COMPONENTES PRINCIPAIS ESTOCASTICAMENTE RESTRITOS.*

Hugo Pedro Boff**

Resumo

O artigo insere a análise convencional dos Componentes Principais Restritos (ACP-R) no contexto aleatório do Modelo Linear Geral (ACP-RA). Na seção 3, matrizes de ACP-RA teóricas e empíricas são obtidas. A estimação destas últimas é discutida no modelo auto-regressivo pelos métodos Máxima Verossimilhança e Aitken Duplo Estágio. Subespaços de referência múltiplos para as componentes são considerados na seção 4. Inferências e outras inserções teóricas da análise são enfim sugeridas.

Abstract

The main scope of this paper is to extend the Restricted Principal Component Analysis (ACP-R) firstly proposed by D'Ambra and Lauro (1982) and Amato (1988) to a stochastic framework. After presenting the ACP-R (Section 2), we show how the stochastic restrictions can be inserted into the classical analysis of the General Linear Model. Theoretical and empirical matrices for stochastically restricted principal components (ACP-RA) are derived for both cases of known and unknown matrix-norm for directed projections of analytical variables on the restriction's subspace. Section 3 present applications of ACP-RA matrices for a particular structure of residual variability, namely, the autoregressive process. In Section 4 the ACP-RA approach is extended to more general structures of residual variability applied in a context of multiple restrictions' subspaces.

1. Introdução.

Na Análise em *Componentes Principais Restritos* (ACP-R), inicialmente proposta por D'Ambra e Lauro (1982), busca-se isolar os fatores componentes principais a partir da matriz de Variâncias e

*O autor agradece o apoio à pesquisa do Prof. Ademir Clemente e o incentivo dos Profs. Mario Diney Correa Bittencourt e Cassio Frederico Camargo Rolim, todos do Depto. de Economia da UFPR, assim que do Prof. Anselmo Chaves Netto do Depto. de Estatística da UFPR. Os erros porventura existentes são de responsabilidade exclusiva do autor.

**Professor do Depto. de Economia e do Curso de Mestrado em Desenvolvimento Econômico da Universidade Federal do Paraná (UFPR).

Covariâncias empíricas, V_x , da melhor imagem de projeção das k variáveis analíticas $x' = (x_1, x_2, \dots, x_k)$ sobre o subespaço R_z gerado por l fatores condicionantes $z' = (z_1, z_2, \dots, z_l)$, no sentido dos *mínimos quadrados*: $X^+ = P_z X$, sendo X, Z as matrizes amostrais (n, k) e (n, l) previamente padronizadas¹ e $P_z = Z(Z'Z)^{-1}Z'$ o projetor ortogonal sobre $R_z \subseteq R^n$. Posteriormente, Amato (1988) enfatiza os aspectos assimétricos da rede de vínculos estabelecida entre x e z e mostra como a ACP-R pode integrar-se no contexto teórico das *Correlações Parciais*, mediante a prévia Padronização Múltipla de X e Z com relação à p fatores exógenos $w' = (w_1, w_2, \dots, w_p)$, $p \geq 1$. As *Variáveis Padronizadas Múltiplas* (VPM), X_w e Z_w , são então consideradas no lugar das variáveis e fatores originais.² A *matriz* de ACP-R vem definida por $V_x = X^{+'}X^+ = X'P_zX$, (ou $V_x = X'_wP_{zw}X_w$) e toda a análise desenvolve-se em um contexto determinístico. Tanto na definição de V_x quanto na construção das VPM, os operadores de Var.Cov. envolvidos definem-se como momentos empíricos das variáveis e não a partir de suas distribuições populacionais.

Neste artigo, é nosso propósito ampliar a metodologia inicialmente proposta, estendendo o quadro analítico da ACP-R ao contexto aleatório, mediante sua integração nas estruturas relacionais estocásticas, convencionalmente hipotetizadas no *Modelo Linear Geral* (MLG) e extensivamente estudadas em Econometria. Esta abordagem, denominada ACP-RA, permitirá então, como veremos na seção 3, modular a *matriz* de ACP-R, V_x dos desvios estimados entre os vínculos relacionais aparentes e as estruturas de dependência formal hipotetizadas. Na interpretação geométrica das *matrizes* de ACP-RA obtidas, as hipóteses formuladas com relação às vinculações seriais entre as variáveis analíticas desempenharão um papel central. Evidenciaremos também como a matriz de ACP-R poderá ser derivada daquela de ACP-RA mediante a simples supressão da hipótese aleatória.

¹Padronização Simples: desvios da média dividido por $n^{1/2}$ desvios padrões para cada coluna de X e Z .

²Veja-se Amato (1976, 1988) e nosso estudo "Variáveis Padronizadas Múltiplas — Construção e Aplicações", Anais do XIV Encontro Brasileiro de Econometria, v. 1, p. 107-125, Campos do Jordão, 1992.

2. Restrições determinísticas (ACP-R).

Consideremos o espaço vetorial (real) de dimensão n , $\mathbf{R}^n$ e dois subespaços de $\mathbf{R}^n$ engendrados pelas k colunas de $X(\mathbf{R}_x)$ e pelas l colunas de $Z(\mathbf{R}_z)$, ambas previamente padronizadas. D'Ambra e Lauro (1982) definem a matriz de ACP-R pela matriz das correlações empíricas das *imagens de projeção de X sobre R_z* ($X^+ = P_z X$):

$$V_x = X^{+'} X^+ = X' P_z X \quad (1)$$

$$P_z = Z(Z'Z)^{-1}Z' \quad (2)$$

Em termos relacionais, a imagem $X^+ = Z(Z'Z)^{-1}Z'X$ pode ser escrita como: $X^+ = Z\hat{G}$, sendo $\hat{G} = (Z'Z)^{-1}Z'X$ a estimativa de *Mínimos Quadrados Ordinários* (MQO) da matriz paramétrica G , de dimensão (l, k) no modelo linear.

$$X = ZG + E, \quad (3)$$

onde E é uma matriz (n, k) de erros não-estocásticos.³ As matrizes X e Z sendo padronizadas, V_x poderá ser ainda expressa como: $V_x = V_{xz} V_{zz}^{-1} V_{zx}$, onde $V_{xz} = X'Z(V_{zx} = V'_{xz})$ e $V_{zz} = Z'Z$ representam as matrizes empíricas de Correlação dos valores amostrais.

Para a obtenção dos Componentes Principais Restritos, a *equação fundamental de diagonalização* a ser considerada é

$$(X'P_zX)v_i = \lambda_i v_i \quad i = 1, 2, \dots, k \quad (4)$$

ou, em termos matriciais,

$$V_x N' = N' \Lambda \quad (5)$$

com $NN' = N'N = I_k$ e $\Lambda = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ matrizes dos auto-vetores (v_i) e auto-valores (λ_i) de V_x , respectivamente.⁴ Os

³ $\hat{G}$ pode ser obtida mediante a execução de $\min_G \text{Tr}(X - ZG)'(X - ZG)$.

⁴ V_x sendo matriz real, simétrica e positiva definida, todos os seus autovalores λ_i serão estritamente positivos ($\lambda_i > 0$).

componentes principais (normalizados) definir-se-ão então por: $c_i = \lambda_i^{-1/2} X^+ v_i$ ou, em termos matriciais,

$$C = \hat{X} N' \Lambda^{-1/2} = \hat{X} F' \quad (6)$$

sendo $F' = N' \Lambda^{-1/2}$ a matriz dos *fatores principais* e com $C' C = I_k$.

As matrizes das correlações (não-simétricas) entre os *componentes principais restritos* (c_i) e as variáveis analíticas x' e entre c_i e os fatores determinantes z' , serão dadas, respectivamente, por (usando-se (7) e (9)): $X' C = N' \Lambda^{1/2} = F' \Lambda$ e

$$Z' C = Z' X F' \quad (7)$$

Como bem enfatizam D'Ambra e Lauro (1982, p.54), sendo $Z' X = V_{zx}$ matriz das correlações dos fatores determinantes com as variáveis originais, a expressão (7) evidencia a reconstituição da imagem destas correlações efetuada pela ACP-R, diretamente sobre os eixos principais. Contribuições individuais e indicadores de representação da variabilidade total poderão ser convencionalmente derivados a partir de (4)-(6), considerando-se os componentes restritos não-normalizados $C_i = \lambda^{1/2} c_i$. Também, outras propriedades dos componentes poderão ser derivadas (ortogonalidade, invariância às transformações regulares de Z em R_z etc.), sobre as quais nos poderemos remeter ao texto dos autores.

3. Restrições aleatórias (ACP-RA).

Uma das maiores limitações da ACP-R deriva do caráter determinístico das restrições formalizadas em (3), o qual não permite levar em consideração as imprecisões porventura subsistentes quanto à verdadeira natureza dos vínculos pressupostos entre X e Z , nem tampouco ponderar a qualidade das imagens obtidas pelos termos aleatórios representativos destas imprecisões.

3.1. Fundamentos geométricos e estatísticos.

Consideremos os subespaços vetoriais R_x e R_z definidos na seção anterior e tomemos dois subespaços de R_x , R_+ e R_E , de maneira que $R_x = R_+ \oplus R_E$, isto é, que $\forall x_i \in R_x$ exista uma única decomposição

$x_i^+ \in \mathbf{R}_+$, $\varepsilon_i \in \mathbf{R}_E$ tal que $x_i = x_i^+ + \varepsilon_i$. Assim, para $i = 1, 2, \dots, k$ temos, em forma matricial:

$$X = X^+ + E \quad (8)$$

Definamos agora sobre $\mathbf{R}^n$ o *produto interno*:

$$(b * a) = a' \Sigma^{-1} b \quad (9)$$

com $a, b \in \mathbf{R}^n$ e Σ^{-1} matriz real, positiva definida, de ordem n . Por outro lado, um *projetor ortogonal* em $\mathbf{R}_z$ define-se como (Rao, 1973, p.47):⁵

$$P_z = Z(Z' \Sigma^{-1} Z)^{-1} Z' \Sigma^{-1} \quad (10)$$

A utilização do projetor P_z dado em (10) para a definição de X^+ definida em (8), ou seja, $X^+ = P_z X$, conduz de fato à ortogonalização de $\mathbf{R}_x$ segundo o *produto interno* definido em (9), de maneira que $\mathbf{R}_E = \mathbf{R}_+^\perp$ será o complemento ortogonal de $\mathbf{R}_+$ em $\mathbf{R}_x$: $(X^+ * E) = X'(I - P_z)' \Sigma^{-1} P_z X = 0$. Temos também $\mathbf{R}_z \perp \mathbf{R}_E$ ($Z * E = 0$). De um ponto de vista estatístico, X^+ pode ser interpretada como matriz-imagem ($\hat{X}$) dos valores calculados de X em uma estimativa de regressão sistêmica, na hipótese desta variável estar vinculada à Z pelo *Modelo Linear Geral* (MLG) descrito em (3) e onde E é uma matriz aleatória de média nula, $\mathcal{E}(E) = 0$, e matriz Var.Cov.:

$$V(E) = \Omega \otimes \Sigma \quad (11)$$

Aqui, Σ introduzida com (9) será simétrica, $\otimes$ é o produto *Kronecker* e Ω é a matriz das Var.Cov. dos resíduos $\varepsilon' = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k)$ entre as k equações $x_t = G' z_t + \varepsilon_t$ ($t = 1, 2, \dots, n$) implícitas em (3).⁶ Com efeito, com base em (10), podemos escrever:

$$X^+ = P_z X = Z(Z' \Sigma^{-1} Z)^{-1} Z' \Sigma^{-1} X = Z \hat{G} = \hat{X} \quad (12)$$

$$\hat{G} = (Z' \Sigma^{-1} Z)^{-1} Z' \Sigma^{-1} X \quad (13)$$

⁵Sendo um projetor ortogonal, P_z será por construção idempotente e $\Sigma^{-1} P_z$ será uma matriz Hermitiana (simétrica) de ordem n .

⁶Definimos $V(E)$ em (11) formando o operador vetor, ou seja, $\text{vec}(E)$ pelas colunas de E . Alternativamente, se tomarmos $\text{vec}(E)$ pelas linhas, teremos: $V(E) = \Sigma \otimes \Omega$. Pelo seu lado, Ω será por definição uma matriz simétrica e positiva definida.

A matriz $\hat{G}$ é o estimador de *Mínimos Quadrados Generalizados* (MQG) de G no modelo (3) segundo (11), ou ainda o seu estimador de *Máxima Verossimilhança* (MV) sob a hipótese da *Normalidade* de E .

A abordagem da ACP-R pressupõe, de fato, $\Sigma = I_n$. A abordagem ACP-RA, além de tomar E como aleatória, admite eventualmente $\Sigma \neq I_n$.

3.2. Projeções e transformações.

O emprego do projetor P_z dado em (10), assim que o cálculo de $\hat{G}$ em (13) pressupõe o conhecimento da matriz Σ . Quando Σ é conhecida (ou estimada fora de R_x), vemos em (12) que a *imagem de projeção* (X^+) coincide com a imagem da regressão de X sobre $Z(\hat{X})$ no modelo (3). Todavia, quando Σ é *desconhecida*, sua estimativa $\hat{\Sigma}$ não será independente de R_x ⁷ e então as coordenadas direcionais das projeções de X sobre R_z não serão mais definidas por uma norma fixa e única, mas serão variáveis segundo os pontos $x_1, x_2, \dots, x_k$ da projeção. O estimador $\hat{P}_z = Z(Z'\hat{\Sigma}^{-1}Z)^{-1}Z'\hat{\Sigma}^{-1}$ não poderá ser visto como um projetor para X sobre R_z e $\hat{X} = \hat{P}_z X$ deverá ser interpretada como a *imagem de uma transformação* de X sobre R_z .

3.3. Matrizes de ACP-RA.

Definiremos diferentemente as matrizes de ACP-RA segundo que Σ seja conhecida ou desconhecida.

1. Para Σ conhecida (ou estimada fora de R_x)

Construimos a *matriz de ACP-RA* ($\mathcal{V}_x$) pelo valor esperado do produto interno das imagens de projeção por P_z de X sobre R_z :

$$\mathcal{V}_x = \mathcal{E}(X^+ * X^+) \quad (14)$$

A partir de (9) e da ortogonalidade induzida por P_z sobre R_x , temos: $X^+ * X^+ = X^+ \Sigma^{-1} X^+ = X' P_z' \Sigma^{-1} P_z X = X' \Sigma^{-1} P_z X$. Usando-se então (12) e (3), podemos escrever: $X^+ * X^+ = E' \Sigma^{-1} P_z E + E' \Sigma^{-1} P_z Z G + G' Z' \Sigma^{-1} P_z E + G' Z' \Sigma^{-1} Z G$. Visto que $\mathcal{E}(E) = 0$ obtemos $\mathcal{E}(X^+ * X^+) = G' Z' \Sigma^{-1} Z G + \mathcal{E}(E' \Sigma^{-1} P_z E)$. Sob a hipótese

⁷Tal acontece com as estimativas de Σ consideradas na seção 4.

(11) obtemos $\mathcal{E}(E'\Sigma^{-1}P_zE) = \Omega \text{Tr}(P_z) = l\Omega$, de maneira que, face à (14), a matriz de ACP-RA *teórica*, para Σ conhecida, será dada por:

$$\mathcal{V}_x = G'Z'\Sigma^{-1}ZG + l\Omega \quad (14a)$$

Notemos que $\mathcal{V}_x$ preserva a simetria da covariação entre X^+ e X , destacada por Amato (1988, p.721) para o caso determinístico. Esta covariação é também igual à medida de variação entre as colunas de X^+ : $\mathcal{V}_x = \mathcal{E}(X'\Sigma^{-1}X^+) = \mathcal{E}(X^+\Sigma^{-1}X)$. Com base em (14a) podemos deduzir uma matriz de ACP-RA *empírica* (para Σ conhecida) substituindo-se G e Ω pelos seus estimadores, ou seja, $\hat{G}$ e $\hat{\Omega}$ (vide seção 3.4):

$$\hat{\mathcal{V}}_x = \hat{G}'(Z'\Sigma^{-1}Z)\hat{G} + l\hat{\Omega} \quad (15)$$

Para $\Sigma = I_n$, o projetor P_z dado em (10) coincide com aquele dado (2), de maneira que, a partir de (13) e de (15) obtemos $\hat{\mathcal{V}}_x = X'P_zX + l\hat{\Omega}$. Reencontramos assim a *matriz* de ACP-R (determinística) V_x , dada em (1), colocando $\hat{\Omega}$ (ou Ω) = 0.

2. Para Σ *desconhecida* (ou estimada sobre $\mathbf{R}_x$)

Quando as colunas de X e de Z são padronizadas (simples ou multiplamente), parece-nos apropriado inscrever a definição da matriz de ACP-RA no contexto teórico das *correlações parciais*. Neste sentido, definiremos esta matriz ($\mathcal{V}_{\hat{x}}$) pela esperança matemática de $\hat{X}'\hat{X}$ ⁸ condicionalmente ao valor ($\hat{\Sigma}$), estimado para Σ :

$$\mathcal{V}_{\hat{x}} = \mathcal{E}(\hat{X}'\hat{X}|\Sigma = \hat{\Sigma}) \quad (16)$$

Substituindo-se agora Σ por $\hat{\Sigma}$ em (10), (12) e (13) e usando (3), obtemos: $\hat{X}'\hat{X} = X'\hat{P}_z'\hat{P}_zX = G'Z'ZG + G'Z'\hat{P}_z'\hat{P}_zE + E'\hat{P}_z'\hat{P}_zZG + E'\hat{P}_z'\hat{P}_zE$. Considerando-se agora a hipótese (11) condicionalmente a $\Sigma = \hat{\Sigma}$, temos $\mathcal{E}(E'\hat{P}_z'\hat{P}_zE) = \Omega \text{Tr}\{(Z'\hat{\Sigma}^{-1}Z)^{-1}Z'Z\}$. Face à (16),

⁸Esperança matemática eventualmente assintótica, segundo os estimadores de G , Ω e Σ utilizados.

a matriz de ACP-RA *teórica*, para Σ estimada sobre R_x , será dada por:

$$\mathcal{V}_{\hat{x}} = G'Z'ZG + c_s\Omega \quad (16a)$$

$$c_s = \text{Tr}(Z'\hat{\Sigma}^{-1}Z)^{-1}Z'Z \quad (17)$$

Notemos que a simetria das covariações, obtida anteriormente para $\mathcal{V}_x$ em (14a), é aqui preservada: $\mathcal{V}_{\hat{x}} = \mathcal{E}(\hat{X}'X) = \mathcal{E}(X'\hat{X})$.

Podemos também obter uma matriz de ACP-RA *empírica* substituindo G e Ω pelos seus estimadores, ou seja:

$$\hat{\mathcal{V}}_{\hat{x}} = \hat{G}'Z'Z\hat{G} + c_s\hat{\Omega} \quad (18)$$

onde c_s é definido em (17). Evidentemente, obtemos a matriz de ACP-R (determinística) $\mathcal{V}_x$, dada em (1), substituindo-se em (18) $\hat{\Sigma}$ e $\hat{\Omega}$ por I_n e O , respectivamente.

4. Estimativas paramétricas.

As matrizes paramétricas Σ e Ω são habitualmente desconhecidas. Duas estruturas estocásticas notáveis reduzem o número de parâmetros de Σ a serem estimados: *A Heterocedasticidade e a Auto-correlação* residual. White (1984, p. 174-85) apresenta as condições sob as quais é possível obter, em ambos os casos, estimativas consistentes para todos os parâmetros, com o concurso de *variáveis instrumentais*. Sob a hipótese de *Normalidade* dos resíduos $E(E \sim N_{nk}\{0; \Omega \otimes \Sigma\})$, o estimador de Ω que otimiza a *fdp* conjunta de $x_1, x_2, \dots, x_k$, quando Σ é conhecida, pode ser aproximado como:

$$\hat{\Omega} = \left(\frac{1}{n-l}\right)\hat{E}'\Sigma^{-1}\hat{E} \quad (19)$$

onde $\hat{E} = X - Z\hat{G}$, sendo $\hat{G}$ dado em (13) e com $\hat{\Omega}$ corrigindo o viés do estimador $\left[\frac{n-l}{l}\hat{\Omega}\right]$ nas amostras finitas. Por outro lado, propriedades assintóticas desejáveis podem ser obtidas na estimação dos parâmetros do modelo (*consistência, normalidade em distribuição*,

eficiência) quando Σ se expressa como função de um m -vetor paramétrico Θ estimável ($m < n$)⁹. Modelos de *Auto-regressão vetorial* generalizada também são encontrados na literatura econométrica, onde a estimação das correlações residuais envolve estimativas MQO e MQG em várias etapas. Para o contexto presente (ACP-RA) concebemos, a título ilustrativo, um modelo auto-regressivo simplificado a partir do qual, estimativas por *Máxima Verossimilhança* (MV) são obtidas para os seus parâmetros. Paralelamente, estimadores do tipo de Aitken também serão considerados (2SAE: *Two Stage Aitken Estimators*).

Resíduos *Auto-Regressivos* (MLG-A)

Consideremos o modelo *Markoviano* de 1ª ordem:

$$\varepsilon_i + \rho_i \varepsilon_{i-} = \eta_i \quad (i = 1, 2, \dots, k), \quad (20)$$

onde $\varepsilon'_{i-} = (\varepsilon_{i0}, \varepsilon_{i1}, \varepsilon_{i2} \dots \varepsilon_{in-1})$ é o resíduo defasado (1 período) correspondente à *iésima* equação $x_i = Zg_i + \varepsilon_i$, onde ρ_i é o coeficiente de auto-correlação serial da *iésima* equação ($|\rho_i| < 1$) e $\{\eta_i\}$ é um processo estocástico *Normal* estacionário, de média nula e matriz Var.Cov. constante: $E(\eta_i) = 0$ e $\text{Cov}(\eta_i, \eta_j) = \omega_{ij} I_n$. Coloquemos $\varepsilon_{i0} = \varepsilon_{i(n+1)} = 0$ ($i = 1, 2, \dots, k$). Definindo-se a matriz (n, n) regular:

$$S_i = \begin{bmatrix} 1 & 0 & & 0 \\ \rho_i & 1 & & \\ & \ddots & \ddots & \\ 0 & & & 1 & 0 \\ & & & \rho_i & 1 \end{bmatrix} \quad (21)$$

podemos escrever a forma (20) como:

$$S_i \varepsilon_i = \eta_i \quad i = 1, 2, \dots, k \quad (21a)$$

Construamos agora o kn -vetor $N'_v = [\eta'_1, \eta'_2, \dots, \eta'_k]$ e um sistema vetorial correspondente ao sistema (3):

$$X_v = (I \otimes Z)G_v + E_v, \quad (22)$$

⁹Veja-se JUDGE *et. al.* 1985, p. 169-182 para o caso $k = 1$.

onde $X_v = \text{vec}(X)$, $G_v = \text{vec}(G)$, $E_v = \text{vec}(E)$. Podemos escrever as k equações (21a) da maneira seguinte: $SE_v = N_v$, onde $S = \sum_{i=1}^n (e_i e_i' \otimes S_i)$. Efetuando a transformação inversa $E_v = S^{-1}N_v$, obtemos $\mathcal{E}(E_v) = 0$ e:

$$V(E_v) = \sum_{i,j}^{k,k} [\omega_{ij} e_i e_j' \otimes (S_j' S_i)^{-1}] \quad (23)$$

O inverso de $V(E_v)$ é dado por: $V^{-1}(E_v) = \sum_{i,j}^{k,k} (\omega^{ij} e_i e_j' \otimes S_i' S_j)$, onde ω^{ij} é o elemento (i, j) de Ω^{-1} . Por outro lado, $|S_i| = 1$ ($i = 1, 2, \dots, k$) de maneira que $|S| = 1$ e $|V(E_v)| = |V(N_v)| = |\Omega \otimes I_n| = |\Omega|^n$.

A verossimilhança amostral para G_v , Ω e S correspondente à (22) escrever-se-á, então:

$$L(S, \Omega, G_v | X_v, Z) = (2\pi)^{-\frac{1}{2}nk} |\Omega|^{-1} |\Omega|^{-\frac{1}{2}n}.$$

$$\exp -\frac{1}{2} [X_v - (I \otimes Z)G_v]' \left[\sum_{i,j}^{k,k} \omega^{ij} e_i e_j' \otimes S_i' S_j \right] [X_v - (I \otimes Z)G_v]$$

A maximização de L com relação à G_v , ω^{ij} e ρ_i conduzirá aos seguintes estimadores (MV):

$$\hat{G}_v = \left[\sum_{i,j}^{k,k} \hat{\omega}^{ij} e_i e_j' \otimes Z' \hat{S}_i' \hat{S}_j Z \right]^{-1} \left[\sum_{j=1}^k e_j \otimes \left(\sum_{p=1}^k \hat{\omega}^{jp} Z' \hat{S}_j' \hat{S}_p x_p \right) \right] \quad (24)$$

$$\hat{\Omega} = \left(\frac{1}{n-l} \right) \sum_{i,j}^{k,k} (e_i e_j' \otimes \hat{\varepsilon}_i' \hat{S}_i' \hat{S}_j \hat{\varepsilon}_j) \quad (\text{corrigido}) \quad (25)$$

$$\hat{\rho}_i = -\frac{1}{2} [\hat{\varepsilon}_i' B \hat{\varepsilon}_i / \hat{\varepsilon}_i' \hat{\varepsilon}_i + (\hat{\varepsilon}_i' \sum_{j(\neq i)}^k \hat{\omega}^{ij} \hat{A}_j \hat{\varepsilon}_j) / \hat{\omega}^{ii} \hat{\varepsilon}_i \hat{\varepsilon}_i] \quad (26)$$

onde $B = [b_{rs}]$ e $A = [a_{rs}]$ são as matrizes (n, n) regulares definidas como:

$$a_{rs} = \begin{cases} 1 & \text{se } s = r + 1 \\ 0 & \text{se } s \neq r + 1 \end{cases}; \quad b_{rs} = \begin{cases} 1 & \text{se } s = r \pm 1 \\ 0 & \text{se } s \neq r \pm 1 \end{cases}$$

e com $\hat{A}_j = \hat{\rho}_j I_n + A$. Notemos que quando as equações $x_i = Zg_i + \varepsilon_i$ não são correlacionadas entre si, $\omega_{ij} = \omega^{ij} = 0 (i \neq j)$, o estimador de ρ_i em (26) reduzir-se-á ao estimador MV obtido para a i ésima equação isolada: $\hat{\rho}_i = -(1/2)\hat{\varepsilon}'_i B \hat{\varepsilon}_i / \hat{\varepsilon}'_i \hat{\varepsilon}_i = -\hat{\varepsilon}'_i \hat{\varepsilon}'_{i-} / \hat{\varepsilon}'_i \hat{\varepsilon}_i$. Podemos também escrever a solução conjunta para $\hat{\rho}' = (\hat{\rho}_1, \hat{\rho}_2, \dots, \hat{\rho}_k)$ nas k equações (26), unicamente em função de $\hat{\varepsilon}$ e dos elementos de $\hat{\Omega}^{-1}$. Notando

$$\hat{\delta}_i = (2\hat{\varepsilon}'_i \hat{\varepsilon}_{i-}) / (\hat{\varepsilon}'_i \hat{\varepsilon}_i) + \left(\sum_{j(j \neq i)}^k \hat{\omega}^{ij} \hat{\varepsilon}'_{i-} \hat{\varepsilon}_j \right) / (\hat{\omega}^{ii} \hat{\varepsilon}'_i \hat{\varepsilon}_i) \quad (27)$$

e

$$\hat{\gamma}_{ij} = \begin{cases} \hat{\omega}^{ji} \hat{\varepsilon}'_i \hat{\varepsilon}_j / \hat{\omega}^{jj} \hat{\varepsilon}'_j \hat{\varepsilon}_j & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases} \quad (28)$$

e definindo-se o vetor $(1, k)$, $\hat{\delta}' = [\hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_k]$ e a matriz (k, k) , $\hat{\Gamma} = [\hat{\gamma}_{ij}]$, poderemos escrever o sistema (26) como:

$$\hat{\rho} = -(2I_k + \hat{\Gamma})^{-1} \hat{\delta}. \quad (29)$$

Para a solução conjunta de $\hat{G}_v, \hat{\Omega}$ e $\hat{\rho}$ em (24), (25) e (29), respectivamente, dois métodos alternativos são usados, o *MV-Iterativo* e o *Aitken Duplo Estágio* (2SAE). Ambos pressupõem uma primeira estimativa por MQO dos parâmetros g_i de cada equação $x_i = Zg_i + \varepsilon_i$, de maneira a obter-se uma estimativa inicial de ρ_i e os resíduos calculados $\hat{\varepsilon}_i$. Com este objetivo, vários procedimentos são recomendados: MV, *Duplo Estágio Iterativo*, *Durbin*, etc.¹⁰ Pelo *MV-Iterativo*, com base na estimativa $\hat{S}_i$ de (21) e nos resíduos $\hat{\varepsilon}_i = x_i - Z\hat{g}_i$, poderemos estimar Ω por (25) e, em seguida, G_v em (24). Em uma segunda etapa, utilizar-se-á os resíduos calculados $\hat{E}_v = X_v - (I \otimes Z)\hat{G}_v$ de (22) e o estimador $\hat{\Omega}$ de (25) para calcularmos $\hat{\delta}_i$ em (27) e $\hat{\gamma}_{ij}$ em (28), $(i, j = 1, 2, \dots, k)$ e assim, reestimarmos ρ por (29). Em condições regulares ($|\hat{\rho}_i| < 1$ e convergência), as iterações $(24) \rightarrow (29) \rightarrow (25) \rightarrow (24) \rightarrow \dots$ poderão prosseguir até um grau de aproximação numérica

¹⁰Veja-se (Kmenta 1978, p. 302-16 ou Johnston 1972, p. 259-64.)

suficiente para $\hat{G}_v$.¹¹ Pelo método *Aitken Duplo Estágio* (2SAE) as equações (27)-(29) não serão usadas. Após a 1ª etapa, efetua-se a transformação de cada equação pré-multiplicando ambos os lados de $x_i = Zg_i + \varepsilon_i$ por $\hat{S}_i$: $\hat{S}_i x_i = \hat{S}_i Zg_i + \hat{S}_i \varepsilon_i$ de maneira a podermos escrever (22) como:

$$X_v^\circ = Z^\circ G_v + N_v^\circ \quad (22a)$$

onde $X_v^\circ = [x_1^\circ, x_2^\circ \dots x_k^\circ]$, $x_i^\circ = \hat{S}_i x_i$, $\sum_{i=1}^k (e_i e_i' \otimes Z_i^\circ)$, $Z_i^\circ = \hat{S}_i Z$; $N_v^\circ = [\eta_1^\circ, \eta_2^\circ \dots \eta_k^\circ]$, $\eta_i^\circ = \hat{S}_i \varepsilon_i$. Supomos que a transformação acima haja suprimido as auto-correlações seriais e assim, $\mathcal{E}(\eta_i^\circ) \equiv 0$; $\text{Cov}(\eta_i^\circ, \eta_j^\circ) \equiv \omega_{ij} I_n$ ($i, j = 1, 2 \dots k$). Assim, (22a) poderá ser tratado como um *Sistema Aparentemente Independente*. O estimador de Aitken para G_v em (22a) será, como sabemos, idêntico ao estimador MV sob a hipótese *normal*: $N_v \sim N_{nk}[0, \Omega \otimes I_n]$, a saber:

$$\begin{aligned} \hat{G}_v^\circ &= [Z^{\circ'}(\Omega \otimes I)^{-1} Z^\circ]^{-1} Z^{\circ'}(\Omega \otimes I)^{-1} X_v^\circ \\ &= \left[\sum_{i,j}^{k,k} (\omega^{ij} e_i e_j' \otimes Z_i^{\circ'} Z_j^\circ) \right]^{-1} \left[\sum_{j=1}^k e_j \otimes \left(\sum_{p=1}^k \omega^{jp} Z_j^{\circ'} x_p^\circ \right) \right] \end{aligned} \quad (24a)$$

Com base nas transformações efetuadas, para Ω conhecida, os estimadores (24a) e (24) são idênticos ($\hat{G}_v = \hat{G}_v^\circ$). O segundo estágio do método envolve um procedimento em duas etapas. Inicialmente, estima-se g_i em cada equação $x_i^\circ = Z_i^\circ g_i + \eta_i^\circ$, aplicando-se MQO à equação (22a). A partir dos resíduos $\hat{\eta}_i^\circ = x_i^\circ - Z_i^\circ \hat{g}_i$ estimamos, então, Ω por:

$$\hat{\Omega}^\circ = \left(\frac{1}{n-l} \right) \sum_{i,j}^{k,k} (e_i e_j' \otimes \hat{\eta}_i^{\circ'} \hat{\eta}_j^\circ) \quad (30)$$

Na segunda etapa, o estimador 2SAE de G_v será obtido substituindo-se em (24a) Ω por $\hat{\Omega}^\circ$ dado em (30). Sua matriz Var.Cov. *Assintótica* poderá ser consistentemente estimada por $[Z^{\circ'}(\hat{\Omega}^\circ \otimes I)^{-1} Z^\circ]^{-1}$.

¹¹Caso a estabilidade numérica de $\hat{G}_v$ não puder ser imediatamente alcançada, pode-se fixar um número (razoável) de iterações (3 ou 4) e reter-se às estimativas que maximizam, por exemplo, a forma quadrática $\hat{F} = \hat{G}_v' [\sum_{i,j}^{k,k} e_i e_j' \otimes \hat{\omega}^{ij} Z' \hat{S}_i \hat{S}_j Z] \hat{G}_v$, ou então, que minimizam a soma das variâncias estimadas de $\hat{G}_v$ ou ainda, $\hat{E}_v' \hat{E}_v$.

Ambos estimadores de G_v (MV-Iterativo e 2SAE) possuem propriedades assintóticas idênticas: não-viés, eficiência e normalidade em distribuição. Além disto, são consistentes e assintoticamente equivalentes ao estimador BLU de Aitken (Kmenta, 1978, p.567-571).

Para a obtenção da matriz de ACP-RA, entretanto, as expressões dadas em (14a) e (16a) não são diretamente aplicáveis, pois neste caso, como vemos por (23), a matriz-norma das projeções de X sobre R_z varia segundo as coordenadas consideradas: $\Sigma = (S'_j S_i)^{-1}$.

5. Subespaços de referência múltiplos.

Consideremos então a projeção de $x' = [x_1, x_2, \dots, x_k]$ sobre k subespaços de referência R_{z_i} , gerados pelas colunas das matrizes Z_i de dimensão (η, l_i) , eventualmente distintas ($i = 1, 2, \dots, k$). O modelo de restrições estocásticas implícito será $x_i = Z_i g_i + \varepsilon_i$, o qual escrever-se-á, em forma sistêmica equivalente:

$$X_v = \bar{Z} G_v + E_v, \quad (31)$$

onde $\bar{Z} = \sum_1^k (e_i e'_i \otimes Z_i)$ é de dimensão $(kn, \bar{l})$, $\bar{l} = \sum_{i=1}^k l_i$ e G_v é o vetor paramétrico $(\bar{l}, 1)$. O vetor estocástico E_v vem suposto *Normal* de média nula e matriz Var.Cov. $V(E_v)$ igual à:

$$\Xi = \sum_{i,j}^{k,k} (e_i e'_j \omega_{ij} \otimes \Sigma_{ij}), \quad (32)$$

onde $\Omega = [\omega_{ij}]$ é simétrica de ordem k e Σ_{ji} de ordem n ($\Sigma_{ij} = \Sigma'_{ji}$), ambas positivas definidas ($i, j = 1, 2, \dots, k$). Consideremos agora R_x^v o subespaço gerado por X_v sobre R^{nk} , e os subespaços de dimensão l_i gerados sobre R^n pelas k matrizes Z_i , ou seja, R_{z_i} . Formemos o espaço-produto $R_z^v = \otimes_1^k R_{z_i}$. Se elegermos Ξ^{-1} por matriz-norma diretora do produto interno $(*)$ [definido em (9)] sobre R^{nk} , temos que um projetor ortogonal sobre R_z^v poderá ser construído como: $\bar{P}_z = \bar{Z} [\bar{Z}' \Xi^{-1} \bar{Z}]^{-1} \bar{Z}' \Xi^{-1}$. Como vimos em 3.1, quando Ξ é uma matriz conhecida, a imagem da projeção de X_v sobre R_z^v , ou seja, $X_v^+ = P_z X_v$ é idêntica à imagem calculada ($\hat{X}$) na regressão (31),

quando G_v é avaliado pelo estimador MV (ou Aitken), ou seja, $\hat{X}_v = \bar{Z}\hat{G}_v$, onde $\hat{G}_v = (\bar{Z}'\Xi^{-1}\bar{Z})^{-1}\bar{Z}'\Xi^{-1}X_v$.

Quando há ausência de correlação entre os resíduos das k equações ($\omega_{ij} = 0$, $i \neq j$) verificamos que $\bar{P}_z = \sum_{i=1}^k (e_i e_i' \otimes Z_i(Z_i'\Sigma_{ii}^{-1}Z_i)^{-1}Z_i'\Sigma_{ii}^{-1}) = \text{Diag}(P_{z_1}, P_{z_2}, \dots, P_{z_k})$, onde $P_{z_i} = Z_i(Z_i'\Sigma_{ii}^{-1}Z_i)^{-1}Z_i'\Sigma_{ii}^{-1}$, de maneira que teremos $x_i^+ = [P_{z_i}]x_i$ e a projeção de x_i sobre R_z^v estará circunscrita ao subespaço R_{z_i} .

Quando apenas Ω for *desconhecida*, poderá ser substituída pelo seu estimador consistente, seja $\hat{\Omega}$, calculado através dos resíduos de *Mínimos Quadrados Generalizados* ($\hat{\varepsilon}_i$) obtidos na estimativa paramétrica das equações $x_i = Z_i g_i + \varepsilon_i$ ($i = 1, 2, \dots, k$), tomando inicialmente Ω como uma matriz diagonal. Assim, $\hat{\varepsilon}_i = x_i - Z_i \hat{g}_i$, $\hat{g}_i = (Z_i'\Sigma_{ii}^{-1}Z_i)^{-1}Z_i'\Sigma_{ii}^{-1}x_i$ e, em uma segunda fase, obteremos estimadores consistentes de ω_{ij} através de $\hat{\omega}_{ij} = \frac{1}{\tau_{ij}}\hat{\varepsilon}_i'\Sigma_{ij}^{-1}\hat{\varepsilon}_j$, com $\tau_{ij} = \text{Tr}M_i'\Sigma_{ij}^{-1}M_j\Sigma_{ji}$, e $M_i = I - P_{z_i}$ ¹² e desde que $\tau_{ij} > 0$. Valores consistentes para $\hat{G}_v$ poderão então ser obtidos substituindo-se Ξ por $\hat{\Xi}$ [dado através de (32)] no estimador de Aitken. Também, a matriz Var.Cov. assintótica de $\hat{G}_v$ será consistentemente estimada por $[\bar{Z}'\hat{\Xi}^{-1}\bar{Z}]^{-1}$.

Matrizes de ACP-RA

Para a obtenção da matriz de ACP-RA neste contexto, deveremos considerar o seu elemento típico (v_{ij}) o qual, quando Σ_{ij} são *conhecidas*, será obtido como v_x , dada em (14a): $v_{ij}^+ = \mathcal{E}(\hat{x}_j'\Sigma_{ji}^{-1}\hat{x}_i) = \mathcal{E}(\hat{g}_j'Z_j'\Sigma_{ji}^{-1}Z_i\hat{g}_i) = \text{Tr}[Z_j'\Sigma_{ji}^{-1}Z_i\mathcal{E}(\hat{g}_j\hat{g}_i')]$. Designemos agora por $C_{ij} = \text{Cov}(\hat{g}_i, \hat{g}_j)$ a matriz (l_i, l_j) da covariância entre os estimadores $\hat{g}_i$ e $\hat{g}_j$. Quando $\hat{g}_i, \hat{g}_j$ são não-viesados, temos $\mathcal{E}(\hat{g}_i\hat{g}_j') = C_{ij} + g_i g_j'$ e assim podemos definir o elemento típico v_{ij}^+ da matriz *teórica* v_x como:

$$v_{ij}^+ = g_j'Z_j'\Sigma_{ji}^{-1}Z_i g_i + \text{Tr}[Z_j'\Sigma_{ji}^{-1}Z_i C_{ij}] \quad (33)$$

Para o termo genérico da matriz ACP-RA quando Σ_{ij} são *desconhecidas*, ou seja, $v_x = [\hat{v}_{ij}]_{(k,k)}$ dada em (16a), consideramos a esperança

¹²Para o não-viés de $\hat{\omega}_{ij}$, temos: $\mathcal{E}(\hat{\varepsilon}_i'\Sigma_{ij}^{-1}\hat{\varepsilon}_j) = \mathcal{E}\text{Tr}(\varepsilon_i'M_i'\Sigma_{ij}^{-1}M_j\varepsilon_j) = \omega_{ji}\text{Tr}M_i'\Sigma_{ij}^{-1}M_j\Sigma_{ji} = \omega_{ij}\tau_{ij}$.

assintótica do produto $\hat{x}_i' \hat{x}_j$. Para estimadores $\hat{g}_i$ consistentes obtemos, semelhantemente a v_{ij}^+ :

$$\hat{v}_{ij} = g_j' Z_j' Z_i g_i + \text{Tr}[Z_j' Z_i C_{ij}], \quad (34)$$

com $C_{ij} = \text{Cov}(\hat{g}_i, \hat{g}_j)$. Quando $\Sigma_{ij} = I_n(v_{ij}^+ = \hat{v}_{ij})$, o modelo (31) poderá ser tratado como um *Sistema Aparentemente Independente*, como já assinalamos anteriormente em 3.4, para o caso $Z_i = Z$.

No caso do MLG-A (*Auto-regressivo*), temos $Z_i = Z$ e, face a (23) e (32), $\Sigma_{ij} = (S_j S_i)^{-1}$. Então, $\Sigma_{ji} = \Sigma_{ij}' = (S_i' S_j')^{-1}$ e assim $Z_j' \Sigma_{ji}^{-1} Z_i = Z' S_i' S_j Z$. A submatriz C_{ji} será o bloco (l, l) correspondente de $[\sum_{i,j}^{k,k} (e_i e_j' \omega^{ij} \otimes Z' S_i' S_j Z)]^{-1}$.

Quando ρ_i são conhecidos, utilizaremos (34), substituindo ali Z_i e Σ_{ij} (como anteriormente indicado) e g_i e C_{ij} pelos seus estimadores $\hat{g}_i$ e $\hat{C}_{ij}$ correspondentes.¹³ Quando os ρ_i são estimados pela amostra, valeremo-nos de (35). Notemos que aqui, $\hat{v}_{ij}$ será estimado substituindo-se g_i e C_{ij} por $\hat{g}_i$ e $\hat{C}_{ij}$ ou $\hat{g}_i^o$ e $\hat{C}_{ij}^o$, segundo o método adotado, *MV-Iterativo* ou *2SAE*, respectivamente.

6. Matriz assintótica, inferências e outras aplicações.

Para as grandes amostras, poderemos deduzir uma matriz de ACP-RA assintótica ($\mathcal{V}_{as}$) a partir de uma matriz $\hat{\mathcal{V}}_x$ definida em (15) no caso $Z_i = Z$, a qual, quando $\Sigma = I_n$ poderá escrever-se (usando-se (13) e (19)) $\hat{\mathcal{V}}_x = X' P_z X + l \hat{\Omega} = X' (\frac{n-2l}{n-l} P_z + \frac{l}{n-l} I_n) X$ onde P_z é o projetor já definido em (2). Supondo-se agora a convergência em *probabilidade* das matrizes $(X' X)$, e $(X' Z)$ para matrizes constantes, e a convergência *simples* de $Z' Z$ para uma matriz regular, mostra-se que a matriz de ACP-R ($X' P_z X$) dada em (1) será um estimador *consistente* de $\mathcal{V}_{as}$.

Para a inferência estatística, supomos a *Normalidade* do resíduo E no modelo (3). No plano das estimativas paramétricas, testes de hipóteses e regiões de confiança convencionais podem ser construídos a partir das distribuições de $\hat{G}$ e $\hat{\Omega}$ definidos em (13) e (19)

¹³ Estimadores MV ou 2SAE, os quais neste caso coincidem, dado que as equações (26) e (29) poderão ser suprimidas.

para $\Sigma = I_n$ (Anderson 1958, p.183; Theil 1971, p.402) e dos estimadores $\hat{G}_v^\circ$ e $\hat{\Omega}^\circ$, definidos em (24a) e (30) (Judge *et. al.* 1985, p.493; Kmenta 1978, p.571). Também a nulidade dos coeficientes ρ_i poderá ser testada (Judge *et. al.* 1985, p.493-94). No plano da ACP-RA, com $\Sigma = I_n$ e $Z_i = Z$, a matriz $\hat{V}_x$ expressa no parágrafo anterior desdobra-se como a soma de duas matrizes aleatórias *Wishart*- k independentes, sendo a primeira ($X'P_zX$) não-central de parâmetro Ω e l graus de liberdade e a segunda ($l\hat{\Omega} = \frac{l}{n-l}\hat{E}'\hat{E}$) central, de parâmetro $\frac{l}{n-l}\Omega$ e $n-l$ graus de liberdade¹⁴. Notando D e $\hat{D}$, para as diferenças das matrizes de ACP-RA *teóricas* e *empíricas*, dadas em (14) e (15), com relação à matriz de ACP-R (V_x) dada em (1), respectivamente, temos $D = l\Omega$ e $\hat{D} = l\hat{\Omega}$. Se todos os auto-valores de D (ou Ω) forem distintos, poderemos construir testes de hipóteses e regiões de confiança para as funções significativas dos auto-valores e auto-vetores de D (ou Ω), a partir de uma decomposição espectral de $\hat{D}$ (ou $\hat{\Omega}$) com base na distribuição (assintótica) *Normal* (Anderson 1963)¹⁵. Neste mesmo contexto, poderemos também avaliar a ortogonalidade das variações estocásticas com relação às restrições determinísticas representadas pelo modelo (3), testando diretamente a diagonalidade de Ω a partir das distribuições dos seus estimadores dados em (25) ou em (30) para o caso *auto-regressivo*, e na seção 4, para o caso dos subespaços de referência múltiplos¹⁶.

No contexto $Z_i = Z$, uma extensão notável da análise formal consiste na integração de informações *a priori* sobre G e Ω , em um enfoque *Bayesiano*. Nesta perspectiva, a utilização de uma fdp *a priori Normal-Conjugada* sobre G e Ω , condicionalmente à Σ (Drèze e Fisher, 1974), conduzir-nos-á a uma fdp *a posteriori* conjunta decomponível em uma *Normal-kl* sobre G e uma *Wishart-k* de matriz-parâmetro Ω . Matrizes de ACP-RA *Bayesianas*, *teóricas* e *empíricas*

¹⁴Para a segunda, o teorema de Cochran (1934) aplica-se a $(\frac{l}{n-l})E'(I - P_z)E$, sendo $\hat{E}$ interpretada como uma matriz amostral (veja-se Malinvaud 1969 p.240. Sobre a *Wishart*, Mardia *et. al.* 1982, p.68 e Rao 1973, p.533-43).

¹⁵O teorema de Anderson não pressupõe a Normalidade (veja-se Mardia *et. al.* 1982, p.230-37).

¹⁶Teste dos multiplicadores de Lagrange ou Teste do Quociente de Verossimilhança (Judge *et. al.* 1985, p.476-77).

poderão então ser definidas a partir dos momentos (*a posteriori*) destas distribuições.

Submetido em agosto de 1992. Revisado em março de 1993.

Referências

- Amato, V. 1976. "A proposito di correlazione multipla generalizzata." *Annali Dell'Istituto di Statistica* 40:3-32.
- . 1988. "Riflessioni su un'analisi non simmetrica delle componenti principali." *Rassegna Economica* 52:719-730.
- Anderson, T.W. 1958. *An Introduction to Multivariate Statistical Analysis*. New York: J. Wiley.
- D'Ambra, L. & Lauro, N. 1982. "Analisi in componenti principali in rapporto a un Sottospazio di Riferimento." *Rivista di Statistica Applicata* 15:51-67.
- Drèze, J.H. & Fisher, W.D. 1974. "Bayesian analysis of a single equation." In *CORE Discussion Papers*, nº 7415. Université Catholique de Louvain.
- Goldberger, A.S. 1964. *Econometric Theory*. New York: J. Wiley.
- Johnston, J. 1972. *Econometric Methods*. New York: McGraw-Hill, 2nd. ed.
- Judge, G.G., W.E. Griffiths, R. Carter-Hill, H. Lütkepohl & T.C. Lee. 1985. *The Theory and Practice of Econometrics*. New York: J. Wiley, 2nd. ed.
- Kmenta, J. 1978. *Elementos de Econometria*. São Paulo: Editora Atlas.
- Malinvaud, E. 1969. *Méthodes Statistiques de l'Économétrie*. Paris: Dunod.
- Mardia, K.V., Kent, J.T. & Bibby, J.M. 1982. *Multivariate Analysis*. London: Academic Press.
- Rao, C.R. 1973. *Linear Statistical Inference and its Applications*. New York: J. Wiley, 2nd. ed.
- Theil, H. 1971. *Principles of Econometrics*. New York: J. Wiley.
- White, H. 1984. *Asymptotic Theory for Econometricians*. Orlando: Academic Press.