

Fazendo

$$\underline{B} = \begin{bmatrix}
3x1 & 3x3 \\
\underline{\beta} & \underline{0} \\
\underline{0} & \underline{I}_{3}
\end{bmatrix}, \quad \underline{\Lambda} = \begin{bmatrix}
1x3 \\
\underline{\alpha}_{3} \\
\underline{I}_{3}
\end{bmatrix}, \quad \underline{\Phi} = \underline{\Sigma}_{XX}, \quad (4.52)$$

$$\Psi = \text{diag}(1,0,0,0) \in \Theta = \begin{bmatrix} 3x3 & 3x3 \\ \frac{\sigma}{2} & 0 \\ 0 & 0 \\ 3x3 & 3x3 \end{bmatrix}, \text{ onde}$$

$$\underline{\sigma} = \begin{bmatrix} \sigma_{\mu_1} & 0 & 0 \\ 0 & \sigma_{\mu_2} & 0 \\ 0 & 0 & \sigma_{\mu_3} \end{bmatrix}, \text{ verificamos que } \underline{\Sigma} \text{ satisfaz}$$

a condição (4.2).

Pois,

$$\underline{\underline{\Sigma}} = \begin{bmatrix} \underline{\Sigma}_{xx} & \underline{\Sigma}_{xy} \\ \underline{\Sigma}'_{xy} & \underline{\Sigma}_{yy} \end{bmatrix} = \begin{bmatrix} \underline{\beta} & \underline{\alpha}' \underline{\Sigma}_{xx} & \underline{\alpha} & \underline{\beta}' + \underline{\beta} & \underline{\beta}' + \underline{\sigma}^2 & \underline{\beta} & \underline{\alpha}' & \underline{\Sigma}_{xx} \\ (\underline{\beta} & \underline{\alpha}' \underline{\Sigma}_{xx})' & \underline{\Sigma}_{xx} \end{bmatrix} = \underline{\Sigma}_{xx}$$

$$= \begin{pmatrix} \underline{\beta} & \underline{\alpha}^{\dagger} & \underline{\Sigma}_{XX} & \underline{\alpha} & \underline{\beta}^{\dagger} & \underline{\beta} & \underline{\alpha}^{\dagger} & \underline{\Sigma}_{XX} \\ \underline{\Sigma}_{XX} & \underline{\alpha} & \underline{\beta}^{\dagger} & \underline{\Sigma}_{XX} \end{pmatrix} + \begin{pmatrix} \underline{\beta} & \underline{\beta}^{\dagger} & \underline{0} \\ \underline{0} & \underline{0} \end{pmatrix} +$$

$$\begin{bmatrix} \dot{\underline{\sigma}}^2 & & \underline{O} \\ \underline{O} & & \underline{O} \end{bmatrix} = \begin{bmatrix} \underline{\beta} & \underline{\alpha}' & \underline{\Sigma}_{xx} \\ & & \underline{\Sigma}_{xx} \end{bmatrix} \begin{bmatrix} \underline{\alpha} & \underline{\beta}' & \underline{I}_3 \end{bmatrix} + \begin{bmatrix} \underline{\beta} & \underline{\beta}' & \underline{O} \\ 0 & 0 & 0 \\ 3\bar{x}3 & 3\bar{x}3 \end{bmatrix} + \begin{bmatrix} \underline{C} & \underline{C} & \underline{C} & \underline{C} & \underline{C} & \underline{C} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 3x3 & 3x3 \\ \underline{\sigma}^2 & \underline{O} \\ \underline{O} & \underline{O} \\ 3x3 & 3x3 \end{bmatrix} = \begin{bmatrix} \underline{\beta} & \underline{\alpha} \\ \underline{I}_3 \end{bmatrix} \underline{\Sigma}_{xx} \begin{bmatrix} \underline{\alpha} & \underline{\beta} \\ \underline{I}_3 \end{bmatrix} +$$

$$\begin{bmatrix}
3x1 & 3x3 \\
\underline{\beta} & \underline{0} \\
\underline{0} & \underline{0} \\
3x1 & 3x3
\end{bmatrix}
\begin{bmatrix}
\underline{\beta}' & \underline{0} \\
\underline{0} & \underline{0}
\end{bmatrix}
+
\begin{bmatrix}
\underline{\sigma}^2 & \underline{0} \\
\underline{0} & \underline{0}
\end{bmatrix}
=
\begin{bmatrix}
\underline{1} & \underline{0} & \underline{0} & \underline{0}
\end{bmatrix}$$

$$\left\{ \begin{array}{ccc} \underline{\beta}' & \underline{O} \\ \underline{O} & \underline{O} \end{array} \right\} + \left\{ \begin{array}{ccc} \sigma^2 & \underline{O} \\ \underline{O} & \underline{O} \end{array} \right\} = \underline{B} \left[\underline{\Lambda} \stackrel{\Phi}{-} \underline{\Lambda}' + \underline{\Psi}^2 \right] \stackrel{B'}{-} + \stackrel{\Theta}{-} \frac{2}{2}$$

Significa dizer, que o modelo do exemplo (2.3) é um caso par ticular do modelo (4.2).

49) Analise de Curvas de Crescimento

Seja N o número de observações de uma variável aleatória e m o número de grupos independentes, onde cada grupo tem n_{α} observações ($\sum_{\alpha=1}^{m} n_{\alpha} = N$). Cada observação é medida em posições diferentes $t_1, t_2, \dots t_p$, num certo intervalo de tempo. Assim podemos considerar a curva de crescimento

$$E[x_t] = \xi_{so} + \xi_{sl}t + \cdots + \xi_{sh}t^h$$
,

s = 1, 2...m, isto \tilde{e} , estamos de posse de um modelo do tipo (2.58) onde A (Nxm) \tilde{e} a matriz com as n_1 linhas $(1, 0, ... 0), n_2$ linhas (0, 1, ... 0) ... e n_m linhas (0, 0, ... 1), a matriz ξ (mxh+1) = (ξ_{sk}) e P(h+1xp) = (t_1^j) .

O valor esperado da matriz das observações X é dado por

$$E\left(\begin{array}{c} X \end{array}\right) = \underline{A} \ \underline{\xi} \ \underline{P} \tag{4.53}$$

Fazendo $\xi = \Xi$, observamos que (4.53) tem a mesma forma de (4.3).

Suponhamos, agora, que as variáveis X_t também satisfaçam ao Processo de Wiener

$$X_t = \mu_t + Z_1 + Z_2 + \cdots + Z_t + \varepsilon_t$$
 (4.54)

onde $\mu_t = E(X_t)$, os Z_t (t=1,2,...p) são variáveis com incrementos independentes e ε_t é a variável êrro de medida. Este modelo também pode ser apresentado na forma matricial

$$X = \mu + L Z + \varepsilon , \qquad (4.55)$$

onde
$$X' = \left(X_1, X_2, \dots, X_p\right), Z' = \left(Z_1, Z_2, \dots, Z_p\right), \mu' = \left(\mu_1, \mu_2, \dots, \mu_p\right),$$

$$\underline{\varepsilon}' = \left(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p\right) \in \underline{L} \quad \text{uma matriz triangular inferior.}$$

A matriz de variâncias e covariâncias de X tem a forma

$$\sum = L \phi L' + \psi^2 \tag{4.56}$$

onde Φ e a matriz de variâncias e covariâncias de Ξ , Ψ^2 e a matriz de variância e covariância de E. Λ expressão (4.56) tem a mesma forma (4.2) fazendo, B = I, $\Theta = 0$ e $\Lambda = L$.

CAPÍTULO 5 ÁLGUNS EXEMPLOS

5.1 - EXEMPLO NUMERICO

5.1.1 - Introdução

dos simulados são usados para a obtenção dos estimadores e para elaboração de testes.

5.1.2 - Dados Iniciais

Consideremos as seguintes matrizes:

$$\mathbf{A} = \begin{bmatrix}
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1
\end{bmatrix}, \quad \Phi = \mathbf{I}_{2}, \quad \Xi = \begin{bmatrix}
0, 2 & 1, 2 \\
0, 8 & 2, 8
\end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix}
0, 9 & 0 \\
0, 7 & 0, 2 \\
0, 5 & 0, 4 \\
0, 3 & 0, 6 \\
0, 1 & 0, 8
\end{bmatrix}, \quad \Psi = \operatorname{diag} \begin{bmatrix}
\Psi_{1} \\
\Psi_{2} \\
\Psi_{3} \\
\Psi_{4} \\
\Psi_{5}
\end{bmatrix} = \begin{bmatrix}
0, 3 \\
0, 4 \\
0, 5 \\
0, 6 \\
0, 7
\end{bmatrix}$$

$$\underline{A} \ \underline{A}' = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix} \quad \underline{B}' = \begin{bmatrix} 0,18 & 0,30 & 0,42 & 0,55 & 0,66 \\ & & & & \\ 1,08 & 1,40 & 1,72 & 2,04 & 2,36 \end{bmatrix} \quad e$$

$$\underline{\mathbf{V}} = \begin{bmatrix}
1,11 \\
0,63 & 0,93 \\
0,45 & 0,43 & 0,91 \\
0,27 & 0,33 & 0,39 & 1,05 \\
0,09 & 0,23 & 0,37 & 0,51 & 1,35
\end{bmatrix}$$

onde r = q = 2, p = 5 e n = 200.

As restrições são apresentadas somente nos parâmetros estrurais λ_{12} de Λ e Φ , onde λ_{12} = 0 e Φ = \underline{I}_2 .

5.1.3 - Simulação

A - Os modelos utilizados aqui, são os modelos (3.1) e (3.2), ou seja,

As expressões (5.1) e (5.2) podem também ser apresentadas nas formas

$$Y = \begin{pmatrix} Y_{1} & Y_{2} & \cdots & Y_{100} & Y_{101} & Y_{102} & \cdots & Y_{200} \end{pmatrix} = (5...)$$

$$\begin{cases} \Xi_{1} & \Xi_{1} & \cdots & \Xi_{1} & \Xi_{2} & \Xi_{2} & \cdots & \Xi_{2} \end{pmatrix} + (5...)$$

$$\begin{cases} \Xi_{1} & \Xi_{1} & \cdots & \Xi_{10} & \Xi_{101} & \Xi_{102} & \cdots & \Xi_{200} \end{pmatrix} = (5...)$$

$$X = \begin{pmatrix} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} = (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} = (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + (5...)$$

$$\begin{cases} X_{1} & X_{2} & \cdots & X_{100} & X_{101} & X_{102} & \cdots & X_{200} \end{pmatrix} + ($$

$$Z_{\mathbf{i}}^{\bullet} = \left(Z_{1\mathbf{i}} Z_{2\mathbf{i}} \cdots Z_{5\mathbf{i}}\right), \ \mathbf{i} = 1, 2, \dots, 200, \quad \text{com } \underline{\varepsilon}_{\mathbf{i}} \sim N_{2}\left[\underline{0}, \underline{\phi}\right]$$

$$e \quad Z_{\mathbf{i}} \sim N_{5}\left(\underline{0}, \underline{\psi}\right).$$

B - Para gerar dados da primeira população, faremos o seguin

- 1°) geramos duas variáveis aleatórias N(0,1) (normal padrão) para formar o vetor ε_1 ,
- 2°) em seguida, somamos $\underline{\epsilon}_{i}^{i} + \underline{\epsilon}_{1}^{i} = \underline{Y}_{i}^{i}$, $i=1,2,\ldots,100$, onde $\underline{\epsilon}_{1}$ $\underline{\epsilon}$ o primeiro vetor coluna de $\underline{\epsilon}$ dado em 5.1.2,
- 3°) e para obtermos o vetor Z_1 , geramos cinco variáveis alea tórias N(0,1) (normal padrão), em seguida, multiplica-se cada variável gerada por $\sqrt{\Psi_k}$, $k=1,2,\ldots,5$,
- 4°) e efetuando o produto $\underline{\Lambda}\underline{Y}_{i}$, obtemos $\underline{X}_{i} = \underline{\Lambda} \underline{Y}_{i} + \underline{Z}_{i}$, i = 1, 2, ..., 100.

As variaveis aleatórias normais são geradas em computador, através de subrotinas amplamente conhecidas, em particular, para este problema foi utilizado a subrotina VARGEN, disponível no C.C da Universidade da Carolina do Norte.

C - Da Matriz X

$$\mathbf{A} \ \mathbf{X'} = \begin{bmatrix} 31,2911 & 58,5773 & 63,6883 & 59,1566 & 76,1187 \\ & & & & \\ 110,0359 & 137,2624 & 174,0489 & 213,5507 & 233,8271 \end{bmatrix} (5.4)$$

$$\mathbf{X} \ \mathbf{X'} = \begin{bmatrix} 335,0615 \\ 280,9319 & 390,6948 \\ 295,2279 & 362,0408 & 525,2303 \\ 304,7186 & 410,8762 & 506,5956 & 762,9457 \\ 318,3725 & 430,5187 & 540,4315 & 683,0306 & 881,1164 \end{bmatrix}$$
 (5.5)

D. - Os estimadores de mínimos quadrados, são dados por

$$\frac{\mathbf{B}}{\mathbf{E}} = \mathbf{X} \mathbf{A}' \quad (\mathbf{A} \mathbf{A}')^{-1} \quad \mathbf{e}$$

$$\frac{\mathbf{V}}{\mathbf{V}} = \frac{1}{200} \mathbf{X} \left(\mathbf{I}_{200} - \mathbf{A}' (\mathbf{A} \mathbf{A}')^{-1} \mathbf{A} \right) \mathbf{X}', \quad \text{então}$$

por (5.4) e (5.5) obtemos

$$\overline{\underline{\beta}}^{\bullet} = \begin{bmatrix} 0,3129 & 0,5858 & 0,6369 & 0,5916 & 0,7612 \\ & & & & \\ 1,1004 & 1,3726 & 1,7405 & 2,7405 & 2,3383 \end{bmatrix}$$
 (5.6)

$$\overline{\underline{V}} = \begin{cases} 1,0210 \\ 0,5578 & 0,8399 \\ 0,4198 & 0,4292 & 0,9087 \\ 0,2561 & 0,4155 & 0,4862 & 1,3596 \\ 0,1863 & 0,4349 & 0,4249 & 0,6933 & 1,3821 \end{cases}$$
 (5.7)

5.1.4 - Estimação de Máxima Verossimilhança

A - Estimação dos Parâmetros Estruturais Λ, φ e Ψ.

Na parte de estimação, vimos que Ξ e o unico estimador que possui expressão analítica e que esta e função de Λ, φ e Ψ, ou seja,

$$\tilde{\Xi} = (\Lambda' V^{-1} \Lambda)^{-1} \Lambda V^{-1} X \Lambda' (\Lambda \Lambda')^{-1}, \qquad (5.8)$$

onde V = A o A' + Y . Em seguida, minimizamos a função

$$\tilde{F} = \log(|V|) + tr(V^{-1}T),$$
 (5.9)

onde $\tilde{T} = \frac{1}{200} (X - \tilde{\beta} \Lambda)(X - \tilde{\beta} \Lambda)$ com $\tilde{\beta} = \tilde{\Xi} \tilde{\Lambda}$, que $\tilde{\epsilon}$ função somente de Λ , $\Phi \in \Psi$.

Para a minimização de \tilde{F} , colocamos os parâmetros livres das matrizes Λ , Φ e Ψ em forma de um vetor $\Theta' = (\theta_1, \theta_2, \dots, \theta_m)$, passando a função \tilde{F} a ser encarada como uma função de Θ , Φ

Foram apresentados dois processos de minimização de f(0) o processo de Davidon, W.C, 1959 e o de Fletcher e Powell, 1963.

Para este caso, $\Phi = I_2$ e $\lambda_{12} = 0$, significa dizer que $m = N(\Lambda) + N(\Psi) = 9+5 = 14$ (onde $N(\cdot)$ é o número de parâmetros livres).

O cálculo destes estimadores é bastante trabalhoso e um programa, escrito em Fortran IV-G, é usado para o ajuste do modelo. Este programa pode ser obtido no Centro de Computação da Carolina do Norte. Para minimizar F podemos usar os métodos.

- 1°) "Steepest descent",
- 2°) Fletcher-Powell com matriz inicial E = I
- 3°) Fletcher-Powell com matriz inicial

- $\underline{E} = \frac{1}{2} \left[\underline{I}_{11}(\underline{0}) \underline{I}_{12}(\underline{0}) \ \underline{I}_{22}^{-1} \ \underline{I}_{21}(\underline{0}) \right]^{-1}, \text{ onde } \underline{0} \text{ contem os}$ $\text{parametros livres de } \underline{\Xi}, \quad \underline{\Lambda} \in \underline{\Psi}, \text{ onde } \underline{E} = \frac{1}{2} \ \underline{J}(\underline{0}) \text{ que } \underline{e}$ $\text{a primeira matriz na diagonal da matriz bloco} \quad \underline{I}^{-1}(\underline{0})$ em (3.68).
- 4°) uma combinação do (1°) com o (3°), sendo que o 1° é usado nas primeiras iterações e o 3° para o resto da operação.

Estes processos são representados simbolicamente por letras; SD, FP-I, FP-II e SD/FP-II, respectivamente.

Estes métodos dependem de una condição inicial, $\psi_i > \epsilon$, i = 1,2,...,p, e que o processo iterativo para imediatamente, quando se verifica que todas as derivadas de \tilde{F} são inferiores ao ϵ dado (ϵ pequeno).

A tabela (5.1) mostra o desenvolvimento dos processos S D, F P-I e S D/F P-II com valor inicial ϵ = 0,001. O simbolo G_{max} indica o valor da maior derivada de \tilde{F} , correspondente ao número de iterações e ao valor de \tilde{F} . No final da tabela, \tilde{e} apresentado o tempo de C.P.U. em segundos, gasto durante as iterações pelos métodos, S D, F P-I e S D/F P-II, respectivamente, num sistema IBM/360-75.

stice and a mental energy and a

Tabela 5.1

| Itera | S D | | F P- | -1 | SD/FP-II | | |
|-------|----------------------|------------------|----------|------------------|----------|------------------|--|
| ções | $\tilde{\mathbf{F}}$ | G _{max} | - F | G _{max} | Ĩ | G _{max} | |
| 0 | 5,717378 | 0,895047 | 5,713378 | 0,895047 | 5,717378 | 0,895047 | |
| . 1 | 5,089637 | 0,446847 | 5,089637 | 0,446847 | 5,089637 | 0,446847 | |
| 2 | 4,691579 | 0,681971 | 4,486940 | 0,628014 | 4,691579 | 0,681971 | |
| 3 | 4,440242 | 0,352651 | 4,390288 | 0,548307 | 4,440242 | 0,352651 | |
| 4 | 4,318538 | 0,590315 | 4,258196 | 0,481306 | 4,318538 | 0,590315 | |
| 5 | 4,246603 | 0,298247 | 4,211948 | 0,646492 | 4,068019 | 0,212787 | |
| 6 | 4,210801 | 0,325221 | 4,131595 | 0,400649 | 4,055455 | 0,411462 | |
| 8 | 4,164438 | 0,206044 | 4,092626 | 0,272744 | 4,044735 | 0,027394 | |
| 10 | 4,128719 | 0,182090 | 4,068557 | 0,084633 | 4,044586 | 0,003847 | |
| 12 | 4,099288 | 0,154941 | 4,057798 | 0,167238 | 4,044574 | 0,000889 | |
| 15 | 4,066047 | 0,291028 | 4,045418 | 0,052400 | | 727 (d) | |
| 18. | 4,049712 | 0,037323 | 4,044612 | 0,007396 | | | |
| 21 | 4,046359 | 0,033712 | 4,044574 | 0,000458 | Δ. | ů. | |
| 182 | 4,044577 | 0,000985 | | | | | |
| Tempo | 32,44 | | 4 | ,89 | 1,63 | | |

Observe que neste caso, o SD / FP-II gastou menos iterações e pouco tempo, segnificando dizer, que é o método que apresenta mais vantagens. Logo, deve ser o preferido, embora possa existir outro melhor.

Os estimadores de máxima verossimilhança de Ξ , Λ e Ψ obtidos são

$$\hat{\Xi} = \begin{cases} 0,5078 & 1,3459 \\ 0,7566 & 2,4067 \end{cases} \qquad \hat{\Lambda} = \begin{cases} 0,7956 & 0 \\ 0,6900 & 0,1986 \\ 0,5325 & 0,4233 \\ 0,3265 & 0,7017 \\ 0,2315 & 0,8392 \end{cases}$$

$$\hat{\Psi} = \text{diag} \begin{cases} 0,3875 \\ 0,3237 \\ 0.4831 \\ 0,7088 \\ 0,6469 \end{cases}$$

B - Estimação dos Parâmetros da forma reduzida β e V.

O modelo empregado aquí é o modelo linear latente. Os estimadores de máxima verossimilhança dos parâmetros na forma re
duzida β e V são

$$\hat{\beta} = \begin{cases} 0,4040 & 0,5006 & 0,5907 & 0,6967 & 0,7525 \\ & & & & & \\ 1,0708 & 1,4006 & 1,7354 & 2,1282 & 2,3313 \end{cases}$$

$$\tilde{\mathbf{V}} =
\begin{cases}
1,0205 \\
0,5490 & 0,8392 \\
0,4237 & 0,4515 & 0,9458 \\
0,2598 & 0,3646 & 0,4709 & 1,3078 \\
1,1842 & 0,3664 & 0,4785 & 0,6645 & 1,4047
\end{cases}$$

Os estimadores de mírimos quadrados β e \overline{V} e os de máxima veresimilança β e \overline{V} de β e V, respectivamente, estão bem próximos um do outro, pois, a maior e média diferença absolu

ta entre $\vec{\beta}$ c $\hat{\vec{\beta}}$ são 0,1051 e 0,0419, enquanto que, entre \vec{V} e $\hat{\vec{V}}$ são 0,0518 e 0,0205.

Os erros padrões de ê , î e v são dados por

$$\mathbf{s.e} \left(\hat{\Xi} \right) = \begin{bmatrix} 0,1155 & 0,1700 \\ 0,1440 & 0,2849 \end{bmatrix}, \quad \mathbf{s.e} \left(\hat{\Lambda} \right) = \begin{bmatrix} 0,9370 & 0 \\ 0,0801 & 0,0539 \\ 0,0742 & 0,0495 \\ 0,0825 & 0,0583 \\ 0,0938 & 0,0759 \end{bmatrix} e$$

$$\mathbf{s.e} \left(\begin{array}{c} \widehat{\Psi} \\ \end{array} \right) = \begin{bmatrix} \mathbf{0,1332} \\ \mathbf{0,0733} \\ 0,0601 \\ \mathbf{0,0791} \\ \mathbf{0,1216} \end{bmatrix}$$
, onde os valores destas matrizes são erros padrões das matrizes correspondentes, por exemplo:

$$0.1155 = s.e(\hat{\Xi}_{11}), 0.0495 = s.e(\lambda_{32}) = 0.0791 = s.e(\Psi_4)$$

Estes valores são obtidos através da matriz informação de Fisher.

Os estimadores dos desvios padrões são dados por

$$\widehat{s.e} \left(\widehat{\Xi} \right) = \begin{bmatrix} 0,1285 & 0,1992 \\ 0,1493 & 0,2657 \end{bmatrix}, \quad \widehat{s.e} \left(\widehat{\Psi} \right) = \operatorname{diag} \begin{bmatrix} 0,1205 \\ 0,0708 \\ 0,0590 \\ 0,0973 \\ 0,1198 \end{bmatrix}$$

C - Estimação de E no Modelo Linear (5.1)

A matriz paramétrica E é a matriz de parâmetros desconheci - dos do modelo linear geral dado em (5.1). O processo de si-

mulação tornou a variável latente Y em variável conhecida .

Logo o estimador de mínimos quadrados E de E e o erro padrão de E são determinados, e dados por

$$\overline{\underline{z}} = \begin{bmatrix} 0,3344 & 1,1531 \\ 0,9395 & 2,9170 \end{bmatrix} e s.e(\overline{\underline{z}}) = \begin{bmatrix} 0,0990 & 0,0990 \\ 0,1002 & 0,1002 \end{bmatrix}.$$

Observe que o estimador Ξ está mais próximo da matriz Ξ dada em 5.1.2 do que o estimador $\widehat{\Xi}$, e que os elementos da matriz s.e(Ξ) são inferiores aos da matriz s.e($\widehat{\Xi}$).

Este resultado mostra que durante o processo de simulação, algumas informações são perdidas quando as variáveis são latentes.

5.1.5 - TESTES DE HIPOTESES

A - Adaptação do Modelo

Indicamos por N(•) = o número de parâmetros livres de uma ma triz. Em seguida, apresentamos a tabela 5.2 que da o número de parâmetros livres dos parâmetros estruturais e da forma reduzida.

Tabela 5.2

| | PARAMETROS | | | | | | | |
|--------|-------------|----|-------|----------|-------------------|----|-----|---------|
| N(•) | ESTRUTURAIS | | TOTAL | FO RE | FORMA REDUZIDA | | 4 | |
| | ES) | Ž. | Ψ | m | β | V | .e | g.l=L-m |
| Livres | - 4 | 9 | 5 | 18 | `10 | 15 | 2.5 | 7 |

A estatística $G_n = -2 \log \lambda_n = 9,51$ que da aproximadamente um p-valor = 0,2181. Logo, para os níveis $\alpha < 0,2181$ aceita mos a hipôtese de adaptação do modelo.

,B - Teste de Hipóteses sobre os Parâmetros Estruturais A e Y
Consideremos o seguinte teste

$$H_0$$
: $\Psi = \psi I_5$, contra (5.10)
 H_1 : $\Psi \neq \psi I_5$

A estatística G_n é dada por

$$G_n = -2 \log \lambda_n = 200 \left[F\left(\hat{\hat{\varrho}}_n\right) - F\left(\hat{\varrho}_n\right) \right]$$
 (5.11)

O objetivo é determinar o estimador de máxima verossimilhança sob a restrição de H_0 , $\hat{\underline{\theta}}_n$.

Para calcular $\hat{0}$ foi usado o programa F P-II, usando-se os verdadeiros valores de Λ dados em 5.1.2 e o valor verdadeiro $\Psi = 0.5$. Os valores de $\hat{\Lambda}$ e $\hat{\Psi}$ determinados são

$$\hat{\bar{\Lambda}} =
\begin{cases}
0,7384 & 0 \\
0,6395 & 0,2070 \\
0,5742 & 0,3967 \\
0,3305 & 0,7455 \\
0,2280 & 0,8981
\end{cases}
e \hat{\bar{\Psi}} = 0,50381 I_5.$$

Como
$$\hat{\vec{\xi}}$$
 e função de $\hat{\vec{\lambda}}$ e $\hat{\vec{\psi}}$, então $\hat{\vec{\xi}} = \begin{pmatrix} 0.5489 & 1.4695 \\ 0.6740 & 2.2232 \end{pmatrix}$.

O processo convergiu em 5 iterações determinando o valor de \vec{F} = 4,157385 e o G_{max} = 0,000506 para ε = 0,001 .

O teste de adaptação do modelo linear latente para o caso em que as matrizes $\underline{\Lambda}$ e $\underline{\Xi}$ são dadas em 5.1.2 e $\underline{\Psi}$ = 0,5 \underline{I}_5 , dã uma estatística G_{200} = 32,07, onde G_{200} tem distribuição x_{11}^2 , com ℓ = 25 e valor p = 0,0007, mostrando assim, que o modelo não se adapta aos dados.

Agora vamos testar a Hipótese (5.10) para o modelo linear estrutural, usando a teoria indicada em (3.8.2). Uma vez que os estimadores de máxima verossimilhança $\hat{\Xi}$, $\hat{\Lambda}$ e Ψ de Ξ , Λ e Ψ sob H_0 , então a estatística

$$Z_{n} = -2 \log \lambda_{n} = n \left[\tilde{F} \left(\hat{\tilde{\varrho}} \right) - \tilde{F} \left(\hat{\varrho} \right) \right]$$
 (5.11)

tem distribuição limite χ_4^2 e o valor p = 0,0002. Como Z_n = = 22,56, H_o rejeitada ao nível de significância α = 0,0002. Os números de parâmetros livres e o número de g. ℓ estão na tabela 5.3.

Tabela 5.3

| | PARÂMETI | | | | |
|------------------------|----------|-----|---|-----------------|--|
| N(•) | 11 X | Δ | Ψ | SOMA | |
| não restrito | 4 | . 9 | 5 | 18 ^ℓ | |
| sob II _o | 4 9 1 | | | 14 ^m | |
| | . 4 | | | | |

C - Testes de Hipóteses sobre E

C-I) Seja

$$H_0: \xi_2 - \xi_1 = 0$$
, contra (5.14)
 $H_1: \xi_2 - \xi_1 \neq 0$,

onde
$$\xi_1' = \begin{bmatrix} \xi_{11} & \xi_{21} \end{bmatrix}$$
 e $\xi_2' = \begin{bmatrix} \xi_{12} & \xi_{22} \end{bmatrix}$.

Testar (5.14) é o mesmo que testar

$$H_0: B \subseteq C = 0$$
 , contra (5.15)

$$H_1: \underline{B} \underline{\Xi} \underline{C} \neq \underline{0}$$
,

onde
$$B = I_2$$
 e $C' = \begin{pmatrix} -1 & 1 \end{pmatrix}$.

Os elementos da matriz $\Xi = \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix}$ poden ser colocados na forma vetorial

$$\underline{\xi}' = \begin{bmatrix} \underline{\xi}_1' & \underline{\xi}_2' \end{bmatrix} = \begin{bmatrix} \xi_{11} & \xi_{21} & \xi_{12} & \xi_{22} \end{bmatrix}.$$
 (5.16)

Então a Hipótese (5.15) pode ser representada na forma

$$H_0: (\underline{C}' \otimes \underline{B})\xi = \underline{O}$$
, contra (5.17)

$$H_1: (C' \otimes B)_{\xi} \neq 0$$
,

onde & é dado por (5.16) e

$$C' \otimes B = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$

Com a Hipótese na forma (5.17), podemos aplicar o Teste de Wald.

A matriz estimador de máxima verossimilhança não restrito $\hat{\xi}$ da matriz $\hat{\xi}$, $\hat{\epsilon}$ dado em 5.1.5. Então, $\hat{\hat{\xi}} = \begin{pmatrix} \hat{\xi}_1 & \hat{\xi}_2 \end{pmatrix}$, logo $\hat{\xi}' = \begin{pmatrix} \hat{\xi}_1 & \hat{\xi}_2 \end{pmatrix} = \begin{pmatrix} 0.5078 & 0.7566 & 1.3459 & 2.4067 \end{pmatrix}$ (5.18)

A matriz de covariâncias ξ e dada por

$$\hat{\underline{\Sigma}} = \left[\underline{\underline{I}}_{11}(\hat{\underline{0}}) - \underline{\underline{I}}_{12}(\hat{\underline{0}}) \ \underline{\underline{I}}_{22}(\hat{\underline{0}}) \ \underline{\underline{I}}_{21}(\hat{\underline{0}})\right]^{-1},$$

onde $\hat{\theta}$ é um vetor constituído pelos elementos dos E.M.V. não restritos $\hat{\Xi}$, $\hat{\Lambda}$ e $\hat{\Psi}$ de Ξ , Λ e Ψ , respectivamente.

A matriz 🖔 , é dada por

$$\hat{\Sigma} =
\begin{cases}
0,0165 \\
-0,0041 & 0,0223 \\
0,0090 & -0,0054 & 0,0397 \\
-0,0047 & 0,0186 & -0,0163 & 0,0706
\end{cases}$$
(5.19)

Multiplicando à esquerda de \(\tilde{\Sigma} \) por (C' \(\tilde{\Sigma} \) B) e à direita por C \(\tilde{\Sigma} \) B', obtemos

$$(\underline{C} \otimes \underline{B}) \hat{\underline{\Sigma}} (\underline{C} \otimes \underline{B}') = \begin{cases} 0,0383 & -0,0103 \\ -0,0103 & 0,0558 \end{cases}$$
 (5.20)

Temos que

$$(\underline{C}' \otimes \underline{B}) \hat{\underline{\xi}} = [0,9381 \quad 1,6501]', \qquad (5.21)$$

onde

$$(\hat{\xi}_{12} - \hat{\xi}_{11}) = 0.9381$$
 e $(\hat{\xi}_{22} - \hat{\xi}_{21}) = 1.6501$.

Observe que (5.20) é a matriz de covariância de (C' \otimes B) $\hat{\xi}$. Multiplicando à esquerda da matrizinversa de (5.20) pela transposta de (5.21) e à direita por (5.21), obtemos um valor da estatística de Wald, W_{200} = 84,67.

Pelo teorema 3.6, sabemos que V_n tem como distribuição limite, um chi-quadrado central (χ_v^2) com v=2 graus de liber-

dade. Com o valor W200, verificamos que H é rejeitada.

C-II) O vetor ¿ é dado por

$$\xi' = \begin{bmatrix} 0,2 & 0,8 & 1,2 & 2,8 \end{bmatrix}$$
, (5.22)

então

$$(\underline{C}' \otimes \underline{B}) \xi = \begin{bmatrix} 1 & 2 \end{bmatrix}. \tag{5.23}$$

Considere a alternativa II1 dada por

$$H_1: (\underline{C}' \otimes \underline{B}) \ \underline{\xi} = \begin{bmatrix} 1 & 2 \end{bmatrix} \tag{5.24}$$

Com a alternativa (5.24), o Teorena 3.7 da a distribuição limite de W_n que é um qui-quadrado não central $\chi^2_v(\delta)$, com v=2 e parâmetro de não centralidade $\delta = 103,16$.

Então o poder do teste sob H₁ é dado por

$$\beta_{H_1}(\xi) = P(\chi_2^2(103,16) > 84,67) = 0,8430.$$
 (5.25)

Para obter δ , usamos (3.75) e (5.24), onde $\sum_{\theta} (\underline{\theta})$ é obtido através de (3.68), ou seja $\sum_{\theta} (\underline{\theta}) = J_{11}(\underline{\theta})$

5.2 - Um exemplo de Aplicação do Modelo Geral (4.2)

5.2.1 -Donald M. Miller e Marjorie V. Lutz (1966), desenvolveram um trabalho que foi adaptado ao modelo de um experimento 2⁴.

Este trabalho é uma pesquisa na área de educação, sobre a técnica de ensino e aprendizagem do aluno.

David E. Wiley, William H. Schmidt e William J. Bramble (1973) modificaram o experimento inicial removendo um fator, o fator "Subject Matter" (S) porque foi considerado irrelevante. Assim, o novo experimento tem agora um modelo de um experimento fatorial completo 2³, cujos fatores são:

- 1) <u>Nível de Escolaridade (G)</u>. Os níveis deste fator, o primeiro grau (G₁) e o sexto grau (G₂), foram escolhidos para representarem os extremos dos graus elementares. Desta maneira, foi possível maximizar as oportunidades para observar quaisquer diferenças de julgamento dos professores, que podem ocorrer como resultado da variação de nível de graduação.
- 2) Abordagem do Professor (T). A abordagem centrada no professor (T₁) e a abordagem centrada no aluno (T₂), foram distinguidas como níveis deste fator na base do local de controle de atividade descrita, e a direção de atenção do aluno descrita. No caso do enfoque centrado no professor, o local e direção foram orientados para o professor; no caso do enfoque centrado no aluno, o local e direção foram orientados para o aluno.
- 3) Método de Ensino (M). O primeiro nível deste fator é o treino (M₁), que foi usado estritamente para se referir à atividade de ensino para memorização; o segundo (M₂) foi usa do para se referir a um enfoque do professor, no qual tenta desenvolver a compreensão do aluma através de processo diri-

gido para estimular a compreensão, sem recurso de memoriza - ção ou rotina rígida de aprendizagem .

Wiley, Schimidt e Bramble (1973) propuseram o uso da análise de estrutura de covariâncias para o problema.

O modelo apresentado é dado por

$$\sum_{i} = \underline{\Gamma} \underline{A} \Phi \underline{A}' \underline{\Gamma} + \underline{\theta}^{2}$$
 (5.26)

$$\underline{\underline{\Gamma}} = \underline{\underline{\Gamma}} \left[\underline{\underline{A}} \ \underline{\Phi} \ \underline{\underline{A'}} + \underline{\underline{\Psi}}^2 \right] \underline{\underline{\Gamma}} , \qquad (5.27)$$

onde a matriz A(pxk) é uma matriz conhecida, $\Gamma(pxp)$ é uma matriz diagonal de fatores escalares desconhecidos, $\Phi(kxk)$ é a matriz de covariâncias das variáveis latentes, simétrica, positiva definida; $\Psi^2(pxp)$ e $\Theta^2(pxp)$ matrizes diagonais que representam as variâncias dos erros.

Estes modelos são casos particulares do modelo geral de Yöreskog. De fato, fazendo $B = \Gamma$, $\Lambda = A = \Psi^2 = 0$ em (5.26) e $B = \Gamma$, $\Lambda = A = \Psi^2 = 0$ em (5.27), verificamos que anbos sa tisfazem ao modelo geral (4.2). A tabela abaixo apresenta as 12 combinações dos modelos (5.26) e (5.27).

| | <u> </u> | $\frac{\theta^2}{2} = \underline{\Psi}^2 \sigma^2 \underline{I}$ | g² ou ψ² |
|--------|-----------|--|--|
| | | (1) | (2) |
| | _ 2 ∑ | | |
| | 11 . | $\sum = \underline{A} \underline{\Lambda}^2 \underline{A}^* + \sigma^2 \underline{I}$ | $\Sigma = A \Lambda^2 \Lambda^4 + \Theta^2$ |
| | ∌≀ | $\sum_{n} = \underbrace{A}_{n} \underbrace{\Lambda^{2}}_{n} \underbrace{A'}_{n} + \sigma^{2} \underbrace{I}_{n}$ | $\underline{\Sigma} = \underline{A} \underline{\Lambda}^2 \underline{\Lambda}^* + \underline{\theta}^2$ |
| Hi | žn v | | |
| 11 | | (3) | (4) |
| . tr 1 | | | (4) |
| | | 2 | 2 |
| | ⊕l | $\sum_{n} = A \Phi A' + \sigma^2 I$ | $\sum_{n} = \sum_{n} \Phi \Lambda^{n} + \Theta^{2}$ |
| | | | |
| | | | |
| | 2 | (5) | (6) |
| | 7 ₹ | $\underline{\underline{\Sigma}} = \underline{\Gamma} \underline{A} \underline{\Delta}^2 \underline{A} \underline{\Gamma} + \sigma^2 \underline{I}$ | $\sum = \Gamma \wedge \Delta \wedge \Gamma + \Theta^2$ |
| | II ⊕ } | ou · | ou |
|) | | $\frac{1}{2} = \underline{r} \left[\underbrace{A} \underline{A}^2 \underline{A}^* + \sigma^2 \underline{I} \right] \underline{r}$ | $\sum_{n} = \underline{r} \left[\underline{\Lambda} \ \underline{\Lambda}^{2} \ \underline{\Lambda}^{1} + \underline{\Psi}^{2} \right] \underline{\underline{r}}$ |
| E-1 | | = -(~ - ~ -)- | =(~ -)- |
| - | | (7) | (8) |
| | | $\sum = \underline{r} \wedge \underline{\phi} \wedge \underline{r} + \sigma^2 \underline{I}$ | $\sum_{r} = \sum_{r} A \Phi A' \Gamma + \Theta^2$ |
| | ⊕1, | ou · | ou |
| | | $\sum_{\underline{r}} = \underline{r} \left[\underbrace{\Lambda} \Phi \underline{\Lambda} + \sigma^2 \underline{I} \right] \underline{r}$ | |
| | - | 7 - 1 (V A V + 0 1)1 | $\sum_{r} = \sum_{r} \left[\underbrace{\Lambda}_{r} \Phi_{r} \Lambda'_{r} + \Psi^{2}_{r} \right] \underline{\Gamma}_{r}$ |
| | | | |

onde $\phi = \underline{\Lambda}^2$ (kxk) é uma matriz diagonal.

O teste de adaptação do modelo foi aplicado em cada modelo da tabela anterior, e foi verificado que os modelos de 5 a 8 não se adaptam aos dados, restando os quatro anteriores. Des tes quatros modelos restantes, o menos restrito, o quarto modelo, foi o que melhor se adaptou aos dados.

O modelo (4) é dado por

$$\Sigma = A \Phi \Lambda' + \Theta^2$$
 (5.28)

onde a matriz A é dada por

A matriz de covariância amostral é

$$S = \begin{cases} 18,74 \\ 9,48 \cdots 18,80 \\ 15,51 & 7,32 \cdots 21,93 \\ 3,98 & 15,27 & 4,10 \cdots 26,62 \\ 15,94 & 4,58 & 13,79 & -2,63 \cdots 19,82 \\ 7,15 & 13,63 & 3,83 & 15,33 & 3,65 \cdots 16,81 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

$$S = \begin{cases} 18,74 \\ 9,48 \cdots 18,80 \\ 15,51 & 7,32 \cdots 21,93 \\ 7,15 & 13,63 & 3,83 & 15,33 & 3,65 \cdots 16,81 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

$$S = \begin{cases} 18,74 \\ 9,48 \cdots 18,80 \\ 15,51 & 7,32 \cdots 21,93 \\ 7,15 & 13,63 & 3,83 & 15,33 & 3,65 \cdots 16,81 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

$$S = \begin{cases} 18,74 \\ 15,94 & 4,58 & 13,79 & -2,63 \cdots 19,82 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

$$S = \begin{cases} 18,74 \\ 15,94 & 4,58 & 13,79 & -2,63 \cdots 19,82 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

$$S = \begin{cases} 18,74 \\ 15,94 & 4,58 & 13,79 & -2,63 \cdots 19,82 \\ 11,69 & 6,05 & 10,18 & 1,13 & 13,55 & 5,72 \cdots 16,58 \\ 2,49 & 12,35 & 0,03 & 16,93 & -0,86 & 14,33 & 2,99 \cdots 18,2 \end{cases}$$

de ϕ e θ^2 , a função

$$F(\sum) = \frac{N}{2} \left(\log(|\Sigma|) - \operatorname{tr}(\sum^{-1} S) \right)$$
 (5.29)

é minimizada. O método de Fletcher e Powell é novamente aplicado para este fim, através de um programa de computador escrito na linguagem FORTRAN - IV para I.B.M e

C.D.C 6.500.

Desta maneira, os estimadores de máxima verossimilhança $\hat{\Phi}$, $\hat{\theta}^2$ e $\hat{\Sigma}$ de Φ , $\hat{\theta}^2$ e $\hat{\Sigma}$ são determinados. Estes estimadores são

$$\widehat{\Phi} = \begin{cases} 9,16 \\ 0,75 & 0,70 \\ 0,63 & -0,05 & 0,43 \\ 0,62 & -0,51 & 1,13 & 5,21 \end{cases} , \qquad (5.30)$$

$$\hat{\theta}^2 = \text{diag}\left(1,52; 4,95; 8,25; 5,58; 1,95; 2,34; 5,76; 2,82\right)$$

$$\hat{\Sigma} = \begin{cases}
19,69 \\
7,75 \cdots 23,14 \\
13,88 & 7,96 \cdots 19,57 \\
3,46 & 18,40 & 5,40 \cdots 25,91 \\
16,39 & 3,93 & 11,93 & -0,54 \cdots 19,36 \\
5,97 & 14,36 & 6,01 & 14,40 & 4,94 \cdots 15,67 \\
12,11 & 4,14 & 9,36 & 1,40 & 12,94 & 4,97 \cdots 15,97 \\
1,69 & 14,58 & 3,44 & 16,32 & 0,47 & 13,36 & 2,22 \cdots 1,76
\end{cases}$$

A matriz de correlação estimada das variáveis latentes é dada por

$$S.E\left(\begin{array}{c} \Phi \end{array}\right) = \left(\begin{array}{c} 1,00 \\ 0,30 & 1,00 \\ 0,32 & -0,08 & 1,00 \\ -0,09 & -0,27 & 0,75 & 1,00 \end{array}\right), \quad (5.32)$$

e a matriz dos erros padrões dos parâmetros estimados de $\hat{\theta}^2$ õ S.E $\left[\frac{\theta^2}{2}\right]$ = diag $\left[0,83;1,41;1,88;1,60;0,96;0,97;1,21;0,92\right]$ (5.33)

5.3 - UM EXEMPLO EM ANALISE FATORIAL

A função $f_k(\Psi)$ em (2.15) é uma função somente de Ψ . Sabemos que esta função tem derivadas de primeira ordem continuas e que possue derivadas de segunda ordem. Para determinar os estimadores Ψ_{ii} $i=1,2,\ldots,p$ aplicamos o processo de minimiza ção de Flecther e Powell na função (2.15). Para acompanhar esse método iterativo, precisamos definir a função $f_k(\Psi)$ numa região $R_{\epsilon} = \left\{ \Psi_{ii}, i=1,2,\ldots,p \; \middle | \; \Psi_{ii} \geq \epsilon \right\}$, onde ϵ é positivo e arbitrariamente pequeno.

Se o mínimo de $f_k(\Psi)$ estiver en R_{ϵ} , dizemos que é una solução <u>própria</u>. Se, por outro lado, estiver na front f_{ϵ} , dizemos que a solução é <u>imprópria</u>. Aqui, tonaremos o vetor f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor o processo, tomamos f_{ϵ} vetor f_{ϵ} iniciar o processo, tomamos f_{ϵ} vetor f_{ϵ} vetor cujas coordenadas f_{ϵ} vetor f_{ϵ} vetor cujas coordenadas f_{ϵ} vetor f_{ϵ} vetor cujas coordenadas f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor cujas coordenadas f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento da diagonal da matriz f_{ϵ} vetor f_{ϵ} vetor f_{ϵ} vetor elemento de solução própria que apresentaremos aqui.

Os dados que utilizamos foram usados por Emmet, 1949, com n = 211 observações e p = 9 variáveis.

A matriz de covariâncias amostral é dada por

- 123 -

Para o nível de significância α = 0,10, a solução própria para k = 3 foi aceita. Dado a matriz S em (5.35) e a expressão

$$x_i^{(1)} = (1 - \frac{k}{2p})(\frac{1}{S^{11}}), i=1,2,...,p, os valores$$

de $x_i^{(1)}$ encontrados são

0,424; 0,414; 0,566; 0,250; 0,380; 0,267; 0,394; 0,595, e 0,290, respectivamente (5.36) Com os valores dos $x_i^{(1)}$, $i=1,2,\ldots,p$ em (5.36) e o valor inicial de $E^{(1)}=I_9$, o método de Fletcher e Powell é iniciado com o objetivo de determinar os estimadores dos λ_{ij} e dos ψ_{ii} , $i=1,2,\ldots,9$ e j=1,2,3.

A solução de máxima verossimilhança e a matriz estimada de E foram obtidas depois da quinta interação. O valor dado a e

que define a região R_{ϵ} , é ϵ = 0,005. Como critério de convergência, admitimos que o processo termina quando o valor absoluto de todas as derivadas de primeira ordem, forem inferiores a 0,0005.

A tabela abaixo fornece os resultados dos estimadores de $\underline{\Lambda}$ e $\underline{\Psi}$.

| i | $\hat{\lambda}_{	ext{i1}}$ | î _{i2} | $\hat{\lambda}_{i3}$ | ŷ ii | ^{∂Φ} k ^{∂Ψ} ii | (5.37) |
|---|----------------------------|-----------------|----------------------|---------|-------------------------------------|--------|
| 1 | 0,664 | 0,321 | -0,074 | 0,450 | 0,0000 | |
| 2 | 0,689 | 0,247 | 0,193 | 0,427 | 0,0000 | |
| 3 | 0,493 | 0,302 | 0,222 | 0,617 | -0,0000 | - |
| 4 | 0,837 | -0,292 | 0,035 | 0,212 | -0,0001 | - 1 |
| 5 | 0,705 | -0,315 | 0,153 | 0,381 | -0,0000 | |
| 6 | 0,819 | -0,377 | -0,195 | 0,177 | 0,0001 | 7-11 |
| 7 | 0,661 | 0,396 | 0,078 | 0,400 | 0,0000 | |
| 8 | 0,458 | 0,296 | -0,491 | 0,462 | 0,0000 | |
| 9 | 0,766 | 0,427 | 0,012 | 0,231 | -0,0000 | |

Matriz
$$E = (e_{ij})$$
 estimada (5.38)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------------------------------|--------------------------------|--|-------|---|------------------------|-----------------------|--------------|------|------|
| 1 2 3 4 5 6 7 8 | 0,01 0,01 -0,00 -0,18 | 0,31 -0,02 0,01 -0,02 -0,03 -0,01 0,04 | -0,01 | 0,15 -0,02 -0,11 -0,00 0,10 | -0,00 0,01 -0,14 | 0,26 0,00 -0,32 | 0,26 $-0,01$ | 3,31 | 0,17 |

APÊNDICE

O METODO DE MINIMIZAÇÃO DE FLETCHER E POWELL

$$F\left(\underline{\theta}^{(s+1)},\underline{X}\right) < F\left(\underline{\theta}^{(s)},\underline{X}\right).$$

Para o desenvolvimento do processo, temos que supor que as derivadas parciais de primeira e segunda ordem sejam contínuas, que exista pelo menos um ponto em Ω onde $\frac{\partial F}{\partial \theta} = 0$, e que $\theta^{(1)}$ seja arbitráriamente escolhido.

Teoricamente, o processo iterativo é desenvolvido até que o menor valor possível de F seja determinado. Na prática, alguns critérios podem ser apresentados, com o objetivo de estacionar o processo. Por exemplo: para dado $\varepsilon > 0$, arbitráriamente pequeno, o processo para, quando simultâneamente to

das as
$$\left| \frac{\partial F}{\partial \theta_i^{(s)}} \right| \leq \epsilon$$
, $i = 1, 2, ..., r$.

O vetor gradiente de F em $\theta = \theta^{(s)}$ existe e é dado por

$$g(s) = \left(\frac{\partial F}{\partial \theta}\right)_{\theta=\theta}(s) \tag{(A 1.1)}$$

onde e (s) é o vetor e na s-ésima interação. Definamos a seguinte direção:

$$\dot{g}(s) = -E(s)E(s)$$
 (A 1.2)

onde $\underline{E}^{(s)}$ (rxr) é uma matriz simétrica e positiva definida. Uma vez que estamos de posse da direção $\underline{d}^{(s)}$, então, tanto podemos determinar o ponto $\underline{\theta}^{(s+1)}$, como podemos detectetar o mínimo de F ao longo desta direção, ou seja, minimizar

$$f(\alpha) = F(\theta^{(s)} + \alpha d^{(s)}, X) \qquad (A 1.3)$$

em relação a $\alpha > 0$.

Pela regra da cadeia,

$$h(\alpha) = f'(\alpha) = \underline{d}^{(s)}, \quad g_{\alpha} \quad (\Lambda 1.4)$$

onde

$$\underline{g}_{\alpha} = \left(\frac{\partial F}{\partial \theta}\right)_{\theta} = \underline{\theta}(s) + \alpha \underline{d}(s) \tag{(A 1.5)}$$

Por (A 1.5), observamos que em $\alpha = 0$,

$$g_0 = g^{(s)}$$
, (A 1.6)

$$h(\underline{0}) = f'(\underline{0}) = \underline{d}^{(s)}, \quad \underline{g}^{(s)} = -\underline{g}^{(s)}, \quad \underline{E}^{(s)} \quad \underline{g}^{(s)}$$
 (A 1.7)

Como $E^{(s)}$ é positiva definida, então h(0) < 0, exceto quando $g^{(s)} = 0$. E se ocorrer que $g^{(s)} = 0$, então o mínimo de F é determinado quando $\theta = \theta^{(s)}$.

Se no estágio (s) $g^{(s)} \neq 0$, então para algum $\alpha^{(s)}$ a função $f(\alpha)$ é minimizada. Para a continuação do processo, precisa mos obter $E^{(s+1)}$. Para isto, se faz necessário definir

$$\theta^{(s+1)} = \theta^{(s)} + \alpha^{(s)} d^{(s)}$$
 (A 1.8)

e calcular os vetores

$$y(s) = \theta^{(s+1)} - \theta^{(s)}$$
, (A 1.9)

$$\underline{h}^{(s)} = g^{(s+1)} - g^{(s)}$$
 (A 1.10)

$$\underline{Z}^{(s)} = \underline{E}^{(s)} \cdot \underline{h}^{(s)}$$
 (A 1.11)

$$\beta^{(s)} = \gamma^{(s)}, \quad h^{(s)}$$
 (A 1.12)

$$\gamma^{(s)} = Z^{(s)}, \quad h^{(s)}$$
 (A 1.13)

Usando as equações de (A 1.8) a (A 1.13), obtemos

$$\underline{E}^{(s+1)} = \underline{E}^{(s)} + {1 \choose \beta(s)} \underline{y}^{(s)} \underline{y}^{(s)} - {1 \choose \gamma(s)} \underline{z}^{(s)} \underline{z}^{(s)}.$$
(A 1.14)

Para cada modelo matemático, os valores iniciais $\theta^{(1)}$ e $E^{(1)}$ escolhidos, são aqueles que proporcionam o menor número possível de interações para determinar o mínimo de F.

REFERÊNCIAS

- ACHCAR, J.A. (1976) "Análises de Trajetórias". Univer sidade de São Paulo (IME), Dissertação de Mestrado.
- |2| AFIFI, A.A. and AZEN, S.P. (1972) "Statistical Analysis". Academic Press, New York and London.
- |3| ANDERSON, T.W. (1958) "An Introduction to Multivariate

 Statistical Analysis". John Wiley & Sons, Inc.
- [4] COSTNER, H.L. (1949) "Theory, deduction and rules of correspondence". American Journal Sociology 75, 245-263.
- |5| DAVIDON, W.C. (1959) "A variable metric method for minimization". AEC Research and Development report,

 A.N.L. -
- of maximum Likelihood". Britsh Journal of Psychology,
 Statistical Section, no 2, 90-97.
- de de São Paulo (IME), Dissertação de Mestrado.
- [8] FLETCHER, R. and POWELL, M.J.D. (1963) "A rapidly convergent descent method for minimization". Comput. J., 2, 163-168.

- GRIZZLE, J.E. and ALLEN, D.M. (1969) "Analysis of Growth and Dose Response Curves". Biometrics, vol. 25, 357-381.
- program for minimizing a function of several variables".

 Res. Bull. 70-14. Educational Testing Service. Princeton, New Jersey.
- of unobservable variables in path analysis". Sociological Methodology. H.L. Costner ed 81-117, Jossy-Ross, London.
- | HECK, D.L. (1960) "Charts of some upper percentual points of distribution of the largest characteristic root". Annals of Mathematical Statistics, vol.31, 625-642.
- Holt, New York.
- HOTELLING, H. (1933) "Analysis of a complex of statistical variable into principal components". Journal of Educational Psychology, vol. 24, 417-441.
- 15 HOTELLING, H. (1936) "Simplified calculation of Prince cipal Components". Psycometrika, vol. 1, 27-35.

- | 16 | YÖRESKOG, K.G. (1967) "Some Contributions to Maximum Likelihood Factor Analysis". Psychometrika, vol.32, no. 4, 443-482.
- | 17 | YÖRESKOG, K.G. and GRUVAEUS, G. (1967) R.M.L.F.A.
 "A computer program for restricted maximum likelihood
 factor analysis". Research Memorandum 67-21. Princenton, M.J. Educational Testing Service.
- | 18 | YÖRESKOG, K.G. (1968) "New Methods in maximum likeling hood factor analysis". British Journal of Mathematical and Statistical Psychology, 21, 85-96.
- | 19 | YÖRESKOG, K.G. (1969) "A general approach to confirmatory Maximum Likelihood Factor Analysis". Psychometrica, vol. 34, no 2, 183-202.
- |20| YÖRESKOG, K.G. (1970-a) A general method for analysis of covariance structures". Biometrika, 57, 239-251.
- YORESKOG, K.G. and GOLDBERGER, A.S. (1972) "Factor analysis by generalized least squares". Psychometrika, vol. 37, no 3, 243-285.
- | 22 | YORESKOG, K.G. (1973) "Analysis of Covariance structure". Department of Statistics, University of Uppsala, UPPSALA, Sweden.

- yöreskog, K.G. (1973) "Analysing Psychological Data by Structure Analysis of Covariance Matrices". University Institute of Statistics, Uppsala, Sweden, Selected Publications, vol. 39.
- |24| LAWLEY, D.N. (1938) "A generalization of Fisher's z-test". Biometrika, no 30, 180-187.
- by the method of maximum likelihood". Proceeding of the Royal Society of Edinburgh, Série A, no 60, 64-82.
- |26| LAWLEY, D.N. (1942) "Further investigations in factor estimation". Proceeding of the Royal Society of Edinburg, Série A, no 62, 176-181.
- |27 | LAWLEY, D.N. (1943) "The application of the maximum likelihood method to factor analysis". British Journal of Psycology, no 33, 172-175.
- |28 | LAWLEY, D.N. and MAXWELL, A.E. (1971) "Factor analysis as a Statistical Method". The Butterworth Group.
- |29 | MILLER, D.M. and LUTZ, M.V. (1966) "Item sesign for an inventory of teaching pratices and learning situation". Journal of Educational Measurement, 3, 53-61.
- | MORRISON, D.F. (1976) "Multivariate Statistical Methods". Mc Graw Hill Book Company, 2ª edição.

- NEUDECKER, H. (1969) "Some theorems on matrix differentation with special reference to kronecker matrix products". Journal of the American Statistical Association, 64, 953-963.
- | 32 | PEARSON, K. (1901) "An lines and planes of closest fit to systems of points in space". Philosophical Magazine, Ser. 6, vol. 2,
- |33| POTTHOFF, R.F. and ROY, S.N. (1964) "A generalized multivariate analysis of variance model useful especially for growth curve problems". Biometrika, vol.51, p.p. 313-326.
- | 34 | RAO, C.R. (1973) "Linear Statistical Inference and its Applications". John Wiley & Sons (2. edição).
- | 35 | RODRIGUES, G. (1975) "A Multivariate Linear Model with latent Factor Structure". Tese de doutoramento.

 The Institute of Statistics, The Consolidated University of North Carolina, Chapel Hill, North Carolina.
- ROY, S.N. (1953) "On heuristies method of test construction and its uses in multivariate analysis". Ann.

 Math. Statist. no 24, 220-228.
- | 37| SCHATZOFF, M. (1966) "Exact distributions of Wilk's likelihood ratio criterion". Biometrika, no 53, 347-358.

- | 38 | SINGER, J.M. (1977) "Análise de Curvas de Crescimento". Universidade de São Paulo (IME), Dissertação de Mestrado.
- | 39 | SPEARMAN, C. (1904) "General intelligence objetively determined and measured". American Journal of Psycology, 15, 201-293.
- | 40| SPIVAK, M. (1965) "Calculus on Manifolds". W.A. Benjamin, Inc., New York.
- | 41 | THURSTONE, L.L. (1931) "Multiple Factor Analysis" .

 Psychological Review 38, 406-427.
- | 42 | THURSTONE, L.L. (1949) "Multiple Factor Analysis" .

 Chicago University.
- | 43| VETTER, W.J. (1970) "Derivative Operation on Matrices". IEEE Transaction on Automatic Control, AC, 15,
 241-244.
- | 44| WAAL, D.J. (1975) "Parametric Multivariate Analysis,
 Part I". Department of Statistics, University of North
 Carolina, Chapel Hill.
- | 45| WILKS, S.S. (1932) "Certain generalization in analy sis of Variance". Biometrika, no 24, 471-494.