

INTERPRETAÇÃO GEOMÉTRICA DO TEOREMA DE JAMES  
EM SUA VERSÃO MULTIVARIADA NÃO CENTRAL.

REGINA CÉLIA DE CARVALHO PINTO MORAN

Dissertação apresentada no Instituto de  
Matemática, Estatística e Ciência da Com  
putação da Universidade Estadual de Campi  
nas como requisito parcial para a obten  
ção do título de Mestre em Estatística.

Orientador: Prof. Dr. JOSÉ NORBERTO W. DACKS

CAMPINAS  
1979

UNICAMP  
BIBLIOTECA CENTRAL

ã Milena  
ao Paulo  
e  
a meus pais

A autora agradece:

— ao Prof. Dr. Ubiratan D'Ambrósio, Diretor do Instituto de Matemática Estatística e Ciência da Computação, pelas facilidades concedidas para a realização deste trabalho.

— ao Prof. Dr. Rubens Murillo Marques, Diretor da Fundação SEADE (Serviço Estadual de Análise de Dados Estatísticos), sob a orientação de quem iniciamos nossa carreira, de quem sempre recebemos grande incentivo e apoio, e que coincidentemente primeiro nos apresentou o Teorema de Cochran.

— à Prof. Maria Eliza Fini, Chefe do Departamento de Estatística pelo apoio e incentivo recebidos.

— ao Prof. Eduardo Ferrandis Ballester em especial, pelos cursos de alto nível que ministrou no Bacharelado de Estatística nas áreas de Análise Multivariada e Modelos Lineares, aos quais somou grande entusiasmo, despertando assim o interesse que temos nesta área. Pelo apoio e incentivo recebidos durante todo esse curso de Mestrado. Pela colaboração procurada para o resultante trabalho: "O Contraste de Hipóteses nos Modelos Lineares Multivariados", que, tendo ampliado consideravelmente nossa visão nesta área, sugeriu este trabalho. Ainda, pelo interesse que demonstrou neste trabalho através de discussões, sugestões e incentivo.

— ao Prof. Dr. José Norberto Dacks pela liberdade que nos permitiu na escolha do tema deste trabalho, pelo incentivo e pela confiança depositada.

— ao Prof. Dr. Antonio Paques pela solicitude com que resolveu nossas dúvidas de Álgebra Linear.

— à *Eugenia*, ao *Reinaldo* e ao *Sidnei* por nossas proveito<sub>u</sub> sas discussões, pelo apoio e amizade recebida.

— ao *Sr. Raul dos Santos* pelo paciente e perfeito trabalho datilográfico, bem como a presteza e responsabilidade com que fomos atendidos.

— Muito especialmente ao *Paulo*, a meus pais, a *Jenny* e a *Chris*.

## NOTAÇÕES USADAS:

$n$  : tamanho da amostra

$p$  : número de variáveis

$\vec{X}$  : vetor coluna  $p \times 1$ ,  $\vec{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$

$\vec{\mu}$  : vetor das médias,  $\vec{\mu}_{p \times 1} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$

$\Sigma_{p \times p}$  : matriz covariante (ou matriz de variancias e covariâncias).

$\vec{X} \equiv N(\mu, \Sigma)$  :  $\vec{X}$  é multinormal (p-variada) de parâmetros: vetor média  $\vec{\mu}$  e matriz covariante  $\Sigma$ .

$X_{p \times n}$  : matriz que representa uma amostra de tamanho  $n$  de uma população multinormal p-variada, onde:

$\vec{X}^i$  : a coluna  $i$  da matriz  $X$

$\vec{X}_i$  : a linha  $i$  da matriz  $X$

$$\vec{\bar{X}} = \begin{pmatrix} \bar{X}_1 \\ \vdots \\ \bar{X}_p \end{pmatrix}; \quad \bar{X}_i = \sum_{J=1}^n X_{iJ}, \quad i=1, \dots, p$$

$F'$  : denota a transposta de  $F$  (aparece aplicada a vetores linha, coluna, ou a matrizes).

$F^{-1}$  : denota a inversa de  $F$

$\otimes$  : produto de Kronecker, que aplicado às matrizes  $A$  e  $B$  significa:

$$A \otimes B = \begin{pmatrix} a_{11}^B & \dots & a_{1p}^B \\ \vdots & & \vdots \\ a_{p1}^B & \dots & a_{pp}^B \end{pmatrix}$$

$$- \quad (Z_1 \dots Z_p) \cdot \text{cov} \begin{pmatrix} Z_1 \\ \vdots \\ Z_p \end{pmatrix} = \begin{pmatrix} \text{cov}(Z_1, Z_1) & \dots & \text{cov}(Z_1, Z_p) \\ \vdots & & \vdots \\ \text{cov}(Z_p, Z_1) & \dots & \text{cov}(Z_p, Z_p) \end{pmatrix}$$

$$- \quad V(X') = \begin{pmatrix} \frac{\bar{X}'_1}{\bar{X}'_n} \\ \vdots \\ \frac{\bar{X}'_1}{\bar{X}'_n} \end{pmatrix} \cdot \text{cov} (\bar{X}'_1 \dots \bar{X}'_n), \quad \bar{X}'_i = \begin{pmatrix} \bar{X}_{in} \\ \vdots \\ \bar{X}_{in} \end{pmatrix}$$

—  $A \equiv W(n, \Sigma, \tau)$  :  $A_{p \times p}$  tem distribuição de Wishart não central com parâmetros  $n, \Sigma, \tau (n \geq p)$

—  $A \equiv W(n, \Sigma)$  :  $A_{p \times p}$  tem distribuição de Wishart central com parâmetros  $n$  e  $\Sigma$ .

—  $E(X)$  : Esperança de  $X$

$$[E(X)]_{iJ} = E(X_{iJ})$$

- $\text{tr}(A)$  : traço da matriz  $A$ .
- $r(A)$  : o posto de  $A$ .
- $Y \equiv \chi^2_{(n, \lambda)}$  :  $Y$  tem distribuição  $\chi^2$  não central com parâmetros  $n$  e  $\lambda$ .

$I$  : matriz identidade.

$\phi(A)$  : função característica de  $A$ .

$\phi(A, B)$  : função característica conjunta de  $A$  e  $B$ .

$E$  : um espaço euclidiano real (OBS! Estamos aqui sendo reduntantes pois um espaço real, com produto interno e de dimensão finita é, frequentemente, denominado um espaço euclidiano)

$\overset{\perp}{\oplus}$  : soma direta ortogonal

$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i$  : projeção ortogonal dos vetores  $\{v_1, \dots, v_n\}$  sobre o subespaço de sub-índice  $i$  envolvido na decomposição do espaço total.

$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_W$  : projeção ortogonal dos vetores  $\{v_1, \dots, v_n\}$  sobre o subespaço  $W$ .

. : produto interno definido sobre E. Sendo que,  
para  $\beta_i \in E$ ,  $i=1, \dots, l$  temos:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_l \end{pmatrix} \cdot (\beta_1 \dots \beta_l) = \begin{pmatrix} (\beta_1 \cdot \beta_1) & \dots & (\beta_l \cdot \beta_l) \\ (\beta_l \cdot \beta_1) & \dots & (\beta_l \cdot \beta_l) \end{pmatrix}$$

( , ) : produto escalar (na página 48 temos  $\|\alpha_K\|^2$ ,  
que é a norma de  $\alpha_K$  ao quadrado em relação ao  
produto escalar).

$N = \left\{ R \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$  : onde  $R_{l \times n}$ , então: N é o subespaço de E gera  
do pelas linhas de  $R \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$

$\subset$  : está contido

$\subsetneq$  : está contido propriamente (está contido e não  
coincide).

$A \in D^g$  : A é uma inversa geral de D.

| | : determinante ou módulo.

|| : módulo do determinante.

Nota: As citações dos autores não estão numeradas por se tra-  
tar de uma correspondência "um a um" com as obras listadas nas  
referências.



## ÍNDICE

|                                                                                    |       |
|------------------------------------------------------------------------------------|-------|
| INTRODUÇÃO .....                                                                   | i     |
| 1 - A DISTRIBUIÇÃO DE WISHART .....                                                | 1-13  |
| Propriedades da Distribuição Multinormal .....                                     | 1-3   |
| A Distribuição de Wishart não central .....                                        | 3-7   |
| A Função Característica de Wishart não central .....                               | 7-13  |
| 2 - TEOREMA DE COCHRAN .....                                                       | 13-21 |
| 3 - TEOREMA DE JAMES .....                                                         | 21-42 |
| 4 - INTERPRETAÇÃO GEOMÉTRICA DO TEOREMA DE JAMES .....                             | 42-51 |
| Algumas Consequências da Interpretação Geométrica<br>do Teorema de James .....     | 46-74 |
| 5 - INTERPRETAÇÃO GEOMÉTRICA DO MODELO LINEAR GERAL .....                          | 51-74 |
| Interpretação dos Estimadores no Modelo Linear Geral .....                         | 51-59 |
| O teste de Hipótese no Modelo Linear Geral e sua<br>Interpretação Geométrica ..... | 59-68 |
| COMENTÁRIO FINAL .....                                                             | 68-74 |

## INTRODUÇÃO

i.

Entre as aplicações mais frequentemente usadas da teoria estatística situam-se a análise de variância, covariância e regressão. Estas técnicas dão lugar a partições de certas "formas quadráticas vetoriais" no caso multivariado, análogas às partições de formas quadráticas no caso univariado.

Estas partições nos levam a "quadrados vetoriais" análogos às somas de quadrados ditas componentes. Não é pela dificuldade de formulação ou raridade prática que usualmente se trabalha no caso univariado, como por exemplo nos textos de Scheffé, Searle e Dugué et Girault, senão pela relativa dificuldade de tratamento e interpretações. Sob certas hipóteses de distribuição apropriadas, no entanto, ao invés de razões de componentes que nos levem a estatística F, teremos razões de determinantes dessas componentes resultando em estatísticas como a de Wilks ou outras, apropriadas para testar certas hipóteses de interesse.

Obviamente, se estamos interessados em funções dessas componentes, precipuamente interessa-nos conhecer suas distribuições e relações de interdependência.

Nosso objetivo se restringe a obter essas distribuições, relações de interdependência e interpretá-las geometricamente. Trata-se portanto de um trabalho meramente introdutório ao estudo dessas técnicas multivariadas.

Os teoremas de Cochran e James são obrigatórios no estudo dessas técnicas e, como de fato, do ponto de vista geométrico, podemos ver por exemplo a análise de variância como um método de

decompor vetores de observações em projeções sobre certos subespaços específicos- espaços estes correspondentes a diferentes fontes de variação e cada qual podendo ter uma significação bastante própria- tratamos também da interpretação geométrica dos citados teoremas.

Necessitamos, no decorrer de nosso trabalho, de certos resultados da Álgebra Linear que decorrem de aplicações sucessivas de resultados básicos; optamos por não os reunir em apêndice e sim na ordem de sua necessidade.

Compusemos o trabalho de cinco partes: a primeira sobre a distribuição de Wishart, parte integrante das seguintes; a segunda sobre o teorema de Cochran na sua versão multivariada não central, que demonstraremos apesar da relativa facilidade devido às sugestões de Anderson. Na terceira parte tratamos do teorema de James em sua versão multivariada não central; baseamo-nos em seu artigo, permitindo-nos uma estruturação de enunciado. Na quarta parte tratamos da interpretação geométrica do teorema de James, onde incluímos como aplicações: interpretação das distribuições de Wishart que figuram no teorema e demonstração da independência das estimativas usuais dos parâmetros de uma multinormal, obtidas através de uma amostra aleatória. A parte cinco trata do modelo Linear Geral, interpretado geometricamente, o que implicitamente é sempre feito, e que procuramos formalizar. Já havíamos tratado parte deste trabalho na publicação interna nº 57, Eduardo Ferrandis Ballester e eu. Resumimos nesta publicação com muita densidade de alguns destes resultados.

## 1 - A DISTRIBUIÇÃO DE WISHART

Durante todo nosso trabalho faremos uso da distribuição normal multivariada e suas propriedades; por estar exhaustivamente tratada em muitos textos, dentre os quais nos utilizamos principalmente dos de Anderson e Krishnirsagar, enunciaremos apenas as propriedades das quais nos utilizamos.

Propriedades da Distribuição Multinormal:

Seja  $\vec{X} \equiv N(\vec{\mu}, \Sigma)$ ,  $\vec{X}_{p \times 1}$ ,  $\vec{\mu}_{p \times 1}$  e  $\Sigma_{p \times p}$ ; então:

PROPRIEDADE 1.1 - Se  $P_{p \times p}$  for uma matriz de constantes e

$$\vec{Z} = P\vec{X}$$

então:

$$\vec{Z} \equiv N(P\vec{\mu}, P \Sigma P')$$

PROPRIEDADE 1.2 - Se  $\vec{K}_{p \times 1}$  for um vetor de constantes e  $P_{p \times p}$  for uma matriz de constantes e  $\vec{Y} = \vec{K} + P\vec{X}$

$$\vec{Y} \equiv N(\vec{K} + P\vec{\mu}, P \Sigma P')$$

PROPRIEDADE 1.3 - Se  $\Sigma$  for diagonal,  $\Sigma_{ii} \neq 0$ ,  $i = 1, \dots, p$  então  $X_i$  é independente de  $X_J \forall i, J = 1, \dots, p$ ,  $i \neq J$ .

PROPRIEDADE 1.4 - Se  $\vec{X}^* = \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix}$  e  $\Sigma = \begin{pmatrix} \Sigma_{11}(K \times K) & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad 1 \leq K < p$

então:

$$\vec{X}^* \equiv N \left( \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_K \end{pmatrix}, \Sigma_{11} \right)$$

PROPRIEDADE 1.5 - Se  $C / \Sigma = C C'$  e  $\vec{Y} = C^{-1} (X - \vec{\mu})$

$$\vec{Y} \equiv N(\vec{0}, I)$$

PROPRIEDADE 1.6 - Se  $\vec{X}^i \equiv N(\vec{\mu}, \Sigma)$ ,  $i=1, 2, \dots, n$ ,  $\vec{X}_{p \times 1}^i$  ;

$$V(X') = I_n \otimes \Sigma \text{ e } \vec{\bar{X}} = \begin{pmatrix} \bar{X}_1 \\ \vdots \\ \bar{X}_p \end{pmatrix}$$

então:

$$\vec{\bar{X}} \equiv N(\vec{\mu}, \frac{1}{n} \Sigma)$$

As aplicações da distribuição de Wishart são muitas e, entre elas, podemos citar que a matriz covariante amostral segue uma distribuição de Wishart e que a análise de variância no caso multivariado dá lugar a uma decomposição de uma distribuição de Wishart em soma de Wishart independentes, cada afirmação sujeita à hipótese de multinormalidade e independência das amostras. Não abordaremos esse assunto, mas salientamos que os testes de hipótese podem ser obtidos, alguns deles como funções de

de distribuições de Wishart independentes. Como não faremos uso das funções densidades (caso central e não central), não as trataremos aqui devido a sua relativa complexidade.

Passamos a definir a distribuição de Wishart e tratar as propriedades de nosso interesse no decorrer do trabalho.

A distribuição de Wishart não central

DEFINIÇÃO 1.7 - Seja  $X_{p \times n}$ ,  $\vec{X}^i \equiv N(\vec{\mu}^i, \Sigma)$ ,  $V(X') = I_n \otimes \Sigma$ ,  $n \geq p$  e  $A = XX'$ . Então A tem a distribuição de Wishart não central com parâmetros  $n$ ,  $\Sigma$ ,  $\tau$  onde  $\tau_{p \times p} = \mu\mu'$ ,  $\mu = (\vec{\mu}^1, \dots, \vec{\mu}^n)$ ,  $\tau$  sendo chamada parâmetro de descentralização. A notação usada será:

$$A \equiv W(n, \Sigma, \tau)$$

PROPRIEDADE 1.8 - Seja  $A \equiv W(n, \Sigma, \tau)$  e  $K_{s \times p}$ ,  $s \leq p$ , uma matriz de constantes com posto de K igual a s. Então:

$$K A K' \equiv W(n, K \Sigma K', K \tau K')$$

PROVA:

$$A = XX', \quad \vec{X}^i \equiv N(\vec{\mu}^i, \Sigma), \quad V(X') = I_n \otimes \Sigma$$

e

$$\mu\mu' = \tau, \quad \mu = (\vec{\mu}^1 | \dots | \vec{\mu}^n)$$

Então:

$$K A K' = K X X' K' = (KX) (KX)' = Y Y'$$

onde:

$$\vec{Y}^i \equiv N(K\vec{\mu}^i, K \Sigma K') \text{ e } V(Y') = I_n \otimes [K \Sigma K']$$

logo

$$YY' \equiv W(n, K \Sigma K', Ku(Ku')')$$

mas

$$Ku(Ku)' = K u u' K' = K \tau K'$$

portanto:

$$K A K' \equiv W(n, K \Sigma K', K \tau K')$$

PROPRIEDADE 1.9 - Sejam  $A \equiv W(m, \Sigma, \tau)$ ,  $B \equiv W(n, \Sigma, \tau^+)$  independentes, então  $C = A+B \equiv W(m+n, \Sigma, \tau+\tau^+)$ .

PROVA:

$$A = XX', \quad X_{p \times m}, \quad \vec{X}^i \equiv N(\vec{\mu}^i, \Sigma), \quad V(X') = I_m \otimes \Sigma$$

e

$$\mu = (\vec{\mu}^1 | \dots | \vec{\mu}^m), \quad \mu\mu' = \tau$$

$$B = YY' , Y_{p \times n} , \vec{Y}^i \equiv N(\vec{\eta}^i, \Sigma) , V(Y') = I_n \otimes \Sigma$$

e

$$\eta = (\vec{\eta}^1 | \dots | \vec{\eta}^n) , \eta \eta' = \tau^+$$

$$\text{Mas } Z = (X | Y)$$

$$Z_{p \times (m+n)} , \vec{Z}^i \equiv N(\vec{\theta}^i, \Sigma) , V(Z') = I_{m+n} \otimes \Sigma$$

e

$$\theta = (\vec{\mu}^1 | \dots | \vec{\mu}^m | \vec{\eta}^1 | \dots | \vec{\eta}^n)$$

$$\therefore Z Z' \equiv W(m+n , \Sigma , \theta \theta')$$

mas

$$Z Z' = X X' + Y Y' = A + B$$

e

$$\theta \theta' = (\mu | \eta) \begin{pmatrix} \mu' \\ \eta' \end{pmatrix} = \mu \mu' + \eta \eta' = \tau + \tau^+$$

PROPRIEDADE 1.10: Seja  $A \equiv W(n, \Sigma, \tau)$  então  $E(A) = n \Sigma + \tau$ .

PROVA:

$$A = XX' , X_{p \times n} , \vec{X}^i \equiv N(\vec{\mu}^i, \Sigma) , V(X') = I_n \otimes \Sigma$$

Mas

$$\begin{aligned} A &= ((X-\mu) + \mu) ((X-\mu) + \mu)' = \\ &= (X-\mu)(X-\mu)' + (X-\mu)\mu' + \mu(X-\mu)' + \mu\mu' \end{aligned}$$



Mas

$$E(X-\mu) = 0 \quad \text{e} \quad (X-\mu)(X-\mu)' \equiv W(n, \Sigma)$$

$$\therefore E(A) = n \Sigma + \mu\mu' = n \Sigma + \tau.$$

Observemos ainda, que, quando  $p = 1$ , a distribuição de Wishart não central se reduz a um múltiplo da  $\chi^2$  não central (Hogg Graig).

De fato: Se

$$p = 1, \Sigma = \sigma^2 \quad \text{e} \quad A \equiv W(n, \sigma^2, \mu\mu'),$$

onde

$$\mu = (\mu_1, \dots, \mu_n)$$

mas

$$A = XX' = \sum_{i=1}^n \tilde{X}^i \tilde{X}^{i'} = \sum_{i=1}^n X_i^2 \quad \text{onde } X_i \equiv N(\mu_i, \sigma^2)$$

$$\therefore \frac{1}{\sigma^2} A \equiv \chi^2(n, \lambda) \quad \text{onde } \lambda = \frac{\sum_{i=1}^n \mu_i^2}{\sigma^2}$$

O tratamento da distribuição de Wishart no caso central é inteiramente análogo, apenas as multinormais independentes são centradas na origem, e portanto os parâmetros serão  $n$  (tamanho da amostra) e  $\Sigma$  (matriz covariante das multinormais que compõem a amostra). Se  $A_{p \times p}$  tem distribuição de Wishart central com parâmetros  $n$ , ( $n \geq p$ ), e  $\Sigma$ , denotaremos:

$$A \equiv W(n, \Sigma).$$

# A Função Característica de Wishart não Central

Antes de obtermos a função característica de uma matriz aleatória com distribuição de Wishart não central verificaremos o seguinte lema:

LEMA 1.11 - Se  $\theta$  de ordem  $p$  é simétrica, e  $M$ , simétrica de mesma ordem, é positiva definida, então existe  $B$  não singular tal que:

$$B' M B = I \quad \text{e} \quad B' \theta B = D, \text{ onde:}$$

$I$  é a identidade de ordem  $n$  e  $D$  é diagonal de mesma ordem.

Verificação:

Como  $M$  é p.d,  $\exists E$  não singular /  $E' M E = I$ .

Seja  $E' \theta E = B^*$ , então  $B^*$  é simétrica e  $\exists C, CC' = I$  tal que:

$$C' B^* C = D, \quad D \text{ diagonal.}$$

Tomando  $B = E C$  temos;  $B$  não singular e:

$$B' M B = C' E' M E C = C' I C = I$$

e

$$B' \theta B = C' E' \theta E C = C' B^* C = D$$

Função característica de uma matriz aleatória com distribuição de Wishart não central:

Obtemos para o caso não central, sendo o caso de Wishart central inteiramente análogo e desenvolvido no texto de Wilks.

A função característica de uma distribuição de Wishart pode ser obtida mais facilmente da distribuição das observações.

Se  $A \equiv W(n, \Sigma_{p \times p}, \tau)$  então temos:

$$A = Z Z', \quad Z_{p \times n}, \quad \vec{Z}^i \equiv N(\vec{\mu}, \Sigma)$$

e

$$V(Z') = I_n \otimes \Sigma$$

Seja  $\theta_{p \times p}$ , simétrica, então a função característica de:

$$(A_{11}, A_{22}, \dots, A_{pp}, 2A_{12}, 2A_{13}, \dots, 2A_{p-1,p})$$

é obtida por:

$$E\{\exp[i \operatorname{tr}(A\theta)]\}, \quad \text{onde} \quad \operatorname{tr}[C_{p \times p}] = \sum_{i=1}^p C_{ii}$$

Mas:

$$E\{\exp[i \operatorname{tr}(A\theta)]\} = E\{\exp\{i \operatorname{tr}\left[\left(\sum_{\alpha=1}^n \vec{Z}^{\alpha} \vec{Z}^{\alpha'}\right)\theta\right]\}\} =$$

$$E\{\exp(i \sum_{\alpha=1}^n \operatorname{tr}(\vec{Z}^{\alpha} \vec{Z}^{\alpha'} \theta))\} =$$

$$= E\{\exp(i \sum_{\alpha=1}^n \operatorname{tr}(\vec{Z}^{\alpha'} \theta \vec{Z}^{\alpha}))\} =$$

$$= E\left\{\exp\left(i \sum_{\alpha=1}^n \vec{Z}^{\alpha'} \theta \vec{Z}^{\alpha}\right)\right\} =$$

$$E\left\{\prod_{\alpha=1}^n \exp\left(i \vec{Z}^{\alpha'} \theta \vec{Z}^{\alpha}\right)\right\} = \prod_{\alpha=1}^n E\left\{\exp\left(i \vec{Z}^{\alpha'} \theta \vec{Z}^{\alpha}\right)\right\} =$$

$$= \left[E \exp\left(i \vec{Z}' \theta \vec{Z}\right)\right]^n, \text{ onde } \vec{Z} \equiv N(\mu, \Sigma)$$

Pelo lema 1.11, temos que  $B$  não singular tal que:

$$B' \Sigma^{-1} B = I$$

e

$$B' \theta B = D$$

Tomando  $Y = B^{-1} \vec{Z}$  temos  $\vec{Y} \equiv N(B^{-1} \vec{\mu}, B^{-1} \Sigma B^{-1})$

ou seja

$$\vec{Y} = N(B^{-1} \vec{\mu}, I)$$

logo:

$$E\left\{\exp\left(i \vec{Z}' \theta \vec{Z}\right)\right\} = E\left\{\exp\left(i (\vec{Y}' B' \theta B \vec{Y})\right)\right\} =$$

$$E\left\{\exp\left(i \vec{Y}' D \vec{Y}\right)\right\} = E\left\{\exp\left(i \sum_{J=1}^p d_{JJ} Y_J^2\right)\right\} =$$

$$= E\left\{\prod_{J=1}^p \exp\left(i d_{JJ} Y_J^2\right)\right\} = \prod_{J=1}^p E\left\{\exp\left(i d_{JJ} Y_J^2\right)\right\}$$

Mas  $E\left\{\exp\left(i d_{JJ} Y_J^2\right)\right\}$  é a função característica de uma  $\chi^2$ . Como um

grau de liberdade e parâmetro de não centralidade  $\eta_J$ , calculada no ponto  $d_{JJ}$ . Notemos que:

$$\eta_J = \frac{1}{2} [E(Y_J)]^2 = \frac{1}{2} [(B^{-1})_J \mu]^2$$

logo:

$$E\{\exp(i d_{JJ} Y_J^2)\} = \frac{1}{(1-2id_{JJ})^{1/2}} \exp\{-\eta_J [1-(1-2id_{JJ})^{-1}]\}$$

logo temos:

$$\prod_{J=1}^P E\{\exp(i d_{JJ} Y_J^2)\} = \prod_{J=1}^P \left\{ \frac{1}{(1-2id_{JJ})^{1/2}} \cdot \exp\{-\eta_J [1-(1-2id_{JJ})^{-1}]\} \right\}$$

Mas:

$$\prod_{J=1}^P \frac{1}{(1-2id_{JJ})^{1/2}} = \frac{1}{|I - 2i D|^{1/2}}$$

e

$$|I - 2i D| = |B' \Sigma^{-1} B - 2i B' 0 B| =$$

$$|B' [\Sigma^{-1} - 2i 0] B| = |B'| |\Sigma^{-1} - 2i 0| |B| = |B|^2 |\Sigma^{-1} - 2i 0|$$

mas

$$B' \Sigma^{-1} B = I \Rightarrow |B'| |\Sigma^{-1}| |B| = |I| \therefore$$

$$|B|^2 = \frac{1}{|\Sigma^{-1}|}$$

$$\therefore |I - 2iD| = \frac{|\Sigma^{-1} - 2i\theta|}{|\Sigma^{-1}|}$$

e temos:

$$\frac{1}{|I - 2iD|^{1/2}} = \frac{|\Sigma^{-1}|^{1/2}}{|\Sigma^{-1} - 2i\theta|^{1/2}}$$

Ainda:

$$\prod_{J=1}^p e^{-\eta_J [1 - (1 - 2id_{JJ})^{-1}]} = e^{-\sum_{J=1}^p \eta_J [1 - (1 - 2id_{JJ})^{-1}]}$$

Definindo:

$$\eta = \begin{bmatrix} \eta_1 & & 0 \\ & \ddots & \\ 0 & & \eta_p \end{bmatrix}, \text{ diagonal}$$

temos:

$$\sum_{J=1}^p \eta_J [1 - (1 - 2id_{JJ})^{-1}] = \text{tr}\{[I - (I - 2iD)^{-1}]\}$$

$$\begin{aligned} e, [I - 2iD]^{-1} &= [B' [\Sigma^{-1} - 2i\theta] B]^{-1} = \\ &= B^{-1} [\Sigma^{-1} - 2i\theta] B'^{-1} \end{aligned}$$

como:

$$I = B' \Sigma^{-1} B \text{ temos } B^{-1} \Sigma B'^{-1} = I$$

logo:

$$[I - [I - 2iD]^{-1}] = B^{-1} [I - [I - 2i\theta]^{-1}] B'^{-1}$$

logo:

$$\text{tr}[\eta [I - [I - 2iD]^{-1}]] =$$

$$= \text{tr}[\eta B^{-1} [I - [I - 2i\theta]^{-1}] B'^{-1}] =$$

$$= \text{tr}\{B'^{-1} \eta B^{-1} [I - [I - 2i\theta]^{-1}]\}$$

$$\therefore E\left\{\prod_{J=1}^p \exp i d_{JJ} Y_J^2\right\} = E\{\exp i Z' \theta Z\} =$$

$$= \frac{|\Sigma^{-1}|^{1/2}}{|\Sigma^{-1} - 2i\theta|^{1/2}} \exp - \text{tr}\{B'^{-1} \eta B^{-1} [I - [I - 2i\theta]^{-1}]\}$$

$$\therefore E\{\exp[i \text{tr}(A\theta)]\} =$$

$$= \frac{|\Sigma^{-1}|^{1/2n}}{|\Sigma^{-1} - 2i\theta|^{1/2n}} \exp - n \text{tr}\{B'^{-1} \eta B^{-1} [I - 2i\theta]^{-1}\}$$

Consequência: Se  $A \equiv W(n, \Sigma)$ , ou seja  $\eta = 0$ ,

temos:

$$E\{\exp[i \operatorname{tr}(A\theta)]\} = \frac{|\Sigma^{-1}|^{1/2n}}{|\Sigma^{-1} - 2i\theta|^{1/2n}}$$

que é a função característica de uma distribuição de Wishart cen  
tral com parâmetros  $n$  e  $\Sigma$ .

## 2 - TEOREMA DE COCHRAN

Este teorema, publicado em 1934, prova os principais resul-  
tados sobre distribuições conjuntas de formas quadráticas num sis-  
tema normal univariado. Um histórico deste teorema pode ser encon-  
trado no artigo de James. Vale lembrar que a versão não central  
deste teorema foi dada por Madow em 1940. Encontrado em sua ver-  
são univariadas em muitos textos (por exemplo, Cramér, Rao, Scheffé e ou-  
tros) e em sua versão multivariada central, por exemplo, no Anderson,  
figura já numa versão multivariada generalizada conhecida como Teo-  
rema de James em textos como o de Kshirsagar.

Este teorema não é em geral demonstrado em cursos de gradua-  
ção, embora comumente utilizado, pelas não menos comuns dificulda-  
des com a Álgebra. Por esta razão permaneceu para nós como algo  
muito forte e nada intuitivo. Muito forte porque dado um certo de-  
lineamento e aplicada a técnica de análise de variância, deparamo-  
-nos com partições de somas de quadrados expressando diferentes  
fontes de variação e, pelo teorema de Cochran, seguem suas distri-



buições e independência possibilitando os testes desejados, sob determinadas hipóteses. Quanto ao nada intuitivo, isto é muito pessoal...

Daremos uma versão multivariada não central do citado teorema que demonstraremos, apesar da relativa facilidade devida às sugestões encontradas no texto de Anderson. Para procedermos a sua demonstração trataremos dois lemas e um teorema cujas demonstrações seguem de aplicações sucessivas de resultados fundamentais da Álgebra Linear, e um teorema que trata o efeito de uma transformação ortogonal em multinormais independentes.

LEMA 2.1 - Se  $E_{n \times n}$  for simétrica e não singular existe  $F_{n \times n}$  não singular tal que:

$$F' E F = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \text{ onde:}$$

a ordem de  $I(-I)$  é o número de autovalores positivos (negativos) de  $E$ .

PROVA:

$$\exists G / G' E G = \begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_n \end{pmatrix}, \text{ onde:}$$

$$h_1 \geq \dots \geq h_g > 0 > h_{g+1} \geq \dots \geq h_n.$$

Definimos K como segue:

$$K = \begin{bmatrix} 1/\sqrt{h_1} & & & & 0 \\ & \ddots & & & \\ & & 1/\sqrt{h_g} & & \\ & & & 1/\sqrt{h_{g+1}} & \\ & 0 & & & \ddots & \\ & & & & & 1/\sqrt{h_n} \end{bmatrix}$$

Então:

$$K' G' E G K = (GK)' E (GK) = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

logo basta tomarmos  $F = KG$ .

LEMA 2.2 - Seja  $C_{n \times n}$  simétrica de posto  $r$  ( $r \leq n$ ).

Então existe uma matriz não singular A, tal que:

$$A' C A = \begin{pmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ onde:}$$

a ordem de  $I(-I)$  é igual ao número de autovalores positivos (negativos) de C.

PROVA: Como posto de C é r temos,  $\exists A_2 / C A_2 = 0$ ,

$$A_2 [n \times (n-r)]$$

Seja  $B_{n \times r} / (B|A_2)$  é não singular, então:

$$\begin{pmatrix} B' \\ A_2' \end{pmatrix} C (B|A_2) = \begin{pmatrix} B'CB & 0 \\ 0 & 0 \end{pmatrix}$$

onde  $B'CB$  é não singular de posto  $r$ .

Pelo lema 2.1 ]  $F_{r \times r}$  não singular tal que:

$$F' B' C B F = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$$

Portanto tomando  $A = (BF|A_2) = (B|A_2) \begin{pmatrix} F & 0 \\ 0 & I \end{pmatrix}$  temos o resultado desejado.

TEOREMA 2.3 - Sejam  $A_i$  de ordem  $n$  e simétricas,  $r(A_i) = n_i$ ,  $i = 1, \dots, K$ . Suponhamos  $\sum_{i=1}^K A_i = I$ .

Então uma condição suficiente para que exista uma matriz ortogonal  $B_{n \times n}$ ,  $B = (B_1 | \dots | B_K)$ ,  $B_i$   $(n \times n_i)$ ,  $i = 1, \dots, K$  tal que  $A_i = B_i B_i'$  é que  $\sum_{i=1}^K n_i = n$ .

DEMONSTRAÇÃO: Pelo lema 2.2,  $\forall i = 1, \dots, K$  ]  $D_i$  não singular tal que:

$$D_i \begin{pmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{pmatrix} D_i' = A_i$$

onde

$$r(I) + r(-I) = n_i$$

$$\text{Seja } D_i = [D_{1i} \mid D_{2i}], \quad D_{1i} (n \times n_i)$$

e sejam:

$$B_i = D_{1i}$$

e

$$Z_i = Y B_i, \text{ onde } Y = (Y_1, \dots, Y_n)$$

e

$$(B_i)_{\beta\alpha}, \quad \beta = 1, \dots, n \text{ e } \alpha = n_1 + \dots + n_{i-1} + 1, \dots, n_1 + \dots + n_i$$

Então se considerarmos a forma quadrática:

$$q_i = Y A_i Y', \text{ teremos:}$$

$$q_i = Y B_i \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} B_i' Y' = \vec{Z}_i \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \vec{Z}_i'$$

ou seja:

$$q_i = \sum_{\alpha = n_1 + \dots + n_{i-1} + 1}^{n_1 + \dots + n_i} C_{\alpha} z_{\alpha}^2, \quad C_{\alpha} = 1 \text{ ou } C_{\alpha} = -1$$

onde

$$\vec{Z}_{\alpha} = Y (\vec{B}_i^{\alpha}), \quad \alpha = n_1 + \dots + n_{i-1} + 1, \dots, n_1 + \dots + n_i$$

Mas se tomarmos  $q = YY' = Y(\sum_{i=1}^K A_i)Y'$  teremos:

$$q = \sum_{i=1}^K q_i = \sum_{\alpha=1}^n C_{\alpha} z_{\alpha}^2$$

e como  $q$  é positiva definida temos  $C_{\alpha} = 1, \alpha = 1, \dots, n$ .

Ainda:

$$A_i = B_i I B_i' = B_i B_i'$$

e como:

$$Y Y' = Y(\sum_{i=1}^K A_i)Y' = Y(\sum_{i=1}^K B_i B_i')Y' = Y B B' Y'$$

temos

$$B B' = I \therefore B \text{ é ortogonal.}$$

TEOREMA 2.4 - Suponhamos  $\vec{X}^{\alpha} \equiv N(\vec{\mu}^{\alpha}, \Sigma), \alpha = 1, \dots, n$ .

$$X = (\vec{X}^1) | \dots | (\vec{X}^n) \text{ e } V(X') = I_n \otimes \Sigma.$$

Seja  $B_{n \times n}$  uma matriz ortogonal, então:

$$Y_{p \times n} = X_{p \times n} B_{n \times n} \text{ é tal que:}$$

$$\vec{Y}^{\alpha} \equiv N(\mu B^{\alpha}, \Sigma), \alpha = 1, \dots, n$$

e

$$V(Y') = I_n \otimes \Sigma.$$

PROVA:  $\vec{Y}^{\alpha} = X \vec{B}^{\alpha}$  é multinormal pois é combinação linear de

multinormais.

$$\begin{aligned} E(\vec{Y}^\alpha) &= E(X \vec{B}^\alpha) = E\left(\sum_{\beta=1}^n B_{\beta\alpha} \vec{X}^\beta\right) = \\ &= \sum_{\beta=1}^n B_{\beta\alpha} E(\vec{X}^\beta) = \sum_{\beta=1}^n B_{\beta\alpha} \vec{\mu}^\beta = \\ &= (\vec{\mu}^1 | \dots | \vec{\mu}^n) \begin{pmatrix} B_{1\alpha} \\ \vdots \\ B_{n\alpha} \end{pmatrix} \end{aligned}$$

$$\therefore E(Y) = \mu B.$$

Ainda quaisquer linhas  $\alpha$  e  $\beta$  de  $Y$  teremos:

$$\begin{aligned} \text{cov}(\vec{Y}_\alpha, \vec{Y}_\beta) &= \vec{Y}_\alpha' \text{cov} \vec{Y}_\beta = B' \vec{X}_\alpha' \text{cov} \vec{X}_\beta B = \\ &= B' \sum_{\alpha\beta} I B = \sum_{\alpha\beta} B' B = \Sigma_{\alpha\beta} I \end{aligned}$$

$\therefore V(Y') = I_n \otimes \Sigma$ , ou seja, as colunas de  $Y$  são multinormais e independentes.

Teorema de Cochran:

Suponhamos  $\vec{X}^\alpha \equiv N(\vec{\mu}^\alpha, \Sigma_{p \times p})$ ,  $\alpha = 1, \dots, n$ ,  $X = (\vec{X}^1 | \dots | \vec{X}^n)$ ,

$$V(X') = I_n \otimes \Sigma \text{ e } E(X) = \mu.$$

Sejam  $A_i$ ,  $n \times n$ ,  $i = 1, \dots, K$  matrizes simétricas tais que

$$I_{n \times n} = \sum_{i=1}^K A_i, \quad r(A_i) = n_i, \quad i = 1, \dots, K.$$

Então a condição necessária e suficiente para que  
 $X A_i X' \equiv W(n, \Sigma, \mu A_i \mu')$ , independente de  $X A_J X'$ ,  $i \neq J$ ,  $i, J=1, \dots, K$ ,  
 é que  $n = \sum_{i=1}^K n_i$ .

PROVA: Para mostrarmos a suficiência estamos nas condições  
 do teorema 2.3, logo:  $\exists C$ , ortogonal de ordem  $n$  tal que:

$$C = (C_1 | \dots | C_K) \text{ e } A_i = C_i C_i', (C_i)_{n \times n_i}.$$

Seja  $Y = XC$ , então pelo teorema 2.4 temos:

$$V(Y') = I_n \otimes \Sigma, E(Y) = \mu C \text{ e } \vec{Y}^\alpha \text{ multinormais, } \alpha=1, \dots, n$$

Então se  $Y_i = X_{p \times n} C_{i \text{ } n \times n_i}$ , temos:

$$E(Y_i) = \mu C_i, V(Y_i') = I_{n_i} \otimes \Sigma$$

e logo:

$$Y_i Y_i' \equiv W(n_i, \Sigma, \vec{\mu} C_i C_i' \vec{\mu}')$$

ou seja:

$$Y_i Y_i' \equiv W(n, \Sigma, \vec{\mu} A_i \vec{\mu}')$$

e como  $Y_i$  independe de  $Y_J$  temos:

$$Y_i Y_i' \text{ independe de } Y_J Y_J' \text{ mas:}$$

$$X A_i X' = X C_i C_i' X' = X C_i (X C_i') = Y_i Y_i'.$$

A condição é necessária pois:

$$X X' = X \left( \sum_{i=1}^K A_i \right) X' = \sum_{i=1}^K (X A_i X')$$

Mas

$$X X' \equiv W(n, \Sigma, u u')$$

e como  $X A_i X'$  é independente de  $X A_j X'$  temos

$$n = \sum_{i=1}^K r(A_i) = \sum_{i=1}^K n_i$$

### 3 - TEOREMA DE JAMES

Nosso objetivo principal está na interpretação geométrica deste teorema que estende o teorema de Cochran, no sentido de que sob as mesmas hipóteses fornece outras condições equivalentes para garantir os mesmos resultados com mais força, ou sejam, distribuições de Wishart e independência.

A importância das condições equivalentes reside mais na sugestão da interpretação geométrica do que no ponto de vista computacional, porque sempre que nos interessamos pelas distribuições de Wishart interessar-nos-á conhecer os postos das matrizes envolvidas. A força maior a que nos referimos repousa no fato de precisarmos supor apenas que as "formas quadráticas vetoriais" tenham



distribuição de Wishart ou (exclusivo) que sejam independentes para obtermos a implicação da condição referente aos postos que aparece no teorema de Cochran.

Antes de procedermos ao enunciado e demonstração do Teorema de James, passamos a obter os resultados necessários.

### Propriedades de Matrizes Idempotentes:

PROPRIEDADE 3.1 - A matriz simétrica  $A_{n \times n}$  é idempotente se e só se os autovalores de A são 1 ou 0.

PROVA: ( $\Rightarrow$ ) Seja P a matriz ortogonal de autovetores de A, ou seja:

$$P' A P = \lambda \text{ onde } \lambda = \left( \begin{array}{ccc|ccc} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_s & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array} \right), r(A) = s$$

então

$$\lambda^2 = P' A P P' A P = P' A I A P = P' A^2 P = P' A P = \lambda$$

logo

$$\lambda_{ii}^2 = \lambda_{ii}, i = 1, \dots, s \text{ e temos } \lambda_{ii} = 1, i=1, \dots, s$$

$$(\Leftarrow) \lambda_{ii} = 1, i=1, \dots, s \text{ e } \lambda_{ii} = 0, i = s+1, \dots, n$$

então:

$$\lambda^2 = \lambda \text{ ou seja } P' A P \cdot P' A P = P' A^2 P = P' A P$$

$$\Rightarrow P P' A^2 P P' = P P' A P P' \Rightarrow A^2 = A$$

pois  $P P' = I$ .

PROPRIEDADE 3.2 - Se a matriz simétrica  $A_{n \times n}$  é idempotente então:

$$\begin{cases} A_{ii} \geq 0 \\ \text{Se } A_{ii} = 0 \Rightarrow A_{iJ} = A_{Ji} = 0, \forall J \leq n. \end{cases}$$

PROVA:  $A = A^2$  então  $A_{ii} = \sum_{J=1}^n A_{iJ} A_{Ji} = \sum_{J=1}^n A_{iJ}^2 \geq 0$

e se  $A_{ii} = 0 \Rightarrow \sum_{J=1}^n A_{iJ}^2 = 0 \Rightarrow A_{iJ} = 0, \forall J \leq n$ .

PROPRIEDADE 3.3 - Seja  $I = \sum_{i=1}^K A_i$ , de ordem  $n$ , idempotentes e simétricas,  $r(A_i) = n_i$ , então

$$\exists P, P P' = I / P' A_i P = \lambda_i, \text{ onde } \lambda_i$$

é uma matriz diagonal cujos elementos diagonais são 1 ou 0.

PROVA: Vamos proceder por indução sobre o número de matrizes que se somam.

O teorema vale quando  $K = 1$ .

Vamos supor válido para  $(K-1)$  matrizes parcelas.

Então:

Seja  $P_1$  a matriz de autovetores ortonormais de  $A_1$ , então:

$$P_1' A_1 P_1 = \lambda_1 = \left( \begin{array}{c|c} I_{n_1 \times n_1} & 0 \\ \hline 0 & 0 \end{array} \right), \quad n_1 = r(A_1)$$

e

$$\begin{aligned} I &= P_1' P_1 = P_1' I P_1 = P_1' \left( \sum_{i=1}^K A_i \right) P_1 = \\ &= P_1' A_1 P_1 + \sum_{i=2}^K P_1' A_i P_1 \end{aligned}$$

logo

$$I = \lambda_1 + \sum_{i=2}^K P_1' A_i P_1$$

Seja

$$P' A_i P = \left( \begin{array}{c|c} B_i & C_i \\ \hline C_i' & D_i \end{array} \right),$$

$$(B_i)_{n_1 \times n_1}, (C_i)_{n_1 \times (n-n_1)}, (D_i)_{(n-n_1) \times (n-n_1)}$$

então:

$$I_{n \times n} = \left( \begin{array}{c|c} I_{n_1 \times n_1} & 0 \\ \hline 0 & 0 \end{array} \right) + \sum_{i=2}^K \left( \begin{array}{c|c} B_i & C_i \\ \hline C_i' & D_i \end{array} \right)$$

logo:

$$\sum_{i=2}^K B_i = 0, \quad \sum_{i=2}^K C_i = 0 \quad e$$

$$\sum_{i=2}^K D_i = I_{(n-n_1) \times (n-n_1)}$$

Mas  $P_1' A_i P_1$  são idempotentes pois:

$$P_1' A_i P_1 P_1' A_i P_1 = P_1' A_i^2 P_1 = P_1' A_i P_1$$

portanto:

$$(B_i)_{KK} \geq 0, K = 1, \dots, n_1$$

mas

$$\sum_{i=2}^K (B_i)_{KK} = 0$$

$$\text{logo } (B_i)_{KK} = 0 \text{ e temos } \begin{cases} (B_i)_{KJ} = (B_i)_{JK} = 0, & \forall K \leq n_1 \\ (C_i)_{KJ} = (C_{JK}) = 0 & \forall n_1 \leq K \leq n \end{cases}$$

logo

$$B_i = 0, C_i = 0 \text{ e } C_i' = 0.$$

Então:

$$I = \left( \begin{array}{c|c} I_{n_1 \times n_1} & 0 \\ \hline 0 & 0 \end{array} \right) + \sum_{i=2}^K \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_i \end{array} \right)$$

onde

$$\sum_{i=2}^K D_i = I_{(n-n_1) \times (n-n_1)}$$

Mas

$$P_1' A_i P_1 = \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_i \end{array} \right)$$

e como  $P_1' A_1 P$  são idempotentes temos  $D_1$  idempotentes e simétricas pois:

$$\left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_1 \end{array} \right) \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_1 \end{array} \right) = \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_1^2 \end{array} \right)$$

$$\therefore D_1 = D_1^2$$

Podemos portanto aplicar a hipótese de indução:

$$\exists P_{2(n-n_1)} \times (n-n_1) / P_2' P_2 = I \quad e$$

$$P_2' D_i P_2 = \lambda_i^+, \quad i = 2, \dots, K, \quad \text{onde } \lambda_i \text{ é diagonal e seus}$$

elementos diagonais são 1 ou 0.

$$\text{Seja agora } P = P_1 \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right)$$

então:

$P$  é ortogonal pois:

$$P P' = P_1 \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2' \end{array} \right) P_1' =$$

$$= P_1 \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 P_2' \end{array} \right) P_1' = P_1 I P_1' = P_1 P_1' = I$$

E temos que  $P$  diagonaliza todas as  $A_i$  pois:

$$\begin{aligned}
 P' A_1 P &= \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P'_2 \end{array} \right) P'_1 A_1 P_1 \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \\
 &= \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P'_2 \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \\
 &\left( \begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \left( \begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right) = \lambda_1
 \end{aligned}$$

e para  $i \geq 2$  temos:

$$\begin{aligned}
 P' A_i P_1 &= \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P'_2 \end{array} \right) P'_1 A_i P_1 \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \\
 &= \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P'_2 \end{array} \right) \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & D_i \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \\
 &= \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & P'_2 D_i \end{array} \right) \left( \begin{array}{c|c} I & 0 \\ \hline 0 & P_2 \end{array} \right) = \\
 &\left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & P'_2 D_i P_2 \end{array} \right) = \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & \lambda_i^+ \end{array} \right)
 \end{aligned}$$

Consequências:

Mas  $\lambda_i^+$ ,  $i \geq 2$  são sucessivas matrizes diagonais de  $(n-n_1) \times (n-n_1)$  com 1 e 0 na diagonal e como:

$$\sum_{i=2}^K D_i = I_{(n-n_1) \times (n-n_1)}$$

temos:

$$\sum_{i=2}^K \lambda_i^+ = I_{(n-n_1) \times (n-n_1)}$$

logo se  $(\lambda_i^+) \ell \ell = 1 \Rightarrow (\lambda_J^+) \ell \ell = 0$ ,  $J \neq i$ ;  $J, i=2, \dots, K$   
portanto

$$\lambda_i \lambda_J = 0, \forall i, J, i \neq J; i, J = 1, \dots, K.$$

Teorema de James:

Suponhamos  $\vec{X}^\alpha \equiv N(\vec{\mu}^\alpha, \Sigma)$ ,  $X_{p \times n}$ ,  $E(X) = \mu$  e  $V(X') = I_n \otimes \Sigma$ .

Sejam  $A_i$ , de ordem  $n$ ,  $r(A_i) = n_i$ ,  $i = 1, \dots, K$ ,  $K$  matrizes simétricas tais que  $I = \sum_{i=1}^K A_i$ .

Então as cinco condições a seguir são equivalentes:

1.  $A_i$  idempotente,  $i = 1, \dots, K$
2.  $A_i A_J = 0$ ,  $i \neq J$ ;  $i, J = 1, \dots, K$
3.  $n = \sum_{i=1}^K n_i$
4.  $X A_i X' \equiv W(n_i, \Sigma, \tau_i)$ ,  $i = 1, \dots, K$
5.  $X A_i X'$  independente de  $X A_J X'$   $i \neq J$ ,  $i, J = 1, \dots, K$

PROVA:

(1  $\Rightarrow$  2) Estamos em condições de aplicar a propriedade 3.3

logo:

$$\exists P \mid PP' = I \text{ e } P'A_iP = \lambda_i, i = 1, \dots, K \text{ e}$$

$$\lambda_i \lambda_J = 0 \forall i \neq J, i, J = 1, \dots, K.$$

portanto  $\forall i \neq J, i, J = 1, \dots, K$  temos:

$$0 = P'A_iPP'A_JP = P'A_iA_JP = PP'A_iA_JPP' = A_iA_J$$

(2  $\Rightarrow$  1),  $\forall_i$  temos:

$$A_i = A_i I = A_i \sum_{J=1}^K A_J = \sum_{J=1}^K A_i A_J = A_i^2$$

(1  $\Rightarrow$  3) Pela propriedade 3.3 temos:

$$P'A_iP = \lambda_i \text{ e } r(A_i) = r(\lambda_i) = t_r \lambda_i$$

$$\text{Como } P'P = I = \sum_{i=1}^K P'A_iP = \sum_{i=1}^K \lambda_i \text{ temos:}$$

$$n = t_r \left( \sum_{i=1}^K \lambda_i \right) = \sum_{i=1}^K t_r \lambda_i = \sum_{i=1}^K n_i$$

(3  $\Rightarrow$  1) Seja  $P_J$  a matriz de autovetores de  $A_J$ , então para cada  $J = 1, \dots, K$  temos:



$$\begin{aligned}
 P_J' P_J &= I = P_J' I P_J = P_J' \sum_{i=1}^K A_i P_J = \\
 &= P_J' A_J P_J + P_J' \left( \sum_{\substack{i=1 \\ i \neq J}}^K A_i \right) P_J = \\
 &= \left( \begin{array}{c|c} \lambda_J & 0 \\ \hline 0 & 0 \end{array} \right) + \left( \begin{array}{c|c} B & C \\ \hline C' & D \end{array} \right)
 \end{aligned}$$

logo  $C = 0$ ,  $C' = 0$  e  $D = I_{(n-n_J) \times (n-n_J)}$

mas

$$\begin{aligned}
 r \left( \begin{array}{c|c} B & 0 \\ \hline 0 & I \end{array} \right) &= r(P_J' \left[ \sum_{\substack{i=1 \\ i \neq J}}^K A_i \right] P_J) = r \left( \sum_{\substack{i=1 \\ i \neq J}}^K P_J' A_i P_J \right) \leq \\
 &\leq \sum_{\substack{i=1 \\ i \neq J}}^K r(P_J' A_i P_J) \leq \sum_{\substack{i=1 \\ i \neq J}}^K r(A_i) = n - r(A_J) = n - n_J
 \end{aligned}$$

e como

$$\begin{aligned}
 r(I) &= n = r \left[ \left( \begin{array}{c|c} A_J & 0 \\ \hline 0 & 0 \end{array} \right) + \left( \begin{array}{c|c} B & 0 \\ \hline 0 & I \end{array} \right) \right] \leq \\
 n_J + r \left( \begin{array}{c|c} B & 0 \\ \hline 0 & I \end{array} \right) &\leq n + (n - n_J) = n
 \end{aligned}$$

temos

$$\begin{aligned}
 r \left( \begin{array}{c|c} B & 0 \\ \hline 0 & I \end{array} \right) &= n - n_J \Rightarrow B = 0 \\
 \therefore P_J' A_J P_J &= \left( \begin{array}{c|c} \lambda_J & 0 \\ \hline 0 & 0 \end{array} \right) = \left( \begin{array}{c|c} I & 0 \\ \hline 0 & 0 \end{array} \right)
 \end{aligned}$$

$\therefore$  os auto valores de  $A_J$  são 0 ou 1 e pela propriedade 3.1 temos  $A_J$  idempotente.

(2 $\Rightarrow$ 4) Estamos em condições de aplicar o teorema 2.3 e temos:

$A_i = C_i C_i'$   $\therefore X A_i X' = X C_i C_i' X' = Y_i Y_i'$  e pelo teorema temos  $Y = XC = X(C_1 | \dots | C_K) = (Y_1 | \dots | Y_K)$ , as colunas de  $Y_i$  são multinormais independentes logo:

$$Y_i Y_i' \equiv W(n_i, \Sigma, E(Y_i)E(Y_i'))$$

$$E(Y_i)E(Y_i') = \mu C_i C_i' \mu' = \mu A_i \mu'$$

logo

$$\tau_i = \mu A_i \mu' \text{ e temos}$$

$$X A_i X' \equiv W(n, \Sigma, \mu A_i \mu')$$

(2 $\Rightarrow$ 5) Ainda pelos teoremas 2.3 e 2.4 temos:

$$X A_i X' = Y_i Y_i' \text{ independentes } i \neq J, i=1, \dots, K$$

Pois  $Y_i$  é independente de  $Y_J$ .

(5 $\Rightarrow$ 1) Temos  $X A_i X'$  independente de  $X A_J X'$ ,  $i, J=1, \dots, K; i \neq J$ .

Seja  $A_J$ ,  $\forall J$  fixo e consideremos  $P$  ortogonal tal que:

$$P' A_J P = D = \begin{pmatrix} D_J & 0 \\ 0 & 0 \end{pmatrix}, \quad D_J (n_J \times n_J) \text{ diagonal},$$

e seja  $Y = Y P$  então:

$$q_J = X A_J X' = X P D P' X' = Y D Y' = \sum_{i=1}^{n_J} D_{ii} \tilde{Y}^i \tilde{Y}^{i'}$$

Mas

$$\tilde{Y}^\alpha \equiv N(\tilde{\mu}^\alpha P, \Sigma) \quad \text{e} \quad V(Y') = I_n \otimes \Sigma$$

logo,

$$V_\alpha, \alpha = 1, \dots, n_J, \quad \tilde{Y}^\alpha \tilde{Y}^{\alpha'} \equiv W(1, \Sigma, \tau^\alpha)$$

Além disso, como  $I = \sum_{J=1}^K A_J$  temos:

$$q = Y Y' = Y(D + (I - D))Y' = Y D Y' + Y(I - D)Y'$$

e  $Y D Y'$  independente de  $Y(I - D)Y'$ , sendo:

$$Y(I-D)Y' = \sum_{i=1}^n (1-D_{ii}) \tilde{Y}^i \tilde{Y}^{i'}$$

Vamos proceder calculando a função característica conjunta de  $q_J$  e  $q - q_J$ ; usando o fato da independência vamos igualar a conjunta ao produto das duas individualmente e tirar as possíveis implicações.

$$\phi(q_J, q - q_J) = E\{\exp i \operatorname{tr}(q_J, q - q_J) \begin{pmatrix} \theta \\ \theta^* \end{pmatrix}\} =$$

$$= E\{\exp i \operatorname{tr}[q_J \theta + (q - q_J) \theta^*]\} =$$

$$\begin{aligned}
 & E\{\exp i \operatorname{tr}\left\{\left[\left(\sum_{\ell=1}^{n_J} D_{\ell\ell} \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta\right] + \left[\left(\sum_{\ell=1}^n (1-D_{\ell\ell}) \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right]\right\}\} = \\
 & = E\{\exp i \operatorname{tr}\left\{\left[\left(\sum_{\ell=1}^{n_J} D_{\ell\ell} \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta\right] + \left[\sum_{\ell=1}^{n_J} (1-D_{\ell\ell}) \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right] + \right. \\
 & \quad \left. + \left[\left(\sum_{\ell=n_J+1}^n \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right]\right\}\} = \\
 & = E\{\exp \operatorname{tr}\left\{\left(\sum_{\ell=1}^{n_J} \vec{Y}^{\ell} \vec{Y}^{\ell'} \left[D_{\ell\ell} \theta + (1-D_{\ell\ell}) \theta^*\right]\right) + \right. \\
 & \quad \left.\left(\sum_{\ell=n_J+1}^n \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right\}\} =
 \end{aligned}$$

(Como  $V(Y') = I_n \otimes \Sigma$  temos:)

$$\begin{aligned}
 & = E\{\exp i \operatorname{tr}\left\{\left(\sum_{\ell=1}^{n_J} \vec{Y}^{\ell} \vec{Y}^{\ell'} \left[D_{\ell\ell} \theta + (1-D_{\ell\ell}) \theta^*\right]\right)\right\}\} \\
 & E\{\exp i \operatorname{tr}\left\{\left(\sum_{\ell=n_J+1}^n \vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right\}\} = \\
 & = \left\{ \prod_{\ell=1}^{n_J} E\{\exp i \operatorname{tr}\left(\vec{Y}^{\ell} \vec{Y}^{\ell'} \left[D_{\ell\ell} \theta + (1-D_{\ell\ell}) \theta^*\right]\right)\} \right\} \cdot \\
 & \left\{ \prod_{\ell=n_J+1}^n E\{\exp i \operatorname{tr}\left[\left(\vec{Y}^{\ell} \vec{Y}^{\ell'}\right) \theta^*\right]\} \right\}
 \end{aligned}$$

Mas:

$$E\{\exp i \operatorname{tr}\left\{\vec{Y}^{\ell} \vec{Y}^{\ell'} \left[D_{\ell\ell} \theta + (1-D_{\ell\ell}) \theta^*\right]\right\}\}$$

é a função característica de uma Wishart não central,  $W(1, \Sigma, \tau^\ell)$ , no ponto  $[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]$ .  
 $E, E\{\exp i \operatorname{tr}[(\hat{Y}^\ell \hat{Y}^{\ell'})\theta^*]\}$  é a função característica de uma Wishart,  $W(1, \Sigma, \tau^\ell)$  no ponto  $\theta^*$ .

Logo temos o primeiro termo da última igualdade dado por:

$$\prod_{\ell=1}^n \frac{|\Sigma^{-1}|^{1/2}}{|\Sigma^{-1} - 2i[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]|^{1/2}} e^{i \operatorname{tr} \eta_\ell}$$

onde

$$\eta_\ell = -\{B'^{-1} \gamma_\ell B^{-1} [I - [I - 2i[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]]^{-1}\}$$

e

$$B / B' \Sigma^{-1} B = I, \quad B' \theta B = D$$

e

$$\gamma_\ell = \begin{bmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_p \end{bmatrix}, \quad c_i = \frac{1}{2} [E(\hat{Y}_i^\ell)]^2$$

e o segundo termo vem dado por:

$$\prod_{\ell=n_J+1}^n \frac{|\Sigma^{-1}|^{1/2}}{|\Sigma^{-1} - 2i\theta^*|^{1/2}} e^{i \operatorname{tr} \eta_\ell}$$

onde

$$\eta_{\ell} = - \{ B'^{-1} \gamma_{\ell} B^{-1} [I - [I - 2\theta^*]^{-1}] \}$$

$$B / B' \Sigma^{-1} B = I \quad e \quad B' \theta B = D$$

e

$$\gamma_{\ell} = \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_p \end{pmatrix}, \quad c_i = \frac{1}{2} (E(\tilde{Y}^{\ell}))_i^2$$

logo:

$$\phi(q_J, q - q_J) = \frac{|\Sigma^{-1}|^{1/2} n_J}{\prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2i[D_{\ell\ell}\theta + (-D_{\ell\ell})\theta^*]|^{1/2}} e^{i \sum_{i=1}^n \text{tr} \eta_i} \frac{|\Sigma^{-1}|^{1/2} (n - n_J)}{|\Sigma^{-1} - 2i\theta^*|^{1/2} (n - n_J)}$$

Por outro lado:

$$\phi(q_J) = \frac{|\Sigma^{-1}|^{1/2} n_J}{\prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2iD_{\ell\ell}\theta|} e^{i \text{tr} \sum_{\ell=1}^{n_J} \eta_{\ell}}$$

e

$$\phi(q - q_J) = \frac{|\Sigma^{-1}|^{1/2} n}{\prod_{\ell=1}^n |\Sigma^{-1} - 2i(1 - D_{\ell\ell})\theta^*|^{1/2}} e^{i \text{tr} \sum_{\ell=1}^n \eta_{\ell}}$$

$\eta_{\ell}$  definido anteriormente.

Fazendo  $\phi(q_J, q - q_J) = \phi(q_J) \cdot \phi(q - q_J)$  e, tomando o modulo, levando-se em conta que  $|e^{ix}| = 1, \forall x$ , temos:

$\forall \theta, \theta^*$  simétricas:

$$\left| \frac{|\Sigma^{-1}|^{\frac{1}{2} n_J} |\Sigma^{-1}|^{\frac{1}{2} (n-n_J)}}{\left\{ \prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2i[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]|^{\frac{1}{2}} \right\} |\Sigma^{-1} - 2i\theta^*|^{\frac{1}{2} (n-n_J)}} \right| =$$

$$\left| \frac{|\Sigma^{-1}|^{\frac{1}{2} n_J} |\Sigma^{-1}|^{\frac{1}{2} n}}{\left\{ \prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2iD_{\ell\ell}\theta|^{\frac{1}{2}} \right\} \left\{ \prod_{\ell=1}^n |\Sigma^{-1} - 2i(1-D_{\ell\ell})\theta^*|^{\frac{1}{2}} \right\}} \right|$$

ou seja:

$$\left| \left\{ \prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2i[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]|^{\frac{1}{2}} \right\} |\Sigma^{-1} - 2i\theta^*|^{\frac{1}{2} (n-n_J)} \right| \cdot \left| |\Sigma^{-1}|^{\frac{1}{2} n_J} \right| =$$

$$\left| \left\{ \prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2iD_{\ell\ell}\theta|^{\frac{1}{2}} \right\} \left\{ \prod_{\ell=1}^n |\Sigma^{-1} - 2i(1-D_{\ell\ell})\theta^*|^{\frac{1}{2}} \right\} \right|$$

A partir da igualdade acima, mostraremos que  $D_{\ell\ell} = 1, \forall \ell = 1, \dots, n_J$  por redução ao absurdo; como  $D_{\ell\ell} = 0, \ell = n_J + 1, \dots, n$  temos:

$$\left| \left\{ \prod_{\ell=1}^{n_J} |\Sigma^{-1} - 2i[D_{\ell\ell}\theta + (1-D_{\ell\ell})\theta^*]|^{\frac{1}{2}} \right\} \cdot |\Sigma^{-1}|^{n_J} \right| =$$

$$= \left| \prod_{\ell=1}^{n_J} \frac{1}{\pi} \left| \Sigma^{-1} - 2i D_{\ell\ell} \theta \right| \left| \Sigma^{-1} - 2i (1-D_{\ell\ell}) \theta^* \right| \right|$$

Suponhamos algum  $D_{\ell\ell} \neq 1$ , por exemplo  $D_{11}$  e tomemos:

$$\theta^* = \frac{\Sigma^{-1}}{2i(1-D_{11})} \text{ então:}$$

$$\left| \Sigma^{-1} - 2i(1-D_{11})\theta^* \right| = 0 \text{ e teremos:}$$

$$\left| \left\{ \prod_{\ell=1}^{n_J} \frac{1}{\pi} \left| \Sigma^{-1} - 2i \left[ D_{\ell\ell} \theta + (1-D_{\ell\ell}) \frac{\Sigma^{-1}}{2i(1-D_{11})} \right] \right| \right\} \left| \Sigma^{-1} \right|^{n_J} \right| = 0$$

logo:

$$\left\{ \prod_{\ell=2}^{n_J} \frac{1}{\pi} \left| \Sigma^{-1} - 2i \left[ D_{\ell\ell} \theta + (1-D_{\ell\ell}) \frac{\Sigma^{-1}}{2i(1-D_{11})} \right] \right| \right\} .$$

$$\left| \Sigma^{-1} - 2i \left[ D_{11} \theta + (1-D_{11}) \frac{\Sigma^{-1}}{2i(1-D_{11})} \right] \right|$$

Mas

$$\left| \Sigma^{-1} - 2i \left[ D_{11} \theta + (1-D_{11}) \frac{\Sigma^{-1}}{2i(1-D_{11})} \right] \right| =$$

$$\left| - 2i D_{11} \theta \right| \text{ (que é constante)}$$

e temos:



$$\left\{ \prod_{\ell=2}^{n_J} \left| \Sigma^{-1} - 2i \left[ D_{\ell\ell} \theta + (1-D_{\ell\ell}) \frac{\Sigma^{-1}}{2i(1-D_{11})} \right] \right| \right\} = 0$$

Tomando  $\theta = \frac{\Sigma^{-1}}{2}$  temos:

$$\begin{aligned} & \left\{ \prod_{\ell=2}^{n_J} \left| \Sigma^{-1} - 2i \left[ D_{\ell\ell} \frac{\Sigma^{-1}}{2} + (1-D_{\ell\ell}) \frac{\Sigma^{-1}}{4i(1-D_{11})} \right] \right| \right\} = \\ & = \left| \Sigma^{-1} \right|^{n_J} \prod_{\ell=2}^{n_J} \left| \left[ I - 2i \left[ \frac{D_{\ell\ell}}{2} I + \frac{(1-D_{\ell\ell})}{4i(1-D_{11})} I \right] \right] \right| = 0 \end{aligned}$$

implicará:

$$\prod_{\ell=2}^{n_J} \left( 1 - i D_{\ell\ell} + \frac{(1-D_{\ell\ell})}{4i(1-D_{11})} \right)^P = 0$$

logo para algum  $\ell$ ,  $\ell = 2, \dots, n_J$ , seja  $\ell = K$  devemos ter:

$$\left( 1 - i D_{KK} + \frac{(1-D_{KK})}{4i(1-D_{11})} \right) = 0$$

$$4i(1-D_{11}) + 4D_{KK}(1-D_{11}) + (1-D_{KK}) = 0$$

$$\Rightarrow (1-D_{11}) = 0 \Rightarrow D_{11} = 1 \text{ absurdo}$$

logo

$$D_{\ell\ell} = 1, \ell = 1, \dots, n_J \text{ e portanto}$$

$$A_J \text{ idempotente, } J = 1, \dots, K.$$

(4 $\Rightarrow$ 1) Para cada  $i$  fixo,  $i = 1, \dots, K$  temos,  $\exists$   $P$  tal que  $P'P = I$  e

$P' A_i P = D$ ,  $D = \left( \begin{array}{c|c} \lambda_i & 0 \\ \hline 0 & 0 \end{array} \right)$ ,  $\lambda_i$  ( $n_i \times n_i$ ), a matriz de auto valores distintos de zero de  $A_i$ .

logo

$$A_i = P \left( \begin{array}{c|c} \lambda_i & 0 \\ \hline 0 & 0 \end{array} \right) P' \text{ e tomando,}$$

$$P = (P_1 \mid P_2), \quad P_1 (n \times n_i) \text{ e } P_2 (n \times (n-n_i))$$

temos:

$$A_i = (P_1 \mid P_2) \left( \begin{array}{c|c} \lambda_i & 0 \\ \hline 0 & 0 \end{array} \right) \begin{pmatrix} P_1' \\ P_2' \end{pmatrix} = P_1 \lambda_i P_1'$$

logo  $X A_i X' = X P_1 \lambda_i P_1' X'$  e chamando:

$$Z = X P_1 \text{ temos:}$$

$$X A_i X' = Z \lambda_i Z'$$

$$\text{Ainda, o elemento } (Z \lambda_i Z')_{JJ} = \sum_{\ell=1}^{n_J} (\lambda_{\ell\ell} Z_{J\ell})^2$$

Sendo  $m$  o número de auto valores positivos de  $A_i$  temos:

$$P \left[ \sum_{\lambda_{\ell\ell} > 0} \lambda_{\ell\ell} (Z_{J\ell})^2 < \sum_{\lambda_{\ell\ell} < 0} -\lambda_{\ell\ell} (Z_{J\ell})^2 \right] \geq$$

$$P[h \sum (Z_{J\ell})^2 < s \sum (Z_{J\ell})^2] \text{ onde:}$$

— : a soma à esquerda, nos dois termos das desigualdades acima tem  $m$  termos  $\ell$ , à direita,  $(n_i - m)$  termos.

$$s = \min_{\lambda_{\ell\ell} < 0} |-\lambda_{\ell\ell}|$$

e

$$h = \max_{\lambda_{\ell\ell} > 0} |\lambda_{\ell\ell}|$$

Mas como  $Z_{J\ell} \sim N(u_{J\ell}, \Sigma_{JJ})$  e independentes temos que

$\sum_{\ell} (Z_{J\ell})^2$  é proporcional a uma qui-quadrado não central e portanto:

$$P[h \sum_{\ell} (Z_{J\ell})^2 < s \sum_{\ell} (Z_{J\ell})^2] = P[h \psi < s \eta],$$

onde  $\psi$  e  $\eta$  são qui-quadrado não centrais com  $m$  e  $n_i - m$  graus de liberdade respectivamente.

Mas:

$$P_r[h \psi \leq s \eta] \geq P[(\psi \leq 1) \cap (\eta \geq h/s)] > 0$$

pois

$$1 = \frac{s}{h} \frac{h}{s} \leq \frac{s}{h} \eta$$

Mas

$$P\left[\sum_{\ell=1}^{n_i} \lambda_{\ell\ell} (Z_{J\ell})^2 < 0\right] > 0, \text{ absurdo pois:}$$

$$\text{Chamando } B_i = X A_i X' = Z \lambda_i Z' \equiv W(n_i, \Sigma, \tau_i)$$

temos

$$(B_i)_{JJ} = (0 \dots \underset{J}{\underset{\uparrow}{1}} \dots 0) B_i \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \leftarrow J$$

$$\therefore (B_i)_{JJ} \equiv W(n_i, K \Sigma K', K \tau_i K')$$

$$\equiv W(n_i, \Sigma_{JJ}, (\tau_i)_{JJ})$$

que é proporcional a uma qui-quadrado não central.

Ainda como  $A = X A_i X'$  temos:

$$|\phi(A)| = |\phi(X A_i X')|, \text{ usando } |e^{ix}| = 1, \forall x$$

temos:

$$\left| \prod_{\alpha=1}^{n_J} |\Sigma^{-1} - 2i \lambda_{\alpha\alpha} \theta|^{1/2} \right| = \left| |\Sigma^{-1} - 2i\theta|^{1/2} \right|^{n_J}$$

pois

$$X A_i X' = \sum_{\alpha=1}^{n_J} \lambda_{\alpha\alpha} \hat{Z}^\alpha \hat{Z}^{\alpha'}$$

e tomando  $\theta = \frac{\Sigma^{-1}}{2}$

$$\left| |\Sigma^{-1}|^{1/2 n_J} \prod_{\alpha=1}^{n_J} |I - i \lambda_{\alpha\alpha} I|^{1/2} \right| = \left| |\Sigma^{-1}|^{1/2 n_J} \right| \left| |I - iI|^{1/2 n_J} \right|$$

$$\prod_{\alpha=1}^{n_J} |(1 - i \lambda_{\alpha\alpha})^{p/2}| = 2^{p n_J / 2}$$

$$\prod_{\alpha=1}^{n_J} (1 + \lambda_{\alpha\alpha}^2) = 2^{n_J} \therefore \lambda_{\alpha\alpha} = 1, \alpha = 1, \dots, n_J$$

pois

$\lambda_{\alpha\alpha} > 0$  e  $\lambda_{\alpha\alpha} \leq 1$ . De fato:

$$I = \sum_{i=1}^K A_i; P I P' = D_1 + \sum_{i>1} P A_i P'$$

Como  $(P A_i P')_{JJ} = \vec{P}_i A_i \vec{P}_i' \geq 0$ , temos  $\lambda_{\alpha\alpha} \geq 0 \Rightarrow A_i$  é p.s.d.

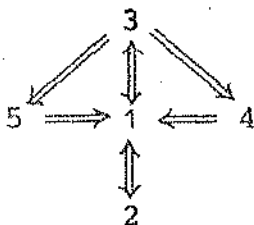
$$\therefore 1 = D_{\alpha\alpha} + (P A_i P')_{\alpha\alpha} \therefore D_{\alpha\alpha} \leq 1$$

e portanto  $(1 + \lambda_{\alpha\alpha}^2) \leq 2$ . Se  $\exists \lambda_{KK} < 1$

$$\Rightarrow (1 + \lambda_{KK}) < 2 \Rightarrow \prod_{\alpha=1}^{n_J} (1 + \lambda_{\alpha\alpha}) < 2^{n_J}$$

$$\therefore \lambda_{KK} = 1, \forall K.$$

Esboçamos abaixo um diagrama que resume como demonstramos o teorema de James:



#### 4 - INTERPRETAÇÃO GEOMÉTRICA DO TEOREMA DE JAMES:

A representação de uma amostra p-variada, com as variáveis vistas como pontos num espaço n-dimensional (n sendo o tamanho da

amostra) é tratada no texto de Dempster, entre outros tantos enfoques geométricos por ele apresentados. Nos textos de Scheffé e Diquê et Girault, entre outros, nota-se a ênfase ao tratamento "vetorial" dos problemas de regressão e análise de variância. Os teoremas de Cochran e James, no seu tratamento formal, não evidenciam a estrutura de espaço vetorial que podemos considerar e as identificações decorrentes dessa estrutura, das hipóteses e condições matriciais.

Julgamos estas identificações importantes pois propiciam interpretações bem significativas das distribuições que surgem nestes teoremas e de como surgem.

Isto veremos mais claramente na parte referente ao Modelo Linear Geral.

Em suma, as condições de idempotência, ortogonalidade e a referente aos postos das matrizes, sugerem respectivamente: projeções ortogonais sobre subespaços que dêem como soma direta um espaço total. É na consideração deste espaço total e sua conveniente decomposição que surge o que chamamos Interpretação Geométrica do Teorema de James, e que passaremos a tratar:

Interpretação Geométrica do Teorema de James (na sua versão multivariada não central).

TEOREMA - Sejam:  $E$  um espaço euclidiano real  $n$ -dimensional e  $\{v_1, \dots, v_n\}$  uma base ortonormal qualquer de  $E$ .

Seja

$$X_{p \times n}, \quad \vec{X}^i \equiv N(\vec{\mu}^i, \Sigma), \quad E(X) = \mu, \quad V(X') = I_n \otimes \Sigma.$$

Então os seguintes conjuntos I e II de condições são equivalentes:

I:  $A_i$ , de ordem  $n$ ,  $r(A_i) = n_i$ ,  $i = 1, \dots, K$  são  $K$  matrizes simétricas tais que:

$$I = \sum_{i=1}^K A_i \text{ e:}$$

1.  $A_i$  idempotente,  $i = 1, \dots, K$
2.  $A_i A_j = 0$ ,  $i \neq j$ ,  $i, j = 1, \dots, K$
3.  $n = \sum_{i=1}^K n_i$
4.  $X A_i X' \equiv W(n_i, \Sigma, r_i)$ ,  $i = 1, \dots, K$
5.  $X A_i X'$  independe de  $X A_j X'$ ,  $i \neq j$ ;  $i, j = 1, \dots, K$

II:  $\exists E_1, \dots, E_K$ , subespaços de  $E$ , tais que,

$$E = E_1 \perp \dots \perp E_K, \quad K \leq n \text{ e}$$

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i = A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad r(A_i) = n_i, \quad i = 1, \dots, K.$$

DEMONSTRAÇÃO:

II  $\Rightarrow$  I

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^K \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i = \left( \sum_{i=1}^K A_i \right) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \Rightarrow I = \sum_{i=1}^K A_i$$

Como  $E_i \perp E_J$ ,  $0 = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i (v_1 \dots v_n)_J =$

$$A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (v_1 \dots v_n)_J A_J' = A_i I A_J' = A_i A_J'$$

e como  $A_i = (I - \sum_{\ell \neq i} A_\ell)$ , temos:

$$(I - \sum_{\ell \neq i} A_\ell) A_J' = 0 \Rightarrow A_J' = A_J A_J'$$

e portanto  $A_i$  são simétricas  $i = 1, \dots, K$ .

Ainda  $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i = A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  é um sistema de geradores do subespaço

$E_i$  de fato:

Se  $\{\alpha_1 \dots \alpha_s\}$  é uma base de  $E_i$ :

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} = \lambda_{s \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \Rightarrow \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} = \lambda_{s \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i$$

$\therefore \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i$  é um sistema de geradores de  $E_i$  e temos 3.

I  $\Rightarrow$  II

Tomando:



$$E_i = \left\{ A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

Então:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \left( \sum_{i=1}^K A_i \right) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^K A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

logo  $E = E_1 \cup E_2 \dots \cup E_K$  (união vetorial) e como:

$$\left\{ A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\} \cdot \left\{ A_j \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}' = A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (v_1 \dots v_n) A_j' =$$

$$A_i A_j = 0 \text{ temos } E_i \perp E_j ; i \neq j; i, j = 1, \dots, K$$

$$\therefore E = E_1 \overset{\perp}{\oplus} E_2 \dots \overset{\perp}{\oplus} E_K$$

Algumas consequências da Interpretação Geométrica do Teorema de James.

Antes de entrarmos na parte que trata do Modelo Linear Geral, vamos tratar duas consequências específicas: a primeira interpreta o que são as distribuições de Wishart que aparecem nos teoremas de Cochran e James; a segunda é uma demonstração da independência entre a matriz covariante amostral e o vetor de médias amostrais (as estimativas usuais) de uma amostra multinormal  $p$ -variada de tamanho  $n$ .

4.1 - Observemos que as distribuições de Wishart independentes dos Teoremas citados (Cochran, James e sua Interpretação Geométrica), nada mais são que produtos de projeções ortogonais, e, por esta razão, não só as "formas quadráticas vetoriais" são independentes, mas também as próprias projeções.

De fato:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i = A_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e portanto:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i \cdot (v_1 \dots v_n)_i = A_i A_i' = A_i$$

e temos,

$$X A_i X' = X \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i \cdot (v_1 \dots v_n)_i X'$$

onde  $X \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_i$  é a projeção da matriz de dados, vista como p

pontos no espaço n-dimensional citado

e

$$\text{cov}[\overrightarrow{(X A_i)_L}, \overrightarrow{(X A_J)_K}] = \text{cov}(A_i \begin{pmatrix} x_{L1} \\ \vdots \\ x_{Ln} \end{pmatrix}, (x_{K1} \dots x_{Kn}) A_J) =$$

$$= A_i \begin{pmatrix} x_{i1} \\ \vdots \\ x_{in} \end{pmatrix} \text{cov} (x_{K1} \dots x_{Kn}) A_J = A_i \Sigma_{\ell_K} I A_J =$$

$$\Sigma_{\ell_K} A_i A_J \therefore \text{cov} [(\overrightarrow{X A_i})_{\ell}, (\overrightarrow{X A_J})_K] = 0 \text{ se } i \neq J, i, J=1, \dots, K.$$

4.2 - Usaremos aqui um conhecido resultado da Álgebra Linear:

— Se um vetor  $\beta$  é uma combinação linear de um conjunto ortogonal de vetores não nulos  $\alpha_1, \alpha_2, \dots, \alpha_n$ ; então  $\beta$  é exatamente a combinação linear:

$$B = \sum_{K=1}^m \frac{(\beta, \alpha_K)}{\|\alpha_K\|^2} \alpha_K.$$

Dada  $X_{p \times n}$ ,  $\vec{X}^i \equiv N(\vec{\mu}, \Sigma)$ ,  $i=1, \dots, n$  e  $V(X') = I_n \otimes \Sigma$ , a matriz covariante amostral  $S$  e o vetor de médias amostrais  $\bar{X}$  são duas estimativas centradas e independentes, dos parâmetros  $\Sigma$  e  $\mu$  respectivamente.

As distribuições dessas estimativas são

$$\bar{X} \equiv N(\vec{\mu}, \Sigma/n) \quad \text{e} \quad S \equiv W(n-1, \Sigma/n-1)$$

De fato:

Seja  $E$  um espaço euclidiano real  $n$ -dimensional  $\ell\{v_1 \dots v_n\}$  uma base qualquer ortonormal de  $E$ . Vamos decompor  $E$  em soma direta ortogonal de dois subespaços, ou seja:

$$E = E_1 \oplus E_2$$

onde,

$$E_1 = \{v_0\}, \quad v_0 = \sum_{i=1}^n v_i$$

Então:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_1 = \frac{1}{\|v_0\|^2} \begin{pmatrix} (v_1, v_0) \\ \vdots \\ (v_n, v_0) \end{pmatrix} v_0$$

e como  $(v_1, v_0) = 1$

e  $\|v_0\|^2 = (v_0, v_0) = n$

temos:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_1 = \frac{1}{n} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} v_0, \quad \text{Mas } v_0 = (1 \dots 1) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$\therefore \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_1 = \frac{1}{n} J \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad (J)_{i,J} = 1; \quad i, J=1, \dots, n$$

tomando:  $A_1 = \frac{1}{n} J$ , teremos:  $A_2 = (I - \frac{1}{n} J)$  e  $X A_1$  independe de  $X A_2$  e portanto de

$$(X A_2)' (X A_2') = X A_2 X'$$

Mas

$$X A_1 = (\bar{X}, \dots, \bar{X})_{p \times n}$$

e

$$X A_2 X' = T$$

onde

$$T_{iJ} = \sum_{i,J=1}^n (x_i - \bar{x}_i) (x_J - \bar{x}_J)$$

$$\text{Seja } K = \frac{1}{n} J_1, \quad J_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}_{n \times 1}, \quad \text{então:}$$

$$X A_1 K = \frac{\vec{x}}{n}, \quad \text{independente de } \frac{1}{n-1} T = S \quad \text{e}$$

$$T \equiv W(n-1, \Sigma, \mu(I - \frac{1}{n} J) \mu')$$

$$\therefore S \equiv W(n-1, \Sigma/n-1, \mu(I - \frac{1}{n} J) \mu' / n-1)$$

pois

$$S = \frac{1}{\sqrt{n-1}} T \frac{1}{\sqrt{n-1}}$$

e

$$\frac{\vec{x}}{n} = X \frac{1}{n} J_1$$

$$\therefore \frac{\vec{x}}{n} \equiv N(\vec{\mu}, \Sigma/n)$$

pois

$$E(\frac{\vec{x}}{n}) = \vec{\mu} \frac{1}{n} J = \vec{\mu}$$

e como

$$\frac{\vec{x}}{n}_i = (x_{i1} \dots x_{in}) \frac{1}{n} J_1$$

$$\begin{aligned} \text{cov}(\bar{X}_i, \bar{X}_J) &= \frac{1}{n} J_1' \bar{X}_1' \text{cov} \bar{X}_J J_1 \frac{1}{n} = \\ &= \frac{1}{n} J_1' \Sigma_{1J} I J_1 \frac{1}{n} = \frac{1}{n^2} \Sigma_{1J} J_1' I J_1 = \frac{\Sigma_{1J}}{n} \end{aligned}$$

## 5 - INTERPRETAÇÃO GEOMÉTRICA DO MODELO LINEAR GERAL

Prolongamos aqui do ponto de vista geométrico o tratamento geral do Modelo Linear encontrado nos textos de Searle e Anderson.

Como é sabido, por escolha conveniente das matrizes que figuram no modelo podemos descrever as mais variadas relações de dependência entre as variáveis envolvidas. Tanto teórica como praticamente o tratamento geral ou específico deste tipo de modelo tem ganho grande atenção, sendo para nós impossível sequer enumerar a variada bibliografia que trata o assunto.

Sem mais justificativas propomo-nos nesta parte a enunciar os resultados já conhecidos de estimação envolvidos no modelo e interpretá-los geometricamente de forma a reafirmar fatos já conhecidos sob outro ponto de vista. Nesta linha analisaremos também, do ponto de vista geométrico, a hipótese linear geral. Não nos vamos ater a tipos específicos de hipótese, e nem tratar as estatísticas usadas para testá-las. Partiremos tão somente do arrazoado de que, feita uma hipótese, em geral as possíveis estatísticas se baseiam na decomposição da estimativa da matriz covariante sob a hipótese; e ainda em geral uma das componentes dessa decomposição é a estimativa da matriz covariante sem a suposição da veracidade da hipótese.

Por esta razão, a priori, para que possamos testar uma hipótese, estas duas estimativas não poderão coincidir; assim, através da interpretação geométrica, obteremos a condição necessária e suficiente para que coincidam, limitando assim inferiormente o conjunto das hipóteses testáveis segundo esse tipo de procedimento. (decomposição da estimativa da matriz covariante sob a hipótese).

#### Interpretação dos Estimadores no Modelo Linear Geral.

Seja:

$$X_{pxn} = B_{pxq} Z_{qxn} + \xi_{pxn}$$

onde:

$B_{pxq}$  é a matriz de parâmetros (desconhecida).

$Z_{qxn}$  é uma matriz conhecida

$$\vec{\xi}^i \equiv N(\vec{0}, \Sigma_{pxp}), \quad V(\xi^i) = I_n \otimes \Sigma \quad n \geq p + r(Z)$$

Pelo método da máxima verossimilhança, os estimadores da matriz de parâmetros e da matriz covariante, que denotaremos respectivamente por  $\hat{B}_G$  e  $n \hat{\Sigma}$ , vêm dados por:

$$\hat{B}_G = X Z' G$$

$n \hat{\Sigma} = X(I - Z' G Z) X'$ , sendo  $n \hat{\Sigma}$  invariante quanto a possíveis escolhas de  $G$  e  $(Z Z')^G$ .

Passamos a um lema que nos levará às interpretações geométricas das quantidades de interesse.

LEMA 5.1 - Dada uma matriz qualquer  $Z_{q \times n}$ , consideremos  $F$  o seguinte subespaço de  $E$

$$F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}, \text{ e } F^\perp \text{ o subespaço de } E \text{ ortogonal a } F,$$

sendo  $\{v_1, \dots, v_n\}$  uma base ortonormal de  $E$ .

Então as matrizes da projeção ortogonal da base  $\{v_1, \dots, v_n\}$  sobre  $F$  e  $F^\perp$  são respectivamente  $Z' G Z$  e  $I - Z' G Z$ , onde  $G \in (ZZ')^g$

PROVA - Como  $Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  é um sistema de geradores de  $F$  temos:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = \lambda_{n \times q} Z_{q \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e como a projeção é ortogonal:

$$\left[ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F \right] \cdot \left[ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right]' = 0$$

ou seja:

$$\left[ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - \lambda Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right] \cdot [(v_1 \dots v_n) Z'] = 0$$



logo

$$Z' - \lambda Z Z' = 0 ,$$

mas

$$Z' = Z' G Z Z' , \quad G \in (Z Z')^G$$

e

$$Z' G Z Z' - \lambda Z Z' = 0 \Rightarrow \lambda Z = Z' G Z$$

(Searle pág 16, Lema 3).

Portanto:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = \lambda_{n \times q} Z_{q \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = Z' G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

se e só se:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = (I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

LEMA 5.2 - Se dada uma decomposição de E,  $E = F \oplus F^\perp$  e as correspondentes projeções da base forem dadas por:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} e ,$$

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} = (I - A) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e ainda,  $A = Z' G Z$ , então  $F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$

PROVA - De fato  $A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  é um sistema de geradores de  $F$  pois:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Seja  $\{\alpha_1, \dots, \alpha_s\}$  uma base de  $F$ ,  
então:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} = \lambda_{s \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix}_F = \lambda_{s \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F$$

logo

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{pmatrix} = \lambda_{s \times n} A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

logo

$$F = \left\{ Z' G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\} \subset \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

e como  $r(Z' G Z) = r(Z)$ , pois:

$$G \in (Z Z')^G \text{ e } r(Z' (Z Z')^G Z) \leq r(Z),$$

e ainda:

$$Z = Z(Z'(Z Z')^G Z) \text{ temos}$$

$$r(Z) \leq r(Z'(Z Z')^G Z)$$

segue-se que:

$$F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

cqd.

Notemos que, dada uma representação geométrica dos dados num espaço euclidiano real E, ou seja:

$$X_{p \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \text{ p pontos de E, teremos:}$$

$$\text{Sendo } F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\} \text{ então}$$

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = Z' G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \text{ e portanto:}$$

$$\begin{aligned} \left( X_{p \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_F &= X_{p \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = \\ &= X Z' G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \end{aligned}$$

Mas  $\hat{B}_G = X Z' G$ , logo:

$$X_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = \hat{B}_G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Podemos por simples aplicação do lema 5.1 visualizar geometricamente os seguintes já conhecidos resultados:

Ao estimarmos, neste caso pelo método da máxima verossimilhança, os parâmetros do modelo, estamos na realidade projetando ortogonalmente as  $p$  variáveis (vistas como  $p$  vetores de  $E$ , resultantes de combinações lineares da base ortogonal  $\{V_1, \dots, V_n\}$ , segundo coeficientes dados pelas linhas de  $X_{pxn}$ ) sobre o subespaço gerado pelas combinações da base segundo as linhas de  $Z$ . E claramente  $\hat{B}_G$  são as coordenadas dessa projeção.

— Sabemos que, no caso do posto de  $Z$  ser completo,  $\hat{B}_G$  é única e de fato neste caso  $Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  é uma base de  $F$ , e portanto  $\hat{B}_G$

são coordenadas em relação à base.

— A não unicidade de  $\hat{B}_G$ , caso o posto de  $Z$  não seja completo, fica explicada, pois, estaríamos calculando coeficientes de projeções ortogonais relativamente a um sistema de geradores e não a uma base.

— A invariabilidade de  $\hat{B}_G Z$  é coerente pois são coordenadas da projeção relativamente à base.

Passamos agora a interpretar a estimativa da matriz covariante, ou seja:

$$n \hat{\Sigma} = X(I - Z' G Z)X'$$

Mas:

$$X_{p \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} = X(I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e o produto dessas projeções ortogonais nos dá:

$$\begin{aligned} & \left\{ X(I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\} \cdot \left\{ X(I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}' = \\ & = X(I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (v_1 \dots v_n) (I - Z' G Z) X' = \\ & = X(I - Z' G Z) X' . \end{aligned}$$

Portanto:

— o estimador da matriz covariante vem dado pelos produtos escalares das componentes ortogonais a  $F$ .

— nas diagonais de  $n \hat{\Sigma}$  estão as normas ao quadrado das componentes ortogonais a  $F$  das projeções de cada variável (vista como um ponto de  $E$ ).

— a terminologia "parte explicada do modelo" e sua negação ficam relacionadas respectivamente com "componente da projeção das variáveis sobre  $F$ " e sua componente ortogonal.

— associando esta decomposição à seguinte decomposição da matriz identidade:

$$I = Z' G Z + (I - Z' G Z)$$

e levando em conta as hipóteses do modelo por aplicação direta da interpretação geométrica temos:

$$n \hat{\Sigma} \equiv W(n - r(Z), \Sigma, B Z (I - Z' G Z) Z' B').$$

E como

$$Z (I - Z' G Z) = 0$$

segue-se:

$$n \hat{\Sigma} \equiv W(n - r(Z), \Sigma).$$

O teste de Hipótese no Modelo Linear Geral e sua Interpretação Geométrica.

Seja H a hipótese linear geral:

$$H: B_{pxq} K_{gx\ell} = M_{px\ell}$$

tal que:

$$r(K) = t \leq \min(g, \ell) .$$

Passamos a um lema que nos fornece a solução geral do sistema acima, para que possamos analisar o modelo sob a hipótese. (Daqui para frente sob a hipótese significará sob a suposição da veracidade da hipótese).

LEMA 5.3 - Se o sistema  $B K = M$  é compatível, a solução geral do mesmo vem dada por:

$$B_{pxq} = M_{pxl} D_{lxq} + \lambda_{pxq-t} U_{q-txq}$$

onde:

—  $D \in K^g$

— As linhas de  $U_{q-txq}$  são uma base do subespaço de  $R^q$  ortogonal ao subespaço gerado pelas colunas de  $K_{qx l}$ . Ou seja:

$$U K = 0 \quad \text{e} \quad r(U) = q - t = q - r(K)$$

—  $\lambda_{pxq-t}$  é uma matriz de parâmetros arbitrários.

PROVA - Da compatibilidade segue-se a existência de  $B_1 \mid B_1 K = M$ .

Então:

$$(M D + \lambda U) K = M D K + \lambda U K = B_1 K D K = B_1 K = M$$

Reciprocamente, se  $B_2$  é uma solução de  $B K = M$ , ou seja,  $B_2 K = M$  então vamos provar que:

$$\exists \lambda_2 \mid B_2 = M D + \lambda_2 U.$$

Mas,

$$B_2 = M D + \lambda_2 U \iff (B_2 - M D) = \lambda_2 U \iff$$

$$(B_2 - M D) = 0$$

Verificando a última equivalência:

$$(B_2 - M D) = \lambda_2 U \iff (B_2 - M D)K = 0$$

$$(\implies) (B_2 - M D)K = \lambda_2 U K = 0$$

$$(\impliedby) (B_2 - M D)K = 0 \implies \exists \lambda_2 \mid (B_2 - M D) = \lambda_2 U$$

De fato:

$$(B_2 - M D)K = B_2 K - M D K = B_2 K - B_2 K D K = 0$$

pois

$$K D K = K.$$

c.q.d.

Pelo Lema 5.3 temos pois que o modelo linear sob a hipótese H pode se escrever como:

$$Y = \lambda (U Z) + \hat{\epsilon}, \text{ onde:}$$

- $-Y = X - M D Z$
- $-D \in K^g$
- $U_{q-txq} / UK = 0$  e  $r(U) = q - r(K)$ .
- $\lambda_{pxq-t}$  é uma matriz de parâmetros arbitrários.

e portanto:

$$n \hat{\Sigma}_H = Y (I - (UZ)' [(UZ) (UZ)']^g (UZ)) Y'$$



por analogia à estimativa do modelo linear geral, pois:

— as colunas de  $y$  são multinormais independentes por serem simples translações das colunas de  $X$ .

— a matriz covariante das colunas de  $Y$  é  $\Sigma$ , posto que variâncias e covariâncias são invariantes quanto a translações.

—  $E(Y) = \lambda(UZ)$ , sendo  $(UZ)$  a nova matriz de delinea-  
mento.

Observemos que:

$$Z = Z Z' G Z, \text{ propriedade de inversa generalizada} \\ \text{(Searle)}$$

logo para qualquer matriz  $S_{p \times q}$  temos:

$$n \hat{\Sigma} = X(I - Z'GZ)X' = (X - SZ) (I - Z'GZ) (X - SZ)'$$

logo tomando  $S = M D$  temos:

$$n \hat{\Sigma} = (X - MDZ) (I - Z'GZ) (X - MDZ)'$$

TEOREMA 5.4 - A condição necessária e suficiente para que as estimativas da matriz covariante sob a hipótese linear geral e sob o modelo linear geral sejam as mesmas é que os subespaços  $F$  e  $F_H$  (explicitados abaixo) coincidam.

$$F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}, F_H = \left\{ U Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

PROVA - Obtivemos anteriormente as seguintes expressões para  $n \hat{\Sigma}$  e  $n \hat{\Sigma}_H$ :

$$n \hat{\Sigma} = Y(I - Z' G Z)Y'$$

e

$$n \hat{\Sigma}_H = Y(I - (UZ)' [(UZ)(UZ)']^G (UZ))Y'$$

onde  $Y = X - M D Z$ , estando todas as matrizes definidas anteriormente.

Logo:

$$n \hat{\Sigma} = n \hat{\Sigma}_H \text{ se e s\o se:}$$

$$I - (UZ) [(UZ)(UZ)']^G (UZ) = I - Z' G Z$$

Mas pelo lema 5.1, a matriz do lado direito da igualdade é a matriz da projeção ortogonal da base  $\{v_1 \dots v_n\}$  sobre o subespaço ortogonal a  $F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$ ; e analogamente do lado esquerdo temos o

ortogonal a  $F_H = \left\{ UZ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$ , ou seja:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} = (I - Z' G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

e

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} = (I - (UZ)' [(UZ)(UZ)']^g (UZ)) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

$$F_H = \left\{ (UZ) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

e portanto sendo:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_F = Z' G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} = (UZ)' [(UZ)(UZ)']^g (UZ) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

, sistemas de geradores de  $F$  e  $F_H$ , temos demonstrado o teorema.

Acabamos de provar que a condição necessária e suficiente para que  $n \hat{\Sigma} = n \hat{\Sigma}_H$  é que os subespaços  $F$  e  $F_H$ , definidos no teorema anterior, coincidam. Uma rápida olhada nos sistemas de geradores desses espaços mostra que:

sendo,

$$F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\} \quad \text{e} \quad F_H = \left\{ UZ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

a condição necessária e suficiente para que  $F$  e  $F_H$  não coincidam é que:

$$r(UZ) \neq r(Z),$$

pois neste caso, e só neste caso, não coincidem os sistemas de geradores de  $F$  e  $F_H$ .

Passamos a investigar, do ponto de vista geométrico, a decomposição de  $n \hat{E}$  quando os subespaços  $F$  e  $F_H$  não coincidem:

$$F_H = UZ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \subsetneq F = \left\{ Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\}$$

se e só se:

$$F_H^\perp \supsetneq F^\perp, \text{ ou seja:}$$

$$\exists S \neq \{\vec{0}\}, S \perp F_H^\perp \mid F_H^\perp = F^\perp \oplus S$$

Pela decomposição acima de  $F_H^\perp$ , temos a seguinte expressão para a projeção ortogonal da base sobre  $F_H$ :

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} = \left[ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} \right]_{F^\perp} + \left[ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} \right]_S$$

Mas  $F^\perp$  e  $S$  são subespaços de  $F_H^\perp$ , portanto:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S$$

ainda:

$$n \hat{\Sigma}_H = Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} \cdot (v_1 \dots v_n)_{F_H}^{\perp} Y' ,$$

e pela identidade acima temos:

$$n \hat{\Sigma}_H = Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} \cdot (v_1 \dots v_n)_{F_H}^{\perp} Y' +$$

$$Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S \cdot (v_1 \dots v_n)_S Y'$$

$$= n \hat{\Sigma} + Q .$$

(Pela expressão encontrada para  $n \hat{\Sigma}$  antes um pouco do teorema 5.4).

Podemos aplicar a interpretação geométrica do teorema de James para obter as distribuições e relações de dependência de  $n \hat{\Sigma}_H$ ,  $n \hat{\Sigma}$  e  $Q$ .

De fato:

$$I = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot (v_1 \dots v_n) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} \cdot (v_1 \dots v_n)_{F_H} +$$

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H}^{\perp} \cdot (v_1 \dots v_n)_{F_H}^{\perp}$$

e por sua vez:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S$$

logo:

$$I = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} \cdot (v_1 \dots v_n)_{F_H} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} \cdot (v_1 \dots v_n)_{F^\perp} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S \cdot (v_1 \dots v_n)_S$$

$$Y Y' = Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H} \cdot (v_1 \dots v_n)_{F_H} Y' + Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} \cdot (v_1 \dots v_n)_{F^\perp} Y' +$$

$$Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S \cdot (v_1 \dots v_n)_S Y'$$

tínhamos que:

$$n \hat{\Sigma} \equiv W(n - r(Z), \Sigma), \quad \text{logo:}$$

$$Q \equiv W(r(Z) - r(UZ), \Sigma, E(Y) A_S [E(Y)]')$$

onde

$$A_S = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S \cdot (v_1 \dots v_n)_S \quad e$$

$n \hat{\Sigma}$  e  $Q$  são independentes.

Pela decomposição da identidade em produtos de projeções ortogonais sobre  $F_H$  e  $F_H^\perp$  e aplicação de raciocínio análogo ao precedente temos:

$$n \hat{\Sigma}_H \equiv W(n - r(UZ), \Sigma, E(Y) A_H^\perp (E(Y))')$$

onde

$$A_H^\perp = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} \cdot (v_1 \dots v_n)_{F_H}^\perp$$

Sob a hipótese teremos:  $E(Y) = \lambda UZ$  e portanto  $E(Y) A_H^\perp = 0$ , e por propriedade de distribuições de Wishart o parâmetro de descentralização de  $Q$  será zero. Resumindo, sob a hipótese:

$$- n \hat{\Sigma} \equiv W(n - r(Z), \Sigma)$$

$$- Q \equiv W(r(Z) - r(UZ), \Sigma)$$

-  $n \hat{\Sigma}_H \equiv W(n - r(UZ), \Sigma)$ , onde sempre  $n \hat{\Sigma}$  e  $Q$  são independentes (sob ou não a hipótese).

Comentário Final:

Searle ao tratar hipóteses não testáveis, considerando que não há teste possível de uma hipótese quando as estimativas da matriz covariante, sob a hipótese e na ausência dela, coincidem; é conclusivo quanto à necessidade da estimabilidade. Ou seja, a não estimabilidade implicaria a não testabilidade da hipótese.

A condição necessária e suficiente desenvolvida no teorema 5.4 amplia, neste contexto, o teto das hipóteses testáveis uma vez que podemos exibir pelo menos uma hipótese que não é estimável e é testável.

Neste sentido, acreditamos estar realmente definido o "limite inferior" do conjunto das hipóteses testáveis, tratando-se agora de estudarmos que hipóteses poder-se-ia classificar entre esse limite e as estimáveis, e com que critérios, de forma a merecerem a atenção especial como as estimáveis.

Passemos pois ao exemplo acima referido.

Consideremos, no modelo linear geral definido nesta parte,  $r(Z) < q$  e seja a hipótese linear H:

$$H: B = B_0$$

H não é estimável pois, se o fosse, deveríamos ter:

$$\exists L_{n \times q} / K_{q \times q} = Z'_{q \times n} L_{n \times q}$$

como

$$K = I_{q \times q} \text{ e } r(K) = q$$

$$\text{se } K = Z' L \Rightarrow r(K) \leq r(Z) < q, \text{ absurdo.}$$

Por outro lado, pelo teorema 5.4 temos:

$$F = \left\{ Z \begin{pmatrix} v \\ \vdots \\ v_n \end{pmatrix} \right\}, \quad F_H = \left\{ UZ \begin{pmatrix} v \\ \vdots \\ v_n \end{pmatrix} \right\}$$

, onde U está definida no lema 5.3 e é tal que:



$$U_{q-txq} \mid UK = 0 \text{ e } r(U) = q - r(K)$$

Mas  $K = I \therefore U = \vec{0}$  e  $r(U) = 0$  implicando:

$$F_H = \{\vec{0}\}$$

Portanto estamos diante de uma hipótese testável pois  $F$  e  $F_H$  não coincidem.

Passamos a obter a decomposição de  $n \hat{\Sigma}_H$  de forma a expressar a componente  $Q$  e verificar a não nulidade da mesma.

Após o teorema 5.4 a seguinte decomposição de  $n \hat{\Sigma}_H$  foi obtida,  $F^\perp$  e  $S$  definidos anteriormente:

$$n \hat{\Sigma}_H = Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} \cdot (v_1 \dots v_n)_{F_H^\perp} Y' =$$

$$= Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} \cdot (v_1 \dots v_n)_{F^\perp} Y' +$$

$$+ Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_S \cdot (v_1 \dots v_n)_S Y'$$

Como  $F_H = \{\vec{0}\}$ ,  $F_H^\perp = E$  logo:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F_H^\perp} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e temos:

$$n \hat{\Sigma}_H = YY' = (X - B_O Z) (X - B_O Z)'$$

pois:

$$Y = X - M D Z, \text{ e neste caso:}$$

$$M = B_O$$

$$D \in K^g \therefore D = I$$

e,

$$Y = X - B_O Z.$$

Como:

$$\begin{aligned} n \hat{\Sigma} &= Y \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{F^\perp} (v_1 \dots v_n)_{F^\perp}' Y' = \\ &= X(I - Z' G Z) X' = (X - \hat{B}_G Z) (X - \hat{B}_G Z)' \end{aligned}$$

Temos pela decomposição de  $n \hat{\Sigma}_H$  :

$$\begin{aligned} Q &= n \hat{\Sigma}_H - n \hat{\Sigma} = (X - B_O Z) (X - B_O Z)' - \\ & (X - \hat{B}_G Z) (X - \hat{B}_G Z)' \end{aligned}$$

Podemos ainda obter Q de forma mais direta, ou seja:

$$S \perp F^\perp \mid F_H^\perp = S \oplus F^\perp$$

Vimos que:

$$\left( X_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_F = \hat{B}_G Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e como:

$$\left( X_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F^\perp} = \left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F^\perp}, \text{ pois}$$

$$\left( B_O Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F^\perp} = 0$$

e

$$\left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F_H^\perp} = \left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_E = (X - B_O Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

e, como tomando-se o novo modelo tivemos:

$$\left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F_H^\perp} = (X - B_O Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

então:

$$\left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_S = \left( Y_{pxn} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F_H^\perp}$$

$$\begin{aligned}
 & \left( Y_{p \times n} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)_{F^\perp} = \\
 & = (X - B_O Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - (X - \hat{B}_G Z) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \\
 & = (\hat{B}_G - B_O) Z \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}
 \end{aligned}$$

logo:

$$Q = (\hat{B}_G - B_O) Z [(\hat{B}_G - B_O) Z]'$$

e temos:

$$n \hat{\Sigma}_H = (X - B_O Z) (X - B_O Z)'$$

$$n \hat{\Sigma}_H = (X - \hat{B}_G Z) (X - \hat{B}_G Z)' + (\hat{B}_G - B_O) Z [(\hat{B}_G - B_O) Z]'$$

$$= n \hat{\Sigma} + Q$$

e pelas conclusões já obtidas, por simples aplicação do teorema da interpretação geométrica, temos que sob H:

$$n \hat{\Sigma}_H \equiv W(n, \Sigma)$$

$$Q \equiv W(r(Z), \Sigma)$$

e, independentemente da verificação de H:

$$n \hat{\Sigma} \equiv W(n - r(Z), \Sigma)$$

sendo Q e n  $\hat{\Sigma}$  independentes.

É usual o procedimento seguinte:

$$H: B = B_0$$

então podemos decompor  $X - B_0 Z$  como:

$$X - B_0 Z = X - \hat{B}_G Z + (\hat{B}_G - B_0) Z$$

e então:

$$n \hat{\Sigma}_H = (X - \hat{B}_G Z)(X - \hat{B}_G Z)' + (\hat{B}_G - B_0) Z [(B_G - B_0) Z]'$$

ou seja, em apenas dois passos poderíamos ter chegado ao resultado desejado apenas dependendo da argumentação intuitiva disponível; o que comumente é usado nos casos particulares para obter as famosas e necessárias decomposições de soma de quadrados.

xxx<sup>x</sup>xxx  
x

## REFERÊNCIAS

- Anderson, T.W (1958). An Introduction to Multivariate Statistical Analysis. John Wiley and Sons, Inc. New York.
- Cochran, W.G. (1934). The distribution of quadratic forms in a normal system, with applications to the analysis of covariance. Proc. Camb. phil. Soc. 30, 178-191.
- Dempster, A.P. (1969). Elements of Continuous Multivariate Analysis. Addison - Wesley.
- Dugué, D. et Girault, M. (1969). Analyse de Variance et Plans D'Expérience. Dunod Paris.
- Ferrandis, Eduardo (1977). O Contraste de Hipóteses nos Modelos Lineares Multivariados. Relatório Interno nº 57.
- Hoffmann, K. and Kunze, R. (1961). Álgebra Linear. Editora Polígono. (USP).
- James, G.S. (1952). Note on a theorem of Cochran, Proc. Camb. phil. Soc. 48 - p.443.
- Kshirsagar, A.M. (1972). Multivariate Analysis. Maral Dekker Inc.
- Rao, C.R. (1965). Linear Statistical Inference and its Applications. John Wiley and Sons, Inc. New York.
- Scheffé, H. (1959). The Analysis of Variance. John Wiley and Sons, Inc.

- Searle, S.R. (1971). Linear Models. John Wiley and Sons, Inc. New York.
- Srivastava, T.N. (1975). A Survey of Statistical Design and Linear Models. North Holland Publishing Company.
- Wilks, Samuel S. (1962). Mathematical Statistics Wiley, New York.