FUNDAÇÃO
GETULIO VARGAS

EPGE
Escola de Pós-Graduação
em Economia

Ensaios Econômicos

Escola de
Pós-Graduação
em Economia
da Fundação
Getulio Vargas

N° 37 ISSN 0104-8910

Modelo de Equações Simultâneas

Fernando de Holanda Barbosa

Janeiro de 1983

URL: http://hdl.handle.net/10438/604

de Holanda Barbosa, Fernando
Modelo de Equações Simultâneas/
Fernando de Holanda Barbosa - Rio de Janeiro : FGV,EPGE, 2010
(Ensaios Econômicos; 37)

Inclui bibliografia.

CDD-330

MODELO DE EQUAÇÕES SIMULTÂNEAS

Fernando de Holanda Barbosa

A interdependência das variáveis e sua determinação simultânea nos modelos econômicos acarreta alguns problemas estatísticos que têm de ser devidamente levados em conta na prática econométrica. Este trabalho tem como objetivo apresentar uma introdução sucinta ao modelo tradicional de equações simultâneas, discutindo-se o problema de identificação e examinando-se alguns métodos de estimação.

1.1 - Conceitos e Definições Básicas

Os conceitos e definições básicas do modelo de equações simultâneas podem ser introduzidos de um modo didático através de um exemplo concreto. Para tal finalidade considere-se um modelo de mercado de um bem em que a quantidade e preço são determinados simultaneamente.

Forma Estrutural do Modelo

A equação de demanda de mercado pelo bem é expressa por:

(1) $$q_t^d = \gamma_{12}\ p_t + \beta_{11}\ y_t + \beta_{12}\ S_t + u_{1t}$$

onde o índice t refere-se ao período que a variável corresponde, q_t^d é a quantidade demandada, p_t é o preço de mercado do bem, y_t é a renda real dos consumidores, S_t é o preço de um bem substituto do bem em estudo, u_{1t} é o termo aleatório cujas propriedades sarão especificadas mais adiante e γ_{12}, β_{11} e β_{12} são os parâmetros estruturais da equação de demanda.

A equação de oferta de mercado do bem é dada por:

(2) $$q_t^s = \gamma_{22}\ p_t + \beta_{23}\ \omega_t + \beta_{24}\ p_{t-1} + u_{2t}$$

onde q_t^s é a quantidade ofertada, ω_t é o preço de um insumo utilizado na produção do bem, p_{t-1} é o preço do produto defasado de um período, u_{2t} é o termo aleatório e γ_{22}, β_{23} e β_{24} são os parâmetros estruturais da equação de oferta. Observe que na especificação (2) a quantidade ofertada é função do preço do bem no período t e do preço com um período de defasagem. Note-se, também, que algumas variáveis que aparecem na equação de demanda não fazem parte da equação de oferta, e vice-versa.

O equilíbrio de mercado se dá quando a quantidade demandada for igual à quantidade ofertada. Em símbolos:

(3) $$q_t^d = q_t^s = q_t$$

onde q_t é a quantidade transacionada no mercado.

As relações (1), (2) e (3) são denominadas equações estruturais em virtude de descreverem a estrutura da economia, no caso a estrutura do mercado do bem em análise. As três equações conjuntamente constituem a forma estrutural do modelo.

As equações estruturais, de um modo geral, podem ser classificadas em três tipos: a) equações de comportamento; b) equações técnicas ou institucionais; e c) identidades ou igualdades. As identidades são definições de variáveis enquanto as igualdades produzem condições de equilíbrio ou ajustamento entre variáveis. As equações técnicas são equações do tipo função de produção e as institucionais refletem peculiaridades de organização institucional como é o caso, por exemplo, de equações de impostos com suas alíquotas expressas em lei. As equações de comportamento, como o nome indica, procuram descrever o comportamento dos agentes econômicos. No modelo de mercado descrito nos parágrafos anteriores, as duas primeiras equações são de comportamento e a terceira equação é uma igualdade que fornece a condição de equilíbrio do modelo.

Variáveis Endógenas, Exógenas e Predeterminadas

As variáveis de um modelo estrutural são de três tipos: endógenas, endógenas defasadas e exógenas. As variáveis endógenas têm seus valores determinados através do modelo. As variáveis exógenas são determinadas "fora" do modelo. As variáveis endógenas defasadas, como seu próprio nome indica, são variáveis endógenas que entram em equações estruturais com valores defasados. As variáveis exógenas e endógenas defasadas são, conjuntamente, denominadas de variáveis predeterminadas. O que caracteriza as variáveis exógenas é o fato de não estarem correlacionadas com os termos aleatórios das equações do modelo dos quais participam.

No modelo de mercado formado pelas equações (1), (2) e (3), as variáveis endógenas são o preço p_t, a quantidade demandada q_t^d e a quantidade ofertada q_t^s. As variáveis exógenas são: a renda y_t, o preço S_t do bem substituto e o preço ω_t do insumo. A variável endógena defasada aparece na equação de oferta e é o preço p_{t-1} do produto defasado de um período. As variáveis predeterminadas y_t, S_t, ω_t e p_{t-1} , por hipótese, não estão correlacionadas comos termos aleatórios u_{1t} e u_{2t}.

Estrutura Estocástica

No tocante à estrutura estocástica do modelo admitiremos que os termos aleatórios u_{1t} e u_{2t} são serialmente independentes e que têm valores esperados iguais a zero:

$$E\ u_{1t} = E\ u_{2t} = 0$$

A matriz de variância-covariância independe do tempo t e é expressa por:

$$E\begin{bmatrix} u_{1t}^2 & u_{1t}u_{2t} \\ u_{2t}u_{1t} & u_{2t}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

e obviamente $\sigma_{21} = \sigma_{12}$.

Forma Reduzida do Modelo

As identidades ou igualdades porventura existentes em um modelo estrutural podem ser eliminadas do modelo diminuindo-se o número de variáveis endógenas através da substituição de uma variável nas demais equações pelo seu valor na igualdade ou identidade. No nosso exemplo, a equação (3) pode ser eliminada do modelo substituindo-se q_t^d e q_t^s, nas duas primeiras equações por q_t. O resultado a que se chega, em notação matricial, é o seguinte:

$$(4) \qquad \begin{bmatrix} 1 & -\gamma_{12} \\ 1 & -\gamma_{22} \end{bmatrix} \begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & 0 & 0 \\ 0 & 0 & \beta_{23} & \beta_{24} \end{bmatrix} \begin{bmatrix} y_t \\ S_t \\ \omega_t \\ p_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

Pós multiplicando-se ambos os lados desta equação pela matriz

$$\begin{bmatrix} 1 & -\gamma_{12} \\ 1 & -\gamma_{22} \end{bmatrix}^{-1} = \frac{1}{\gamma_{12}-\gamma_{22}} \begin{bmatrix} -\gamma_{22} & \gamma_{12} \\ -1 & 1 \end{bmatrix}$$

obtém-se

$$(5) \qquad \begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \end{bmatrix} \begin{bmatrix} y_t \\ S_t \\ \omega_t \\ p_{t-1} \end{bmatrix} + \begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}$$

onde:

$$(6) \qquad \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \end{bmatrix} = \frac{1}{\gamma_{12}-\gamma_{22}} \begin{bmatrix} -\beta_{11}\gamma_{22} & -\beta_{12}\gamma_{22} & \beta_{23}\gamma_{12} & \beta_{24}\gamma_{12} \\ -\beta_{11} & -\beta_{12} & \beta_{23} & \beta_{24} \end{bmatrix}$$

e

$$\begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix} = \frac{1}{\gamma_{12}-\gamma_{22}} \begin{bmatrix} -\gamma_{22}u_{1t}+\gamma_{12}u_{2t} \\ -u_{1t}+u_{2t} \end{bmatrix}$$

O sistema de equações (5) é denominado de forma reduzida do modelo por expressar cada variável endógena como função das variáveis predeterminadas, e os coeficientes π_{ij} são os parâmetros da forma reduzida.

Os termos aleatórios v_{1t} e v_{2t} da forma reduzida não estão correlacionados com as variáveis exógenas em virtude dessas variáveis não estarem correlacionadas com os termos aleatórios u_{1t} e u_{2t} . Consequentemente, desde que as demais hipóteses do modelo de

regressão sejam satisfeitas pode-se aplicar o método de mínimos quadrados ordinários na estimativa dos parâmetros de cada equação da foma reduzida (5).

1.2 - Identificação do Modelo

Uma vez que os parâmetros da forma reduzida do modelo sempre podem ser estimados surge, então, a questão de saber se a partir dessas estimativas os coeficientes da forma estrutural do modelo podem ou não serem estimados. Este tipo de indagação, constitui-se no chamado problema de identificação.

Para estudar esse tipo de problema consideremos, em primeiro lugar, a identificação dos parâmetros da equação de demanda (1). Subtraindo-se dos elementos da primeira linha da matriz que aparece no lado esquerdo de (6) os elementos da segunda linha, depois de multiplicados pelo coeficiente γ_{12}, resulta:

(7) $$\begin{cases} \pi_{11} - \pi_{21}\gamma_{12} = \beta_{11} \\ \pi_{12} - \pi_{22}\gamma_{12} = \beta_{12} \\ \pi_{13} - \pi_{23}\gamma_{12} = 0 \\ \pi_{14} - \pi_{24}\gamma_{12} = 0 \end{cases}$$

Essas equações podem ser escritas na seguinte forma matricial:

(8) $$Q_1 \ \delta_1 = \pi_1$$

onde:

$$\pi_1 = \begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{14} \end{bmatrix}, Q_1 = \begin{bmatrix} \pi_{21} & 1 & 0 \\ \pi_{22} & 0 & 1 \\ \pi_{23} & 0 & 0 \\ \pi_{24} & 0 & 0 \end{bmatrix}, \delta_1 = \begin{bmatrix} \gamma_{12} \\ \beta_{11} \\ \beta_{12} \end{bmatrix}$$

A estimação da forma reduzida do modelo por mínimos quadrados ordinários permite que se obtenha estimativas do vetor π_1 e da matriz Q_1. O problema de identificação consiste, então, em saber sob que condições o vetor δ_1 pode ser conhecido a partir dessas estimativas. Formalmente, o problema reduz-se ao estabelecimento das condições em que o sistema de equações lineares (8) tem solução. Observe-se que a primeira equação do modelo estrutural em estudo, a equação de demanda, contém duas variáveis endógenas, $L_1 = 2$, e duas variáveis predeterminadas (≡ exógenas), $K_1 = 2$. O modelo completo, como formalizado no sistema (4), tem duas variáveis endógenas, $L = 2$, e quatro variáveis predeterminadas, $K = 4$. A matriz Q_1 é, portanto, de ordem $K \times (L_1 - 1 + K_1)$, o vetor coluna π_1 contém K elementos e δ_1 é um vetor com $(L_1 - 1 + K_1)$ elementos.

Condição de Posto Para Identificação

A condição necessária e suficiente para que os parâmetros da equação de demanda sejam identificados é que o posto (≡ rank) da matriz Q_1 seja igual ao seu número de colunas. Isto é:

$$\rho(Q_1) = L_1 - 1 + K_1 = 3$$

onde $\rho(Q_1)$ indica o posto da matriz Q_1. Essa condição é denominada de condição de posto para identificação.

Condição de Ordem Para Identificação

Uma condição necessária, embora não suficiente, para que os parâmetros da equação da demanda sejam identificados, denominada de condição de ordem para identificação, é que o número de colunas da matriz Q_1 seja menor ou igual ao seu número de linhas. Obviamente, esta é uma condição necessária para que o posto da matriz Q_1 seja igual a $L_1 - 1 + K$. Em símbolos, a condição de ordem é a seguinte:

$$K \geq L_1 - 1 + K_1$$

ou ainda:

$$K_2 \geq L_1 - 1$$

Esta última desigualdade foi obtida levando-se em conta que o número de variáveis predeterminadas do modelo é igual à soma do número de variáveis predeterminadas que aparecem na equação de demanda e do número de variáveis predeterminadas que foram excluídas da referida equação: $K = K_1 + K_2$.

A condição de ordem é uma condição de fácil verificação pois requer apenas a contagem do número de variáveis predeterminadas excluidas da equação e do número de variáveis endógenas incluidas na equação. Se o número de variáveis predeterminadas excluídas for maior ou igual ao número de variáveis endógenas incluídas na equação, menos 1, é possível que a equação seja identificada. Caso contrário, pode-se afirmar que a equação não é identificada

Equação Superidentificada

Admita que os coeficientes β_{23} e β_{24} sejam diferentes de zero. Portanto, de acordo com (6), os parâmetros π_{23} e π_{24} são, também, diferentes de zero. Segue-se, então, que o coeficiente γ_{12} pode ser obtido de duas maneiras diferentes, segundo (7) , através das seguintes fórmulas:

$$\gamma_{12} = \frac{\pi_{13}}{\pi_{23}}$$

e

$$\gamma_{12} = \frac{\pi_{14}}{\pi_{24}}$$

Observe-se que, neste caso, o número de linhas da matriz Q_1 é maior do que o seu número de colunas pois $K>L_1 - 1+K_1$. A equação de demanda é então dita superidentificada em virtude de não existir solução única para a estimativa do parâmetro γ_{12} a partir das estimativas dos parâmetros de forma reduzida. Cabe salientar que esse problema pode ser contornado como se mostrará mais adiante.

Equação Exatamente Identificada

Admita-se agora que, por exemplo, o parâmetro β_{24} seja igual a zero. Esta hipótese implica em que o coeficiente π_{24} também é igual a zero. Assim, a última equação de (7) deixa de existir e a matriz Q_1 passa a ser uma matriz quadrada de ordem igual ao número K de variáveis predeterminadas do modelo. Segue-se, então, que o parâmetro estrutural γ_{12} agora é calculado através de uma única expressão, que é a seguinte:

$$\gamma_{12} = \frac{\pi_{13}}{\pi_{23}}$$

A equação de demanda neste caso é dita exatamente identificada em virtude do parâmetro γ_{12} ser obtido por um única via.

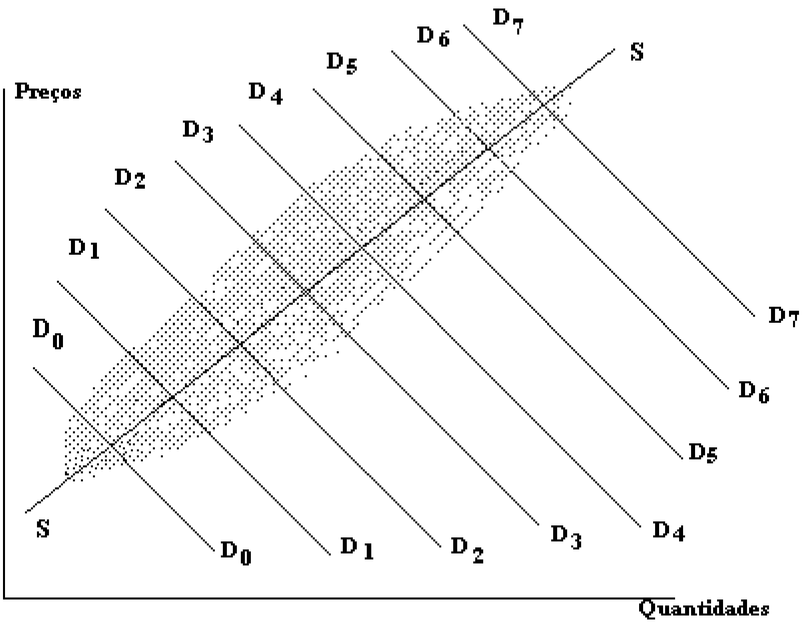
Equação Não Identificada

Suponha que tanto o coeficiente β_{23} como o coeficiente β_{24} sejam iguais a zero, o que significa dizer que $\pi_{23} = \pi_{24} = 0$. Esta hipótese faz com que o sistema de equações (7) reduza-se a um sistema de duas equações e três incógnitas, $\gamma_{12},$ β_{11} e β_{12}. A conclusão óbvia é de que é impossível estimar-se os parâmetros $\gamma_{12},$ β_{11} e β_{12} , individualmente, a partir dos valores conhecidos das estimativas dos parâmetros da forma reduzida do modelo.

É interessante examinar-se com mais detalhes este caso para que se possa compreender melhor o porquê da impossibilidade de identificação dos parâmetros da equação de demanda, quando a quantidade ofertada é função apenas do preço do produto. A Figura 1 mostra que, a não ser por choques aleatórios, a curva de oferta permanece estável, enquanto a curva de demanda se desloca em virtude de mudanças na renda dos consumidores e no preço do bem substituto. Estes deslocamentos da curva de demanda geram observações de preços e quantidades que estão dispersos em torno da curva de oferta. Consequentemente, estes dados não permitirão que se obtenha os parâmetros da equação de demanda mas sim os parâmetros da curva de oferta.

Figura 1

IDENTIFICAÇÃO DA EQUAÇÃO DE OFERTA



Identificação da Equação de Oferta

Subtraindo-se dos elementos da primeira linha da matriz (6) os elementos da segunda linha, depois de serem multiplicados pelo parâmetro estrutural γ_{22}, resulta no seguinte sistema de equações:

$$\begin{cases} \pi_{11} - \gamma_{22}\pi_{21} = 0 \\ \pi_{12} - \gamma_{22}\pi_{22} = 0 \\ \pi_{13} - \gamma_{22}\pi_{23} = \beta_{23} \\ \pi_{14}\gamma_{22}\pi_{24} = \beta_{24} \end{cases}$$

Alternativamente, esta equação pode ser escrita em forma matricial do seguinte modo:

$$\pi_1 = Q_2 \ \delta_2$$

onde:

$$\pi_1 = \begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{14} \end{bmatrix}, Q_2 = \begin{bmatrix} \pi_{21} & 0 & 0 \\ \pi_{22} & 0 & 0 \\ \pi_{23} & 1 & 0 \\ \pi_{24} & 0 & 1 \end{bmatrix}, \delta_2 = \begin{bmatrix} \gamma_{22} \\ \beta_{23} \\ \beta_{24} \end{bmatrix}$$

Observe-se que a matriz Q_2 é diferente da matriz Q_1. De maneira análoga ao que foi estabelecido para a equação de demanda, a condição necessária e suficiente para que os parâmetros da equação de oferta sejam identificados é que o posto de matriz Q_2 seja igual ao seu número de colunas. Isto é:

$$\rho(Q_2) = L_2 - 1 + K_2$$

onde L_2 e K_2 indicam, respectivamente, o número de variáveis endógenas e o número de variáveis predeterminadas incluídas na equação da oferta.

Identificação: Uma Regra Prática

A condição de posto para identificação, como apresentada aqui, requer que se conheça uma matriz que tem como alguns dos seus elementos coeficientes da forma reduzida do modelo. Este procedimento não é prático em virtude das equações de um modelo de equações simultâneas serem especificadas, em geral, sob a forma estrutural. Torna-se, portanto, bastante atrativo o estabelecimento de uma condição de posto que necessite apenas do exame da matriz dos coeficientes da forma estrutural do modelo, o que certamente facilitará o estudo da identificação de cada equação do modelo. O restante desta subseção é dedicado a esse assunto.

Considere-se um modelo de equações simultâneas com L variáveis endógenas e K variáveis predeterminadas. Imagine que se disponha de uma amostra de tamanho T para estimar-se o modelo. A iésima equação do modelo, para o período t, será dada por:

$$\gamma_{il}\ y_{1t} + \gamma_{i2}\ y_{2t} + \ldots + \gamma_{iL}\ y_{Lt} = \beta_{il}\ x_{1t} + \ldots + \beta_{iK}\ x_{Kt} + u_{it}\,, \quad \begin{array}{l} i = 1, \ldots, L \\ t = 1, \ldots, T \end{array}$$

Em notação matricial, este sistema pode ser escrito da seguinte forma:

(9) $\quad Y\,\Gamma = X\,B + U$

onde:

$$Y = \begin{bmatrix} y_{11}\,y_{21} \cdots y_{L1} \\ y_{12}\,y_{22} \cdots y_{L2} \\ ''' \\ ''' \\ ''' \\ y_{1T}\,y_{2T} \cdots y_{LT} \end{bmatrix}, X = \begin{bmatrix} x_{11}\,x_{21} \cdots x_{K1} \\ x_{12}\,x_{22} \cdots x_{K2} \\ ''' \\ ''' \\ ''' \\ x_{1T}\ x_{2T}\ x_{KT} \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} y_{11} y_{21} \cdots y_{L1} \\ y_{12} y_{22} \cdots y_{L2} \\ ''' \\ ''' \\ ''' \\ y_{1L}\, y_{2L} \cdots y_{LL} \end{bmatrix}, B = \begin{bmatrix} \beta_{11} \beta_{21} \cdots \beta_{L1} \\ \beta_{12} \beta_{22} \cdots \beta_{L2} \\ ''' \\ ''' \\ ''' \\ \beta_{1K}\, \beta_{2K}\, \beta_{LK} \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} u_{21} \ldots u_{L1} \\ u_{12} u_{22} \ldots u_{L2} \\ ''' \\ ''' \\ ''' \\ u_{1T}\, u_{2T} \ldots u_{LT} \end{bmatrix}$$

A matriz Y das variáveis endógenas é uma matriz de ordem $T \times L$, X é uma matriz de ordem $T \times K$ das variáveis predeterminadas, Γ é uma matriz quadrada de tamanho L cujos elementos são os coeficientes das variáveis endógenas, B é uma matriz de ordem $K \times L$ dos coeficientes das variáveis predeterminadas e U é uma matriz de ordem $T \times L$ dos termos aleatórios. Observe que estamos adotando notação contrária à usual em relação as matrizes acima: o primeiro índice indica a coluna e o segundo refere-se à linha da matriz.

Pósmultilplicando-se o sistema de equações (9), pela matriz inversa da matriz Γ, obtém-se a forma reduzida do modelo:

(10) $Y = X\, \Pi + V$

onde:

$\Pi = B\, \Gamma^{-1}$ ou $\Pi\, \Gamma = B$

e

$V = U\, \Gamma^{-1}$

onde Π e V são matrizes de ordens $K \times L$ e $T \times L$, respectivamente.

Admitamos que a primeira equação do sistema (9) exclua algumas variáveis endógenas e predeterminadas do modelo e que seja expressa por:

(11) $y_1 = Y_1\ \gamma_1 + X_1\ \beta_1 + u_1 = Z_1\ \delta_1 + u_1$

onde y_1 é um vetor coluna com T elementos, Y_1 é uma matriz de ordem $T \times L_1 - 1$, γ_1 um vetor $L_1 - 1 \times 1$, X_1 uma matriz de ordem $T \times K_1$, β_1 um vetor $K_1 \times 1$, u_1 um vetor $T \times 1$, $Z_1 = \left[Y_1 \vdots X_1 \right]$, $\delta_1' = \left[\gamma_1' \vdots \beta_1' \right]$. As matrizes Y, X, U, Γ e B do sistema de equações (9) estão relacionadas com os vetores e matrizes que aparecem em (11) através das seguintes partições:

$$Y = [y_1 \vdots Y_1 \vdots Y_2], X = [X_1 \vdots X_2], U = [u_1 \vdots U_2]$$

$$\Gamma = \begin{bmatrix} 1 & \Gamma_{21} \\ -\gamma_1 & \Gamma_{22} \\ 0 & \Gamma_{23} \end{bmatrix}, B = \begin{bmatrix} \beta_1 & B_{21} \\ 0 & B_{22} \end{bmatrix}$$

Usando-se esta notação a forma reduzida (10) passa, então, a ser escrita como:

(12) $$[y_1 \vdots Y_1 \vdots Y_2] = [X_1 \vdots X_2] \begin{bmatrix} \Pi_{11} & \Pi_{21} & \Pi_{31} \\ \Pi_{12} & \Pi_{22} & \Pi_{32} \end{bmatrix} + [v_1 \vdots V_1 \vdots V_2]$$

Como Π Γ = B, tem-se que:

$$\begin{bmatrix} \Pi_{11} \Pi_{21} \Pi_{31} \\ \Pi_{12} \Pi_{22} \Pi_{32} \end{bmatrix} \begin{bmatrix} 1 \Gamma_{21} \\ -\gamma_1 \Gamma_{22} \\ 0 \Gamma_{23} \end{bmatrix} = \begin{bmatrix} \beta_1 B_{21} \\ 0 B_{22} \end{bmatrix}$$

Efetuando-se as multiplicações indicadas acima chega-se ao seguinte resultado:

(13) $$\begin{bmatrix} \Pi_{11} - \Pi_{21} \gamma_1 \Pi_{11} \Gamma_{21} + \Pi_{21} \Gamma_{22} + \Pi_{31} \Gamma_{23} \\ \Pi_{12} - \Pi_{22} \gamma_1 \Pi_{12} \Gamma_{21} + \Pi_{22} \Gamma_{22} + \Pi_{32} \Gamma_{23} \end{bmatrix} = \begin{bmatrix} \beta_1 & B_{21} \\ 0 & B_{22} \end{bmatrix}$$

Comparando-se as primeiras colunas das duas matrizes acima conclui-se que:

(14) $$\begin{cases} \Pi_{11} = \Pi_{21} \gamma_1 + \beta_1 \\ \Pi_{12} = \Pi_{22} \gamma_1 \end{cases}$$

Estas equações podem ser escritas como:

(15) $Q_1 \; \delta_1 = \pi_1$

onde:

$$\pi_1 = \begin{bmatrix} \Pi_{11} \\ \Pi_{12} \end{bmatrix}, Q_1 = \begin{bmatrix} \Pi_{21} & I \\ \Pi_{22} & 0 \end{bmatrix}$$

Observe-se que π_1 é um vetor com *K* elementos, Π_{21} e Π_{22} são matrizes de ordem $K_1 \; x \; (L_1\text{-}1)$ e $K_2 \; x \; (L_1 - 1)$, respectivamente. A matriz Q_1 é de ordem $K \; x \; (L_1 - 1 + K_1)$. A

condição necessária e suficiente para identificação da primeira equação do modelo, como vimos anteriormente, é que o posto da matriz Q_1 seja igual $L_1 - 1 + K_1$. Esta condição é equivalente a que o posto da matriz $\left[\Pi_{12} \vdots \Pi_{22}\right]$ seja igual ao posto da matriz Π_{22} e igual a $L_1 - 1$. Isto é:

$$(16) \qquad \rho\left[\Pi_{12} \vdots \Pi_{22}\right] = \rho(\Pi_{22}) = L_1 - 1$$

Considere a matriz Δ* de ordem $(K_2 + L - L_1) \times L$, definida por:

$$\Delta^* = \begin{bmatrix} 0 & B_{22} \\ 0 & \Gamma_{23} \end{bmatrix}$$

cujo posto é,obviamente, igual ao posto da matriz

$$\Delta = \begin{bmatrix} B_{22} \\ \Gamma_{23} \end{bmatrix}$$

A matriz Δ é uma matriz de ordem $(K_2 + L - L_1) \times (L-1)$. Cabe salientar que a matriz Δ é formada pelos coeficientes das variáveis endógenas predeterminadas e endógenas excluídas da equação, que se deseja saber se é ou não identificada, e que entram nas demais equações do modelo.

A matriz Δ* é igual ao produto das seguintes matrizes:

$$\Delta^* = \begin{bmatrix} \Pi_{12}\Pi_{22}\Pi_{32} \\ \\ 0\,0\,1 \end{bmatrix} \begin{bmatrix} 1 & \Gamma_{21} \\ -\gamma_1 & \Gamma_{22} \\ 0 & \Gamma_{23} \end{bmatrix} = \begin{bmatrix} 0 & B_{22} \\ 0 & \Gamma_{23} \end{bmatrix}$$

É fácil verificar-se este resultado com o auxílio de (13).

Pós multiplicando-se a matriz Δ* pela matriz inversa de Γ resulta:

$$\Delta^*\Gamma^{-1} = \begin{bmatrix} \Pi_{12}\Pi_{22}\Pi_{32} \\ \\ 0\,0\,I \end{bmatrix}$$

Como o posto da matriz $\Delta^* \Gamma^{-1}$ é igual ao posto da matriz Δ*, e como o posto de $\Delta^* \Gamma^{-1}$, segundo a expressão anterior é igual ao posto da submatriz $\left[\Pi_{12} \vdots \Pi_{22}\right]$ adicionado ao posto da matriz identidade I, segue-se, então que:

$$\rho(\Delta^*) = \rho(\Delta^*\Gamma^{-1}) = \rho\left(\left[\Pi_{12} \vdots \Pi_{22}\right]\right) + L - L_1$$

Em virtude de (16) concluímos que o posto da matriz Δ é igual a:

$$\rho(\Delta) = \rho(\Delta^*) = L - 1$$

Isto é, o posto da matriz Δ deve ser igual ao número de variáveis endógenas do modelo, menos um, para que a primeira equação estrutural seja identificada.

Observe-se que a matriz Δ tem $(K_2 + L - L_1)$ linhas e $(L-1)$. colunas. Quando a equação for identificada e o número de linhas de Δ for maior que seu número de colunas, $K_2 > L_1 - 1$, a equação é dita superidentificada. Quando a equação for identificada e a matriz Δ tiver um número de colunas igual ao número de linhas a equação é dita exatamente identificada.

Aplicação da Regra ao Modelo de Oferta e Procura

A matriz Δ correspondente à equação de demanda do modelo de mercado (4) reduz-se a um vetor coluna formado pelos coeficientes β_{23} e β_{24}, tendo em vista que as duas variáveis endógenas do modelo são incluídas na equação. Assim, temos que:

$$\Delta = \begin{bmatrix} \beta_{23} \\ \beta_{24} \end{bmatrix}$$

O posto de Δ deve ser igual a $L-1 = 2-1 = 1$ para que a equação de demanda seja identificada. Cabe salientar que se ambos os coeficientes, β_{23} e β_{24}, forem diferentes de zero, a equação de demanda é superidentificada enquanto que se um destes coeficientes for nulo a equação é exatamente identificada.

No caso da equação de oferta a matriz Δ é igual a:

$$\Delta = \begin{bmatrix} \beta_{11} \\ \beta_{12} \end{bmatrix}$$

pois não existe variável endógena excluída da equação de oferta. O posto de Δ será igual a um se pelo menos um dos coeficientes , β_{11} e β_{12}, for diferente de zero. Quando ambos os coeficientes forem nulos a equação de oferta não será identificada. A equação de oferta será superidentificada se ambos os coeficientes forem diferentes de zero e será exatamente identificada se um desses coeficientes for nulo.

É interessante observar que em um modelo com apenas duas variáveis endógenas, a exclusão de uma variável predeterminada de cada equação do modelo garante a sua identificação.

1.3 - Estimação de Parâmetros Estruturais

A literatura econométrica contém um bom número de diferentes estimadores para os parâmetros das equações estruturais de um modelo de equações simultâneas. Quanto à informação que utilizam, estes estimadores podem ser classificados em duas classes: os de informação limitada e os estimadores de informação completa. Os estimadores de informação limitada levam em conta apenas a informação que diz respeito a uma equação particular do modelo, enquanto os estimadores de informação completa incorporam toda a informação contida no mesmo.

Apresentaremos a seguir, três estimadores de informação limitada: mínimos quadrados indireto generalizado, mínimos quadrados em duas etapas e máxima verossimilhança.

O estimador de mínimos quadrados em duas etapas é bastante popular devido à sua fórmula simples comparada com outros estimadores. No que toca ao estimador indireto generalizado de mínimos quadrados a maioria dos livros textos de econometria afirmam, erroneamente, que este estimador não pode ser aplicado a equações superidentificadas pois gerariam estimativas que não são únicas. Discutiremos, também, nesta subseção os problemas que surgem quando se usa o método de mínimos quadrados ordinários na estimação dos parâmetros de uma equação estrutural. Veremos ainda a aplicabilidade do método de mínimos quadrados ordinários quando o modelo de equações simultâneas for de um tipo particular, o chamado modelo recursivo.

<u>Inconsistência dos Estimadores de Mínimos Quadrados</u>

A equação da forma reduzida do modelo de oferta e procura que exprime o preço do produto como função das variáveis predeterminadas, de acordo com (5), é dada por:

$$p_t = \pi_{21}\ y_t + \pi_{22}\ S_t + \pi_{23}\ \omega_t + \pi_{24}\ p_{t-1} + \frac{u_{2t} - u_{1t}}{\gamma_{12} - \gamma_{22}}$$

A partir desta relação é fácil se constatar que a covariância entre o preço p_t e o termo aleatório u_{1t} é diferente de zero pois

$$E\,p_t u_{1t} = E\left[\left(\frac{u_{2t}-u_{1t}}{\gamma_{12}-\gamma_{22}}\right)u_{1t}\right] = \frac{\sigma_{21}-\sigma_{11}}{\gamma_{12}-\gamma_{22}}$$

Na obtenção deste resultado levamos em conta que, por hipótese. as variáveis predeterminadas não estão correlacionadas com u_{1t} . A expressão anterior mostra claramente que o preço p_t e o termo aleatório u_{1t} estão correlacionados. Este fato implica em que os estimadores de mínimos quadrados ordinários dos parâmetros da equação de demanda (1), repetida aqui por conveniência,

$$q_t = \gamma_{12}\ p_t + \beta_{11}\ y_t + \beta_{12}\ S_t + u_{1t,}$$

serão tendenciosos e inconsistentes. O mesmo fato ocorre com os estimadores de mínimos quadrados ordinários dos parâmetros da equação de oferta em virtude da covariância existente entre o preço p_t e o termo aleatório u_{2t}:

$$E\,p_t u_{2t} = E\left[\left(\frac{u_{2t}-u_{1t}}{\gamma_{12}-\gamma_{22}}\right)u_{2t}\right] = \frac{\sigma_{22}-\sigma_{12}}{\gamma_{12}-\gamma_{22}}$$

A conclusão a que se chega quanto à estimação dos parâmetros de uma equação estrutural, através do método de mínimos quadrados ordinários, é de que este método não é, em geral, recomendável face à tendenciosidade e à inconsistência dos estimadores daí resultante.

<u>Modelos Recursivos</u>

Suponha agora que a quantidade ofertada independa do preço p_t do produto no período *t* mas que ainda seja função do preço ω_t do insumo e do preço p_{t-1} do produto com um período de defasagem. Isto é:

(17) $$q_t = \beta_{23}\ \omega_t + \beta_{24}\ p_{t-1} + u_{2t}$$

A equação de demanda de mercado continua sendo dada por (1), porém, preferimos agora escrevê-la de maneira ligeiramente diferente, com preço p_t do produto no lado esquerdo da equação:

(18) $$p_t = \frac{1}{\gamma_{12}} q_t - \frac{\beta_{11}}{\gamma_{12}} \gamma_t - \frac{\beta_{12}}{\gamma_{12}} S_t - \frac{u_{1t}}{\gamma_{12}}$$

Admita também a hipótese de que os termos aleatórios u_{1t} e u_{2t} não estejam correlacionados. As variáveis endógenas do modelo de mercado continuam sendo o preço p_t e a quantidade q_t. As variáveis predeterminadas, como antes, são o preço ω_t do insumo, o preço p_{t-1} do bem defasado de um período, a renda y_t e o preço S_t do bem substituto.

A equação de oferta (17) além de estar na forma estrutural está, também, na forma reduzida, pois o lado direito daquela equação só contém variáveis predeterminadas. Logo, a aplicação de mínimos quadrados ordinários conduz a estimadores não tendenciosos dos seus parâmetros.

O termo aleatório u_{2t}, por hipótese, não está correlacionado com o termo aleatório u_{1t}. Segue-se, então, que a quantidade q_t não está correlacionada com o termo u_{1t} . Consequentemente, a aplicação de mínimos quadrados ordinários na estimação dos parâmetros da equação de demanda conduz a estimadores não tendenciosos dos seus parâmetros. Observe que neste caso é importante prestar atenção para que a regressão seja feita do preço p_t contra as variáveis q_t, y_t e S_t. A ordem das variáveis nos modelos recursivos é bastante importante e daí porque escrevemos a equação de demanda como em (18).

O sistema de equações formado por (17) e (18) é denominado recursivo em virtude do modelo ser resolvido recursivamente: a primeira equação fornece o valor da quantidade q_t e, conhecido este valor, a segunda equação nos dá o valor do preço p_t. Em notação matricial, o modelo de mercado expresso por (17) e (18) pode ser escrito como:

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{\gamma_{12}} & 1 \end{bmatrix} \begin{bmatrix} q_t \\ p_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & \beta_{23} & \beta_{24} \\ -\frac{\beta_{11}}{\gamma_{12}} & -\frac{\beta_{12}}{\gamma_{12}} & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ S_t \\ \omega_t \\ p_{t-1} \end{bmatrix} + \begin{bmatrix} u_{2t} \\ -\frac{u_{1t}}{\gamma_{12}} \end{bmatrix}$$

Duas propriedades caracterizam um modelo recursivo. A primeira é que a matriz dos coeficientes que multiplica o vetor das variáveis endógenas tem um formato particular, qual seja, a de ser uma matriz triangular, que, no nosso exemplo, é igual a:

$$\begin{bmatrix} 1 & 0 \\ -\dfrac{1}{\gamma_{12}} & 1 \end{bmatrix}$$

A segunda propriedade que caracteriza um modelo como recursivo é que a matriz de variância dos termos aleatórios é uma matriz diagonal. No exemplo do modelo de mercado, a hipótese de que $E\, u_{2t}\, u_{1t} = 0$ implica que:

$$E\begin{bmatrix} u_{2t}^2 & \dfrac{u_{2t}u_{1t}}{\gamma_{12}} \\ \dfrac{u_{2t}u_{1t}}{\gamma_{12}} & \dfrac{u_{1t}^2}{\gamma_{12}} \end{bmatrix} = \begin{bmatrix} \sigma_{22} & 0 \\ 0 & \dfrac{\sigma_{11}}{\gamma_{12}} \end{bmatrix}$$

Vale ressaltar que estas duas propriedades mencionadas asseguram a consistência dos estimadores de mínimos quadrados ordinários dos parâmetros das equções de um modelo recursivo.

<u>Estimador de Mínimos Quadrados Indireto Generalizado</u>

A expressão (15), repetida aqui por conveniência,

$$Q_1\ \delta_1 = \pi_1$$

contém um sistema de K equações lineares com $(L_1 + K_1 - 1)$ incógnitas, os elementos do vetor δ_1. A solução desta última é dada por:

$$\delta_1 = (Q_1 ' Q_1)^{-1}\ Q_1 '\ \pi_1$$

O estimador de mínimos quadrados indireto generalizado $\hat{\delta}_1$ é obtido quando substituímos os valores de Q_1 e π_1 na expressão acima pelos estimadores $\hat{Q}_1$ *e* $\hat{\pi}_1$ da forma reduzida. Isto é:

$$\hat{\delta}_1 = (\hat{Q}_1 ' \hat{Q}_1)^{-1}\ \hat{Q}_1 '\ \hat{\pi}_1$$

onde a matriz $\hat{Q}_1$ é igual a

(19) $\hat{Q}_1 = \left[\hat{\Pi}_2 \ \vdots\ D\right]$,

e os estimadores $\hat{\Pi}_2$ *e* $\hat{\pi}_1$ da forma reduzida são:

(20) $\hat{\Pi}_2 = (X'\,X)^{-1}\ X'\,Y_1$

(21) $\hat{\pi}_1 = (X'\,X)^{-1}\ X'\,y_1$

e a matriz D é dada por:

$$D = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Denominamos o estimador $\hat{\delta}_1$ de estimador de mínimos quadrados indireto generalizado por dois motivos. Em primeiro lugar, porque este estimador é obtido indiretamente através dos estimadores $\hat{\Pi}_2$ *e* $\hat{\pi}_1$. Em segundo lugar porque, no caso particular de a equação estrutural ser exatamente identificada, $\hat{\delta}_1$ reduz-se ao tradicional estimador indireto de mínimos quadrados:

$$\hat{\delta}_1 = \hat{Q}_1^{-1}\ \hat{\pi}_1$$

Cuidemos, agora, de obter expressões algébricas para o estimador $\hat{\delta}_1$ em função das matrizes de observações X e $[y_1 \vdots Y_1]$. Usamos a relação (19) para escrever o estimador $\hat{\delta}_1$ da seguinte forma:

$$\hat{\delta}_1 = \begin{bmatrix} \hat{\Pi}_2{}'D & \hat{\Pi}_2{}'D \\ D'\hat{\Pi}_2 & D'D \end{bmatrix}^{-1} \begin{bmatrix} \hat{\Pi}'_2 \\ D' \end{bmatrix} \hat{\pi}_1$$

Em seguida, substituímos os valores de $\hat{\Pi}_2$ *e* $\hat{\pi}_1$ dados em (20) e (21) na expressão anterior e obtemos:

$$(22) \qquad \hat{\delta}_1 = \begin{bmatrix} Y_1^{'}X(X'X)^{-2}X'Y_1 & Y_1^{'}X(X'X)^{-2}X'X_1 \\ X_1^{'}X(X'X)^{-2}X'Y_1 & I \end{bmatrix}^{-1}$$

$$\begin{bmatrix} Y_1^{'}X(X'X)^{-2}X' \\ X_1^{'}X(X'X)^{-2}X' \end{bmatrix} y_1$$

Para se chegar a essa expressão levamos em conta que $D'\,\Pi_2 = X_1^{'}\ X\ (X'X)^{-2}\ X'Y_1$ A demonstração desta igualdade é bastante simples:

$$D'\,\hat{\Pi}_2 = D'\,(X'X)^{-1}\ X'Y_1 = D'\,X'\,X\ (X'X)^{-1}\ (X'X)^{-1}\ X'Y_1 = X_1^{'}\ X\ (X'X)^{-2}\ X'Y_1$$

tendo em vista que $D'X' = X_1^{'}$. De maneira similar calcula-se o produto $D'\,\hat{\Pi}_1$.

Usando-se uma notação mais simples a expressão (22) pode ser escrita como:

$$(23) \qquad \hat{\delta}_1 = \left[Z_1^{'}X(X'X)^{-2}X'Z_1\right]^{-1} Z_1^{'}X(X'X)^{-2}X'y_1$$

levando-se em conta que:

$$X_1^{'}\ X\ (X'X)^{-2}\ X'X_1 = D'\,X'\,X\ (X'X)^{-2}\ X'XD = I$$

pois $D'D = I$.

Propriedades do Estimador $\hat{\delta}_1$

Substituímos o valor de y_1, dado pelo lado direito do segundo sinal de igualdade de (10), em (23), obtemos:

$$(24) \qquad \hat{\delta}_1 = \delta_1 + \left[Z_1' \; X \; (X'X)^{-2} \; X'Z_1 \right]^{-1} Z_1' \; X \; (X'X)^{-2} \; X' u_1$$

De acordo com algumas hipóteses tradicionais no estudo dos estimadores dos parâmetros de um sistema de equações simultâneas temos que[1]

$$\text{limp} \left[\frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} \frac{X'Z_1}{T} \right]^{-1} \quad \text{é finita, e}$$

$$\text{limp} \left[\frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} \frac{X'u_1}{T} \right] \quad \text{é igual a zero.}$$

A notação *limp* denota o limite em probabilidade da variável indicada. Aplicando-se estes resultados a (24) concluímos que:

$$\text{limp} \;\; \hat{\delta}_1 = \delta_1$$

o que significa dizer que o estimador indireto generalizado de mínimos quadrados é consistente.

A distribuição assintótica da variável aleatória $\sqrt{T}\;(\hat{\delta}_1 - \delta_1)$ pode ser obtida com o seguinte procedimento da expressão (24) temos que:

$$\sqrt{T}(\hat{\delta}_1 - \delta_1) = \left[\frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} \frac{X'Z_1}{T} \right]^{-1} \frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} \frac{X'u_1}{\sqrt{T}}$$

Por outro lado, as hipóteses a que nos referimos abaixo da expressão (24) nos possibilita afirmar que:

$$\text{limp} \left[\frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} \frac{X'Z_1}{T} \right]^{-1} \frac{Z_1'X}{T} \left(\frac{X'X}{T} \right)^{-2} = G$$

onde G é uma matriz cujos elementos são finitos. Em seguida, lançamos mão do teorema que afirma que a distribuição assintótica de $X'u_1 \,/\, \sqrt{T}$ é normal com valor esperado zero e

[1]Essas hipóteses estão listadas em H.Theil (1971), Principles of Econometrics, John Wiley, capítulo 10.

matriz de variância-covariância igual a σ_{11} limp $\left(\frac{X'X}{T}\right)$. Podemos, então, concluir que a distribuição assintótica de é $\sqrt{T}\ (\hat{\delta}_1 - \delta_1)$ é normal com média zero e matriz de variância-covariância, Var $\hat{\delta}_1$, igual a:

$$(25) \qquad \sigma_{11} G \left[\text{limp} \left(\frac{X'X}{T} \right) \right] G' = Var \hat{\delta}_1$$

A estatística

$$(26) \qquad s_{11} = \frac{1}{T} (y_1 - Z_1\ \hat{\delta}_1)'\ (y_1 - Z_1\ \hat{\delta}_1)$$

é um estimador consistente da variância σ_{11}. A prova dessa propriedade é bastante simples. Substituindo-se (10) na expressão anterior obtemos:

$$s_{11} = \frac{u_1'\ u_1}{T} - \frac{u_1'\ Z_1}{T} (\hat{\delta}_1 - \delta_1) - \frac{(\hat{\delta}_1 - \delta_1)'\ Z_1'\ u_1}{T} + (\hat{\delta}_1 - \delta_1)' \frac{Z_1'\ Z_1}{T} (\hat{\delta}_1 - \delta_1)$$

A partir desta relação concluímos que:

$$\text{limp}\ \ s_{11} = \text{limp}\ \frac{u_1'\ u_1}{T} = \sigma_{11}$$

baseados no fato de que limp $(\hat{\delta}_1 - \delta_1) = 0$ e de que limp $u_1'\ u_1 / T = \sigma_{11}$.

É fácil derivar-se a partir de (25) e (26) os erros padrões (assintóticos) da estimativa do estimador indireto generalizado de mínimos quadrados. Isto é, a partir destas expressões concluímos que os erros-padrões da estimativa dos elementos do vetor $\hat{\delta}_1$ são dados pelas raízes quadradas dos elementos da diagonal principal da matriz

$$s_{11} = \left[Z_1'\ X\ (X'X)^{-2}\ X'Z_1 \right]^{-1} Z_1'\ X\ (X'X)^{-3}\ X'Z_1 \left[Z_1'\ X\ (X'X)^{-2}\ X'Z_1 \right]^{-1}$$

Mínimos Quadrados Em Duas Etapas

O estimador de mínimos quadrados em duas etapas pode ser derivado de diferentes maneiras. A seguir derivaremos este estimador com o auxílio do enfoque algébrico. da forma reduzida do modelo, expresso em (11), temos que

$$y_1 = [X_1 \vdots X_2] \begin{bmatrix} \Pi_{11} \\ \Pi_{12} \end{bmatrix} + v_1$$

Por outro lado, da equação (14) sabemos que:

$$\begin{bmatrix} \Pi_{11} \\ \Pi_{12} \end{bmatrix} = \begin{bmatrix} \Pi_{21} & I \\ \Pi_{22} & 0 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \beta_1 \end{bmatrix}$$

Substituindo-se esta relação na expressão anterior concluímos que:

$$y_1 = [X_1 \vdots X_2] \begin{bmatrix} \Pi_{21} & I \\ \Pi_{22} & 0 \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \beta_1 \end{bmatrix} + v_1$$

Se os valores de Π_{21} e Π_{22} fossem conhecidos, poderíamos aplicar o método de mínimos quadrados ordinários à equação anterior. O estimador assim obtido seria dado por:

$$d_1 = \left[Q_1' \; X' X \; Q_1 \right]^{-1} \quad Q_1' \; X' \, y_1$$

Entretanto, devido ao fato de que Π_{21} e Π_{22} serem parâmetros que têm de ser estimados, visto não serem conhecidos <u>a priori</u>, o estimador acima não pode ser aplicado na prática. Todavia, o problema é facilmente contornado usando-se a matriz

$$\hat{Q}_1 = \begin{bmatrix} \hat{\Pi}_2 & I \\ \hat{\Pi}_{22} & 0 \end{bmatrix}$$

ao invés da matriz Q_1, onde $\hat{\Pi}_{21}$ *e* $\hat{\Pi}_{22}$ são os estimadores de mínimos quadrados da forma reduzida do modelo. Dessa maneira, obtemos o estimador de mínimos quadrados em duas etapas:

$$\tilde{\delta}_1 = \left[\hat{Q}_1 \,' \, X' X \; \hat{Q}_1 \right]^{-1} \hat{Q}_1 \,' \; X' \, y_1$$

Usamos a relação (19) para escrever este estimador na seguinte forma:

$$\tilde{\delta}_1 = \begin{bmatrix} \hat{\Pi}_2 ' X' X \hat{\Pi}_2 & \hat{\Pi}_2 ' X' X D \\ D' X' X \hat{\Pi}_2 & D' X' X D \end{bmatrix}^{-1} \begin{bmatrix} \hat{\Pi}_2 X' \\ D' X' \end{bmatrix} Y_1$$

Em seguida, substituindo-se o valor de Π_2 dado por (20) na expressão acima resulta:

$$\tilde{\delta}_1 = \begin{bmatrix} Y_1' X (X' X)^{-1} X' Y_1 & Y_1' X_1 \\ X_1' Y_1 & X_1' X_1 \end{bmatrix}^{-1} \begin{bmatrix} Y_1' X (X' X)^{-1} X' \\ X_1' \end{bmatrix} y_1$$

Alternativamente, essa expressão pode ser escrita de maneira mais compacta do seguinte modo:

$$\tilde{\delta}_1 = \left[Z_1 ' X \; (X' X)^{-1} \; X' Z_1 \right]^{-1} Z_1 ' X \; (X' X)^{-1} \; X' \, y_1$$

O enfoque que adotamos na derivação do estimador de mínimos quadrados em duas etapas não mostra o porque do seu nome. Todavia, não é difícil entender qual a origem da palavra duas etapas no nome do estimador. Com efeito, observe que o valor previsto de Y_1

que denominamos por $\hat{Y}_1$, é igual ao produto da matriz X vezes a estimativa de mínimos quadrados ordinários, $\hat{\Pi}_2$, da forma reduzida:

$$\hat{Y}_1 = X\ \hat{\Pi}_2 = X\ (X'X)^{-1}\ X'Y_1$$

Como,

$$\hat{Y}_1'\ \hat{Y}_1 = Y_1'\ X\ (X'X)^{-1}\ X'X\ (X'X)^{-1}\ X'Y_1 = Y_1'\ X\ (X'X)^{-1}\ X'Y_1$$

e:

$$\hat{Y}_1'\ X_1 = Y_1'\ X (X'X)^{-1}\ X'\ X_1 = Y_1'\ X\ (X'X)^{-1}\ X'\ XD = Y_1'\ X\ D = Y_1'\ X_1$$

Segue-se que:

$$\tilde{\delta}_1 = \begin{bmatrix} \hat{Y}_1'\hat{Y}_1\hat{Y}_1'X_1 \\ \\ X_1'\hat{Y}_1X_1'X_1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_1' \\ \\ X_1' \end{bmatrix} y_1$$

Logo, a estimativa $\tilde{\delta}_1$ é também obtida pelo seguinte procedimento em duas etapas:

1ª etapa: a partir da regressão de Y_1
contra X obtém-se $\hat{Y}_1$;

2º etapa: faz-se, então, a regressão de
y_1 contra $\hat{Y}_1$ e X_1 .

Pode-se provar, de maneira análoga ao que foi feito para o estimador indireto generalizado de mínimos quadrados, que o estimador de mínimos quadrados em duas estapas é consistente, limp $\tilde{\delta}_i = \delta_1$, e que $\sqrt{T}\ (\tilde{\delta}_1 - \delta_1)$ tem uma distribuição asintótica normal com média zero e matriz de variância-covariância igual a:

$$\sigma_{11} \text{limp} \left[\left(\frac{Z_1'X}{T} \right) \left(\frac{X'X}{T} \right)^{-1} \left(\frac{X'Z_1}{T} \right) \right]^{-1}$$

Portanto, os erros padrões da estimativa dos elementos do vetor $\tilde{\delta}_1$ são dados pelas raízes quadradas dos elementos da diagonal principal da matriz:

$$s_{11} \left[Z_1'X\ (X'X)^{-1}\ X'Z_1 \right]^{-1}$$

A interpretação que acabamos de descrever do estimador de mínimos quadrados em duas etapas é bastante sugestiva, mas não deve ser levada ao pé da letra, sob pena de se cometer um erro importante no cálculo dos erros padrões das estimativas dos parâmetros do modelo. Com efeito, suponhamos que um pesquisador não dispondo de um pacote de

computação que calculasse diretamente a estimativa de mínimos quadrados em duas etapas resolvesse aplicar, em duas etapas, mínimos quadrados ordinários. Nestas circunstâncias a estimativa da variância σ_{11} seria fornecida pela seguinte expressão:

$$\tilde{\sigma}_{11} = \frac{\tilde{u}_1 ' \tilde{u}_1}{T}$$

onde $\tilde{u}_1 = y_1 - \hat{Y}_1 \tilde{\gamma}_1 - X_1 \tilde{\beta}_1$.

Como $Y_1 = \hat{Y}_1 + \hat{V}_1$, segue-se então que:

$$\tilde{u}_1 = \hat{u}_1 + \hat{V}_1 \tilde{\gamma}_1$$

onde $\hat{u}_1 = y_1 - Y_1 \tilde{\gamma}_1 - X_1 \tilde{\beta}_1$.

Com a finalidade de examinarmos se o estimador $\tilde{\sigma}_{11}$ é consistente procedermos do seguinte modo. O limite em probabilidade de $\tilde{\sigma}_{11}$ é igual a:

$$\text{limp } \tilde{\sigma}_{11} = \text{limp } \frac{\tilde{u}_1 ' \tilde{u}_1}{T} = \lim \frac{(\hat{u}_1 + \hat{V}_1 \tilde{\gamma}_1)' (\hat{u}_1 + \hat{V}_1 \tilde{\gamma}_1)}{T}$$

Alternativamente:

$$\text{limp } \tilde{\sigma}_{11} = \text{limp} \frac{\hat{u}_1 ' \hat{u}_1}{T} + \text{limp } 2 \frac{\hat{u}_1 ' \hat{v}_1 \tilde{\gamma}_1}{T} + \text{limp } \frac{\tilde{\gamma}_1 ' \hat{V}_1 ' \hat{V}_1 \tilde{\gamma}_1}{T} \tag{27}$$

Como

$$\text{limp } \frac{\hat{u}_1 ' \hat{u}_1}{T} = \sigma_{11}$$

e provaremos em seguida que:

$$\text{limp } 2 \frac{\hat{u}_1 ' \hat{V}_1 \tilde{\gamma}_1}{T} = (\text{limp } 2 \frac{u_1 ' V_1}{T}) \; \gamma_1 = \text{limp } \frac{\tilde{\gamma}_1 ' \hat{V}_1 ' \hat{V}_1 \tilde{\gamma}_1}{T} = \gamma_1 ' \text{ limp } \frac{V_1 ' V_1}{T} \; \gamma_1 =$$

segue-se, então, que em geral:

$$\text{limp } \tilde{\sigma}_{11} \neq \sigma_{11}$$

A matriz de resíduos $\hat{V}_1$ é igual a:

$$\hat{V}_1 = Y_1 - \hat{Y}_1 = Y_1 - X \; (X ' X)^{-1} \; X ' Y_1 = \left[I - X (X ' X)^{-1} \; X ' \right] Y_1$$

Como $Y_1 = X \, \Pi_2 + V_1$, a matriz $\hat{V}_1$ pode ser expressa por:

$$\hat{V}_1 = \left[I - X\ (X'X)^{-1}\ X \right] V_1 = M\ V_1$$

Logo,

$$\text{limp}\ \frac{\hat{V}_1'\hat{V}_1}{T} = \text{limp}\ \frac{V_1'\ M'\ M\ V_1}{T} = \text{limp}\ \frac{V'\ M\ V_1}{T}$$

ou ainda:

$$\text{limp}\frac{\hat{V}_1'\hat{V}_1}{T} = \text{limp}\frac{V_1'V_1}{T} - \text{limp}\frac{V_1'X}{T}\left(\frac{X'X}{T}\right)^{-1}\frac{X'V_1}{T}$$

Em virtude de

$$\text{limp}\ \frac{V_1'\ X}{T} = 0$$

pois VΓ = U e limp $\frac{U'\ X}{T}$ = 0, segue-se que:

$$\text{limp}\ \frac{\hat{V}_1'\hat{V}_1}{T} = \text{limp}\ \frac{\mathrm{V}_1'V_1}{T}$$

Por outro lado,

$$\text{limp}\ \frac{\hat{u}_1'\hat{V}_1}{T} = \text{limp}\ \frac{U_1'\ N'\ M\ V_1}{T}$$

onde $\hat{u}_1 = y_1 - Z_1\ \hat{\delta}_1 = N\ u_1$, e

$$N = I - Z_1 \left[Z_1'\ X\ (X'X)^{-1}\ X'Z_1 \right]^{-1} Z_1'\ X\ (X'X)^{-1}\ X'$$

Com um pouco de paciência e álgebra pode-se, então, mostrar que:

$$\text{limp}\ \frac{u_1'\ N'\ M\ V_1}{T} = \text{limp}\ \frac{u_1'V_1}{T}$$

Utilizando-se a partição indicada antes da expressão (11), obtém-se:

$$v_1 - V_1\ \gamma_1 = u_1$$

Consequentemente:

$$\text{limp}\ \frac{u_1'V_1}{T} = \text{limp}\ \frac{(\mathrm{v}_1 - V_1\ \gamma_1)'V_1}{T} = \text{limp}\ \frac{v_1'V_1}{T} - \text{limp}\ \frac{\gamma_1'V_1'V_1}{T}$$

Substituindo-se este resultado em (27), conclui-se que:

$$\text{limp } \tilde{\sigma}_{11} = \sigma_{11} + 2\,(\text{limp } \frac{v_1{}'V_1}{T})\ \gamma_1 - \gamma_1{}'\,(\text{limp } \frac{V_1{}'V_1}{T})\ \gamma_1$$

Admitindo-se que no modelo de equações simultâneas a matriz de variância-covariância dos termos estocásticos da matriz U da forma estrutural é dada por:

$$E\ U'\ U = \Sigma$$

a matriz de variância-covariância da forma reduzida é igual a:

$$E\ V'V = E(\Gamma^{-1})\,'U'U\ \Gamma^{-1} = (\Gamma^{-1})'\ \Sigma\ \Gamma^{-1} = \Omega$$

Logo, se a matriz for particionada como em (11) segue-se que:

$$E\begin{bmatrix} v_1{}' \\ V_1{}' \\ V_2{}' \end{bmatrix}\begin{bmatrix} v_1 \vdots V_2 \vdots V_2 \end{bmatrix} = E\begin{bmatrix} v_1{}'v_1 & v_1{}'V_1 & v_1{}'V_2 \\ V_1{}'v_1 & V_1{}'V_1 & V_1{}'V_2 \\ V_2{}'v_1 & V_2{}'V_1 & V_2{}'V_2 \end{bmatrix}$$

que é igual a:

$$EV'V = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix}$$

Portanto:

$$\text{limp } \frac{v_1{}'V_1}{T} = \Omega_{12}$$

e

$$\text{limp } \frac{V_1{}'V_1}{T} = \Omega_{22}$$

e a inconsistência do estimador $\hat{\sigma}_{11}$ está demonstrada. Isto é:

$$\text{limp } \tilde{\sigma}_{11} = \tilde{\sigma}_{11} + 2\ \Omega_{12}\ \gamma_1 - \gamma_1{}'\ \Omega_{22}\ \gamma_1$$

Máxima Verossimilhança: Informação Limitada

O estimador de máxima verossimilhança de informação limitada dos parâmetros de uma equação estrutural do modelo de equações simultâneas na sua concepção é bastante simples, pois consiste na aplicação direta do método de máxima verossimilhança. Entretanto, sua obtenção requer um pouco de álgebra. Com a finalidade de tornar a dedução desse estimador menos complicada comecemos por introduzir uma notação que facilite as

manipulações algébricas. A equação estrutural (10), que deseja estimar, passa a ser então escrita como:

$$[y_1 \vdots Y_1]\begin{bmatrix} 1 \\ -\gamma_1 \end{bmatrix} = X_1 \beta_1 + u_1$$

ou:

(28) $Y_c \ \gamma = X_1 \ \beta_1 + u_1$

onde:

$$Y_c = [y_1 \vdots Y_1] \text{ e } \gamma = \begin{bmatrix} 1 \\ -\gamma_1 \end{bmatrix}$$

A equação (11), da forma reduzida do modelo de equações simultâneas,permite escrever[2]

(29) $Y_c = X_1 \ \Pi_1 + X_2 \ \Pi_2 + V_c$

onde:

$$\Pi_1 = [\Pi_{11} \ \Pi_{21}]$$

$$\Pi_2 = [\Pi_{12} \ \Pi_{22}]$$

$$V_c = [v_1 \ \ V_1]$$

As restrições (14) repetidas aqui por conveniência,

$$\Pi_{11} = \Pi_{21} \ \gamma_1 + \beta_1$$
$$\Pi_{12} = \Pi_{22} \ \gamma_1$$

podem ser reescritas na nova notação do seguinte modo:

(30) $\Pi_1 \ \gamma = \beta_1$

(31) $\Pi_2 \ \gamma = 0$

A matriz V_C dos termos estocásticos da forma reduzida segue uma distribuição normal com média igual a 0 e variância-covariância igual a Ω_C, cuja função de densidade de probabilidade é igual a:

[2]Mudamos, por conveniência, a notação adotada até aqui: o símbolo Π_2 representa agora outra matriz.

$$p\,(V_c) = (2\Pi)^{-\frac{L_1 T}{2}}\ \left|\Omega_c\right|^{-T/2}\ \exp\,(-\frac{1}{2}\,tr\,V_c\,'V_c\ \Omega_c^{-1})$$

onde exp( ) indica o número natural e elevado ao termo entre parênteses e o símbolo tr representa o traço da matriz.

Como o Jacobiano da transformação de V_c para Y_c é igual a 1, o logaritmo da função de densidade de probabilidade de Y_c é igual a:

(32)
$$\log \mathrm{p}\,(Y_c\,/\,\Pi_1\,,\Pi_2\,,\Omega_c\,,\,X\,) = -\frac{L_1\ T}{2}\log 2\Pi - \frac{T}{2}\log\left|\Omega_c\right|$$
$$-\frac{1}{2}\,tr\,(Y_c - X_1\ \Pi_1 - X_2\,\Pi_2\,)'\,(Y_c - X_1\,\Pi_1 -\ X_2\ \Pi_2\,)\,\Omega_c^{-1}$$

Uma vez conhecidos os valores de Y_c esta função se transforma no logaritmo de função de verossimilhança, isto é:

(33) $$\ell = \log\ p\,(\Pi_1\,,\Pi_2\,,\Omega_c\,/\,Y_c\,,X\,)$$

A equação (30) não implica em nenhuma restrição sobre os elementos da matriz Π_1. Ela indica apenas como calcular o vetor β_1 quando Π_1 e γ forem conhecidos. O mesmo não ocorre com a equação (31), pois ela impõe restrições sobre os elementos da matriz Π_2. Consequentemente, os parâmetros do modelo devem ser estimados levando-se em conta a restrição (31). O problema consiste, então, em maximizar a função

$$\log\ p\,(\Pi_1\,,\Pi_2\,,\Omega_c\,/\,Y_c\,,\ X\,)$$

sujeito a restrição:

$$\Pi_2\ \ \gamma = 0$$

Antes de escrever a expressão de Lagrange deste problema de maximização condicionada é interessante observar que a restrição não involve as matrizes Π_1 e Ω_c. Este fato simplifica um pouco a solução do problema pois podemos derivar parcialmente o logaritmo da função de verossimilhança em relação a Π_1 e Ω_c^{-1}, igualá-las a zero, e obter os valores de Π_1 e Ω_c em função de Π_2, Y_c e X. Isto é:[3]

---

[3]Aplicamos aqui os seguintes resultados:

$$\frac{\partial tr}{\partial \Pi}(Y - X\ \Pi)'(Y - X\ \Pi)\ \Omega^{-1} = -2\ \Omega^{-1}\ (Y'\,X - \Pi'\,X\,'\,X\,)$$

$$\frac{\partial \log}{\partial \Omega_c^{-1}}\left|\Omega_c\right| = -\ \Omega_c$$

$$\frac{\partial tr}{\partial \Omega^{-1}}V_c\,'V_c\ \ \Omega^{-1} = V_c\,'V_c$$

e derivamos em relação a $\Omega_c{}^{-1}$, ao invés de Ω_c, pois a derivada é mais fácil e a propriedade de invariância do método de máxima verossimilhança permite que isto seja feito.

$$\frac{\partial \ell}{\partial \Pi_1} = -\frac{1}{2}(-2\Omega_c^{-1})\left[(Y_c - X_2\,\Pi_2)'X_1 - \Pi_1'Y_1'X_1\right] = 0$$

e

$$\frac{\partial \ell}{\partial \Omega_c^{-1}} = \frac{T}{2}\,\Omega_c - \frac{1}{2}V_c'V_c = 0$$

Portanto,

$$\tilde{\Pi}_1 = (X_1'X_1)^{-1}\,X_1'(Y_c - X_2\,\Pi_2)$$

e

$$\Omega_c = \frac{V_c'V_c}{T}$$

Como

$$V_c = Y_c - X_1\,\Pi_1 - X_2\,\Pi_2 = Y_c - X_1(X_1'X_1)^{-1}\,X_1'(Y_c - X_2\,\Pi_2) - X_2\,\Pi_2$$

$$= \left[I - X_1(X_1'X_1)^{-1}\,X_1'\right](Y_c - X_2\,\Pi_2) = M_1\,(Y_c - X_c\,\Pi_2)$$

onde $M_1 = I - X_1\,(X_1'X_1)^{-1}\,X_1'$, é uma matriz idempotente, segue-se que:

$$\Omega_c = \frac{V_c'V_c}{T} = \frac{(Y_c - X_2\,\Pi_2)'\,M_1(Y_c - X_2\,\Pi_2)}{T}$$

Substituindo-se esta expressão na função de verossimilhança obtém-se a função de verossimilhança concentrada:

$$\log \mathrm{p}\;(\Pi_2 / Y_c, X) = -\frac{L_1 T}{2}\log 2\Pi - \frac{T}{2}\log\left|\frac{(Y_c - X_2\,\Pi_2)'\,M_1\,(Y_c - X_2\,\Pi_2}{T}\right| - \frac{1}{2}\,tr\;I$$

Desprezando-se as constantes que aparecem nesta função, a maximização do log p $(\Pi_2/Y_c, X)$, sujeito à restrição $\Pi_2\,\gamma=0$, é equivalente ao seguinte problema:

minimizar

$$\log\left|(Y_c - X_2\,\Pi_2)'\,M_1\,(Y_c - X_2\,\Pi_2)\right|$$

com relação Π_2, sujeito à restrição:

$$\Pi_2\,\gamma = 0$$

A expressão de Lagrange deste problema é igual a:

(34) $L = \log \left|(Y_c - X_2 \Pi_2)' M_1 (Y_c - X_2 \Pi_2)\right| - 2\lambda ' \Pi_2 \ \gamma$

onde λ é o vetor formado pelos multiplicadores de Lagrange. As derivadas parciais de L com relação a Π_2, λ e γ são dadas por:[4]

(35) $\dfrac{\partial L}{\partial \Pi_2} = 2(X_2 ' M_1 Y_c - X_2 ' M_1 X_2 \Pi_2) W^{-1} - 2\lambda \gamma' = 0$

(36) $\dfrac{\partial L}{\partial \lambda} = 2\Pi_2 \ \gamma = 0$

(37) $\dfrac{\partial L}{\partial \gamma} = -2\Pi_2 ' \lambda = 0$

onde $W = (Y_c - X_2 \Pi_2)' M_1 (Y_c - X_2 \Pi_2)$.

A solução do sistema de equações formado por (35), (36) e (37) é um pouco intrincada e requer algumas manipulações algébricas. A partir da equação (35) temos que:

(38) $\Pi_2 = (X_2 ' M_1 \ X_2)^{-1} \ (X_2 ' M_1 Y_c{}^{-\lambda} \ \gamma' W)$

Como pela equação (36), Π_2 λ = 0, resulta que:

$$(X_2 ' M_1 X_2)^{-1} \ (X_2 \ M_1 \ Y_c - \lambda \ \gamma' W) \ \gamma = 0$$

Logo:

$$\lambda = \frac{X_2 ' M_1 Y_c \ \gamma}{\gamma' W \gamma}$$

Substituindo-se este valor de λ em (38) obtém-se:

(39) $\Pi_2 = \hat{\Pi}_2 - \dfrac{\hat{\Pi}_2 \ \gamma \gamma' W}{\gamma' W \gamma}$

onde:

$$\hat{\Pi}_2 = (X_2 ' M_1 X_2)^{-1} \ X_2 ' M_1 \ Y_c$$

[4]Aplicamos aqui os seguintes resultados:

$$\frac{\partial \log}{\partial X} \left| X' BX \right| = 2BX (X' BX)^{-1} \ , B = B'$$

$$\frac{\partial a' Xb}{\partial X} = a \ b' \quad e \quad \frac{\partial a' x}{\partial x} = a$$

A equação (37) implica em que $\Pi_2 \lambda = 0$. Substituindo-se os valores de Π_2 e λ obtidos anteriormente nesta expressão, resulta que:

$$\left(\hat{\Pi}_2 - \frac{\hat{\Pi}_2 \gamma\gamma' W}{\gamma' W \gamma}\right)' \frac{X_2' M_1 Y_c}{\gamma' W \gamma} \gamma = 0$$

ou alternativamente:

(40) $$\left(\hat{\Pi}_2' X_c' M_1 Y_c - \frac{W\gamma\gamma'\hat{\Pi}_2' X_2' M_1 Y_c}{\gamma' W \gamma}\right)\gamma = 0$$

É fácil verificar-se que

$$\hat{\Pi}_2' X_2' M_1 Y_c = Y_c' M_1 Y_c - \hat{W}$$

onde:

$$\hat{W} = (Y_c - X_2 \hat{\Pi}_2)' M_1 (Y_c - X_2 \hat{\Pi}_2)$$

e denominando-se por:

$$\theta = \gamma' W \gamma$$

$$\theta_1 = \gamma' \hat{\Pi}_2' X_2' M_1 Y_c \gamma$$

a equação (40) transforma-se em:

(41) $$\left(Y_c' M_1 Y_c - \hat{W} - \frac{\theta_1 W}{\theta}\right)\gamma = 0$$

Levando-se o valor de Π_2, da equação (39), na expressão de W, obtém-se:

(42) $$W = \hat{W} + \theta^{-2} \theta_1 W \gamma\gamma' W$$

Pós-multiplicando-se ambos os lados desta equação pelo vetor γ resulta:

$$W\gamma = \hat{W}\gamma + \theta^{-2} \theta_1 W \gamma \gamma' W \gamma$$

e daí obtém-se:

(43) $$W \gamma = \frac{1}{1-\phi} \hat{W} \gamma$$

onde $\phi = \theta_1/\theta$

Substituindo-se (43) em (41) chega-se a:

$$(Y_c' \, M_1 \, Y_c - \hat{W} - \frac{\phi}{1-\phi}\hat{W}) \; \gamma = 0$$

Alternativamente:

$$(Y_c' \, M_1 \, Y_c - \mu \hat{W}) \; \gamma = 0$$

ou ainda:[5]

(44) $$(Y_c \;\, M_1 \, Y_c - \mu \, Y_c' \, M \, Y_c) \; \gamma = 0$$

onde $\mu = 1/1-\phi$. Para que este sistema de equações tenha uma solução diferente da solução trivial $\gamma = 0$ é necessário que o valor de μ seja tal que:

$$\left| Y_c \; M_1 \, Y_c - \mu \, Y_c' \, M \, Y_c \right| = 0$$

Esta equação fornece um polinômio do grau L_1 na variável μ, e portanto teremos L_1 raízes, que são denominadas de raízes características. Resta, então, saber qual dessas raízes minimiza a função de verossimilhança condicionada. A seguir demonstraremos que o mínimo da função ocorre para a menor raiz característica do polinômio. Para tal finalidade comecemos por substituir (43) em (42):

$$W = \hat{W} + \theta^{-2} \;\; \theta_1 \; \frac{\hat{W} \; \gamma \gamma' \, \hat{W}}{(1-\phi)^2}$$

Associado a cada raiz característica μ existe um vetor característico γ que, entretanto, não é único, pois se γ for uma solução é fácil verificar-se que $k\gamma$ também é uma solução para qualquer valor de $k \neq 0$. Consequentemente, para se ter valores únicos para γ é preciso adotar-se uma regra de normalização. Adotaremos aqui a seguinte normalização:

$$\gamma' \;\, \hat{W} \, \gamma = 1$$

O valor de θ passa, então, a ser dado por:

$$\theta = \gamma' \, W \, \gamma = \frac{\gamma' \, \hat{W} \, \gamma}{1-\phi} = \frac{1}{1-\phi}$$

---

[5]A matriz $\hat{W}$ é igual a $Y_c \; ' \; M \; Y_c$,onde, $M = M_1 - M_1 \; X_2 \; (X_2 ' M_1 \, X_2)^{-1} \; X_2 ' \, M_1 = I - X \, (X ' X)^{-1} \; X '$. Para provar a última igualdade basta desenvolver:

$$I - [X_1 \;\, X_2] \begin{bmatrix} X_1 ' X_1 & X_1 ' X_2 \\ X_2 ' X_1 & X_2 ' X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1 ' \\ X_2 ' \end{bmatrix}$$

usando-se resultados da inversa de uma matriz particionada 2x2.

Com esses dois últimos resultados a expressão anterior de W transforma-se em:

(45) $W = \hat{W} + (\mu - 1)\,\hat{W}\,\gamma\,\gamma'\,\hat{W}$

pois $\theta_1 = \phi\ \theta = \mu$-1.

Seja C a matriz formada de acordo com:

$$C = \left[\,\gamma_{(1)}\ ,\ \gamma_{(2)}\,,\ldots,\gamma_{(L_1)}\right]$$

onde $\gamma_{(1)}\ ,\ \gamma_{(2)}\,,\ldots,\gamma_{(L_1)}$ são os vetores característicos correspondentes às raízes características $\mu_1\ ,\ \mu_2\,,\ldots,\mu_{L_1}$. Em virtude da normalização adotada temos que:

$$C'\,\hat{W}\ C = I$$

Utilizando-se uma propriedade de determinantes resulta que:

(46) $\left|C'\,\hat{W} C\right| = \left|C\right|^2\ \left|\hat{W}\right| = 1$

Por outro lado, pré-multiplicando-se (45) por C' e pós-multiplicando-a por C obtemos:

$$C'\,W\ C = C'\,\hat{W} C + (\mu_k\ - 1)\ C'\,\hat{W}\ \gamma_{(k)}\ \ \gamma'_{(k)}\ \hat{W} C$$

que é igual a:

$$C'\,W\ C = I\ + (\mu_k\ - 1)\ e_k\ \ e_k^{'}$$

em virtude da regra de normalização e do fato que $C'\,\hat{W}\ \gamma_{(k)} = e_k$, onde $e_k{}' = (0,\ldots 1,\ldots,0)$, com o valor 1 na késima posição. Segue-se, então, que:

$$\left|C'\,W\ C\right| = \mu_k$$

Como $\left|C'\,\hat{W} C\right| = \left|C\right|^2\ \left|\hat{W}\right|$ e levando-se em conta a expressão (46), conclui-se que:

$$\left|W\right| = \mu_k\ \left|\hat{W}\right|$$

Logo, a menor raiz característica fornecerá o valor mínimo da função de verossimilhança condicionada.

Uma vez determinado o valor de γ os demais parâmetros do modelo, W, Π_2, Π_1 e β_1, são facilmente calculados. Cabe ainda observar que o método de máxima verossimilhança de informação limitada independe de qual é a variável escolhida para ser colocada no lado esquerdo na equação estrutural, fato este que não ocorre com os demais métodos de estimação até aqui apresentados. Esta propriedade decorre do fato já assinalado anteriormente que a estimativa do parâmetro γ é determinada a menos de uma constante de proporcionalidade.

<u>Estimador de Máxima Verossimilhança de Informação Limitada: Enfoque de Minimização da Razão da Variância</u>

A equação (44) possibilita escrever μ do seguinte modo:

$$\mu = \frac{\gamma' Y_c' \, M_1 \, Y_c \, \gamma}{\gamma' Y_c' \, M \, Y_c \, \gamma}$$

Esta expressão pode ser interpretada de uma maneira bastante interessante. Com efeito, como

(47) $\quad Y_c \, \gamma = X_1 \, \beta_1 + \mu_1$

Se γ fosse conhecido, a soma dos quadrados dos resíduos desta regressão seria igual a :

$$SQR_1 = \gamma' \, Y_c' \, M_1 \, Y_c \, \gamma$$

Por outro lado, se acrescentássemos a matriz X_2 nos regressores da equação (47),

$$Y_c \, \gamma = Y_1 \, \beta_1 + X_2 \, \beta_2 + u_1 \, ,$$

a soma dos quadrados dos resíduos, na suposição de um valor conhecido de γ, desta regressão seria igual a:

$$SQR_2 = \gamma' \, Y_c' \, M \, Y_c \, \gamma$$

O valor de μ é justamente igual à razão dessas duas somas de quadrados de resíduos:

$$\mu = \frac{SQR_1}{SQR_2} = \frac{\gamma' Y_c' \, M_1 \, Y_c \, \gamma}{\gamma' \, Y_c' \, M \, Y_c \, \gamma}$$

Como o vetor de coeficientes β_2 é igual a zero, uma idéia de certo modo intuitiva é obter o valor de γ que minimiza a razão das somas dos quadrados dos resíduos, de sorte a se obter o menor valor para μ. Consequentemente, derivando-se μ em relação a γ,

$$\frac{\partial \mu}{\partial \gamma} = \frac{(\gamma' \, Y_c' \, M \, Y_c \, \gamma) \, 2 \, \gamma_c' \, M_1 \, Y_c \, \gamma - (\gamma' \, Y_c' \, M_1 \, Y_c \, \gamma) \, 2 \, Y_c' \, M \, Y_c \, \gamma}{(\gamma' Y_c' \, M \, Y_c \, \gamma)^2}$$

e igualando-se este resultado a zero, obtém-se:

$$(Y_c' \, M_1 \, Y_c - \mu \, Y_c' \, M \, Y_c) \, \gamma = 0$$

que nada mais é do que a equação (44). Uma vez determinado o valor de γ a estimativa de β_1 é facilmente calculada através de:

$$\hat{\beta}_1 = (X_1' \, X_1)^{-1} \, X_1' Y_c \, \gamma$$

Vale ressaltar que estas duas propriedades mencionadas asseguram a consistência dos estimadores de mínimos quadrados ordinários dos parâmetros das equções de um modelo recursivo.

Estimador de Mínimos Quadrados Indireto Generalizado

A expressão (15), repetida aqui por conveniência,

$$Q_1 \ \delta_1 = \pi_1$$

contém um sistema de K equações lineares com *(L_1 + K_1 - 1)* incógnitas, os elementos do vetor δ_1. A solução desta última é dada por:

$$\delta_1 = (Q_1 ' \, Q_1)^{-1} \ Q_1 ' \ \pi_1$$

O estimador de mínimos quadrados indireto generalizado $\hat{\delta}_1$ é obtido quando substituímos os valores de Q_1 e π_1 na expressão acima pelos estimadores $\hat{Q}_1$ *e* $\hat{\pi}_1$ da forma reduzida. Isto é:

$$\hat{\delta}_1 = (\hat{Q}_1 ' \hat{Q}_1)^{-1} \ \hat{Q}_1 ' \ \hat{\pi}_1$$

onde a matriz $\hat{Q}_1$ é igual a

(19) $$\hat{Q}_1 = \left[\hat{\Pi}_2 \ \vdots \ D \right] \quad ,$$

e os estimadores $\hat{\Pi}_2$ *e* $\hat{\pi}_1$ da forma reduzida são:

(20) $$\hat{\Pi}_2 = (X ' \ X)^{-1} \ X ' \ Y_1$$

(21) $$\hat{\pi}_1 = (X ' \ X)^{-1} \ X ' \ y_1$$

e a matriz D é dada por:

$$D = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

Denominamos o estimador $\hat{\delta}_1$ de estimador de mínimos quadrados indireto generalizado por dois motivos. Em primeiro lugar, porque este estimador é obtido indiretamente através dos estimadores $\hat{\Pi}_2$ *e* $\hat{\pi}_1$. Em segundo lugar porque, no caso particular de a equação estrutural ser exatamente identificada, $\hat{\delta}_1$ reduz-se ao tradicional estimador indireto de mínimos quadrados:

$$\hat{\delta}_1 = \hat{Q}_1^{-1} \ \hat{\pi}_1$$

Cuidemos, agora, de obter expressões algébricas para o estimador $\hat{\delta}_1$ em função das matrizes de observações X e $[y_1 \vdots Y_1]$. Usamos a relação (19) para escrever o estimador $\hat{\delta}_1$ da seguinte forma:

$$\hat{\delta}_1 = \begin{bmatrix} \hat{\Pi}_2'D & \hat{\Pi}_2'D \\ D'\hat{\Pi}_2 & D'D \end{bmatrix}^{-1} \begin{bmatrix} \hat{\Pi}'_2 \\ D' \end{bmatrix} \hat{\pi}_1$$

Em seguida, substituímos os valores de $\hat{\Pi}_2$ *e* $\hat{\pi}_1$ dados em (20) e (21) na expressão anterior e obtemos:

$$(22) \qquad \hat{\delta}_1 = \begin{bmatrix} Y_1' X(X'X)^{-2} X'Y_1 Y_1' X(X'X)^{-2} X'X_1 \\ X_1' X(X'X)^{-2} X'Y_1 \ I \end{bmatrix}^{-1}$$

$$\begin{bmatrix} Y_1' X(X'X)^{-2} X' \\ X_1' X(X'X)^{-2} X' \end{bmatrix} y_1$$

Para se chegar a essa expressão levamos em conta que $D'\Pi_2 = X_1' \ X \ (X'X)^{-2} \ X'Y_1$ A demonstração desta igualdade é bastante simples:

$$D'\hat{\Pi}_2 = D'(X'X)^{-1} \ X'Y_1 = D'X'X \ (X'X)^{-1} (X'X)^{-1} \ X'Y_1 = X_1' \ X \ (X'X)^{-2} \ X'Y_1$$

tendo em vista que $D'X' = X_1'$. De maneira similar calcula-se o produto $D' \hat{\Pi}_1$.

Usando-se uma notação mais simples a expressão (22) pode ser escrita como:

$$(23) \qquad \hat{\delta}_1 = \left[Z_1' X(X'X)^{-2} X'Z_1 \right]^{-1} Z_1' X(X'X)^{-2} X'y_1$$

levando-se em conta que:

$$X_1' \ X \ (X'X)^{-2} \ X'X_1 = D'X'X \ (X'X)^{-2} \ X'XD = I$$

pois $D'D = I$.

<u>Propriedades do Estimador $\hat{\delta}_1$</u>

Substituímos o valor de y_1, dado pelo lado direito do segundo sinal de igualdade de (10), em (23), obtemos:

$$(24) \qquad \hat{\delta}_1 = \delta_1 + \left[Z_1' \ X \ (X'X)^{-2} \ X'Z_1 \right]^{-1} Z_1' \ X \ (X'X)^{-2} \ X'u_1$$

De acordo com algumas hipóteses tradicionais no estudo dos estimadores dos parâmetros de um sistema de equações simultâneas temos que[6]

---

[6]Essas hipóteses estão listadas em H.Theil (1971), Principles of Econometrics, John Wiley, capítulo 10.

$$\text{limp}\left[\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2}\frac{X'Z_1}{T}\right]^{-1} \quad \text{é finita, e}$$

$$\text{limp}\left[\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2}\frac{X'u_1}{T}\right] \quad \text{é igual a zero.}$$

A notação *limp* denota o limite em probabilidade da variável indicada. Aplicando-se estes resultados a (24) concluímos que:

$$\text{limp } \hat{\delta}_1 = \delta_1$$

o que significa dizer que o estimador indireto generalizado de mínimos quadrados é consistente.

A distribuição assintótica da variável aleatória $\sqrt{T}\ (\hat{\delta}_1 - \delta_1)$ pode ser obtida com o seguinte procedimento da expressão (24) temos que:

$$\sqrt{T}(\hat{\delta}_1 - \delta_1) = \left[\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2}\frac{X'Z_1}{T}\right]^{-1}\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2}\frac{X'u_1}{\sqrt{T}}$$

Por outro lado, as hipóteses a que nos referimos abaixo da expressão (24) nos possibilita afirmar que:

$$\text{limp}\left[\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2}\frac{X'Z_1}{T}\right]^{-1}\frac{Z_1^{'}X}{T}\left(\frac{X'X}{T}\right)^{-2} = G$$

onde G é uma matriz cujos elementos são finitos. Em seguida, lançamos mão do teorema que afirma que a distribuição assintótica de $X'u_1\ /\sqrt{T}$ é normal com valor esperado zero e matriz de variância-covariância igual a σ_{11} limp $\left(\frac{X'X}{T}\right)$. Podemos, então, concluir que a distribuição assintótica de é $\sqrt{T}\ (\hat{\delta}_1 - \delta_1)$ é normal com média zero e matriz de variância-covariância, Var $\hat{\delta}_1$, igual a:

(25) $\quad \sigma_{11}G\left[\text{limp}\left(\frac{X'X}{T}\right)\right]G' = Var\hat{\delta}_1$

A estatística

(26) $\quad s_{11} = \frac{1}{T}(y_1 - Z_1\ \hat{\delta}_1)'\ (y_1 - Z_1\ \hat{\delta}_1)$

é um estimador consistente da variância σ_{11}. A prova dessa propriedade é bastante simples. Substituindo-se (10) na expressão anterior obtemos:

$$s_{11} = \frac{u_1^{'}\, u_1}{T} - \frac{u_1^{'}\, Z_1}{T}\,(\hat{\delta}_1 - \delta_1) - \frac{(\hat{\delta}_1 - \delta_1)'\, Z_1^{'}\, u_1}{T} + (\hat{\delta}_1 - \delta_1)'\,\frac{Z_1^{'}\, Z_1}{T}\,(\hat{\delta}_1 - \delta_1)$$

A partir desta relação concluímos que:

$$\text{limp}\ \ s_{11} = \text{limp}\ \frac{u_1^{'}\, u_1}{T} = \sigma_{11}$$

baseados no fato de que limp $(\hat{\delta}_1 - \delta_1) = 0$ e de que limp $u_1^{'}\, u_1 \,/\, T = \sigma_{11}$.

É fácil derivar-se a partir de (25) e (26) os erros padrões (assintóticos) da estimativa do estimador indireto generalizado de mínimos quadrados. Isto é, a partir destas expressões concluímos que os erros-padrões da estimativa dos elementos do vetor $\hat{\delta}_1$ são dados pelas raízes quadradas dos elementos da diagonal principal da matriz

$$s_{11} = \left[Z_1^{'}\ X\ (X'X)^{-2}\ X'Z_1\right]^{-1} Z_1^{'}\ X\ (X'X)^{-3}\ X'\,Z_1\left[Z_1^{'}\ X\ (X'X)^{-2}\ X'Z_1\right]^{-1}$$

<u>Mínimos Quadrados Em Duas Etapas</u>

O estimador de mínimos quadrados em duas etapas pode ser derivado de diferentes maneiras. A seguir derivaremos este estimador com o auxílio do enfoque algébrico. da forma reduzida do modelo, expresso em (11), temos que

$$y_1 = [X_1 \vdots X_2]\begin{bmatrix}\Pi_{11}\\ \Pi_{12}\end{bmatrix} + v_1$$

Por outro lado, da equação (14) sabemos que:

$$\begin{bmatrix}\Pi_{11}\\ \Pi_{12}\end{bmatrix} = \begin{bmatrix}\Pi_{21} I\\ \Pi_{22} 0\end{bmatrix}\begin{bmatrix}\gamma_1\\ \beta_1\end{bmatrix}$$

Substituindo-se esta relação na expressão anterior concluímos que:

$$y_1 = [X_1 \vdots X_2]\begin{bmatrix}\Pi_{21} & I\\ \Pi_{22} & 0\end{bmatrix}\begin{bmatrix}\gamma_1\\ \beta_1\end{bmatrix} + v_1$$

Se os valores de Π_{21} e Π_{22} fossem conhecidos, poderíamos aplicar o método de mínimos quadrados ordinários à equação anterior. O estimador assim obtido seria dado por:

$$d_1 = \left[Q_1^{'}\ X'X\ Q_1\right]^{-1}\ \ Q_1^{'}\ X'\,y_1$$

Entretanto, devido ao fato de que Π_{21} e Π_{22} serem parâmetros que têm de ser estimados, visto não serem conhecidos a priori, o estimador acima não pode ser aplicado na prática. Todavia, o problema é facilmente contornado usando-se a matriz

$$\hat{Q}_1 = \begin{bmatrix} \hat{\Pi}_2 & I \\ \hat{\Pi}_{22} & 0 \end{bmatrix}$$

ao invés da matriz Q_1, onde $\hat{\Pi}_{21}$ *e* $\hat{\Pi}_{22}$ são os estimadores de mínimos quadrados da forma reduzida do modelo. Dessa maneira, obtemos o estimador de mínimos quadrados em duas etapas:

$$\tilde{\delta}_1 = \left[\hat{Q}_1 ' X' X \hat{Q}_1\right]^{-1} \hat{Q}_1 ' X' y_1$$

Usamos a relação (19) para escrever este estimador na seguinte forma:

$$\tilde{\delta}_1 = \begin{bmatrix} \hat{\Pi}_2 ' X' X \hat{\Pi}_2 & \hat{\Pi}_2 ' X' X D \\ D' X' X \hat{\Pi}_2 & D' X' X D \end{bmatrix}^{-1} \begin{bmatrix} \hat{\Pi}_2 X' \\ D' X' \end{bmatrix} Y_1$$

Em seguida, substituindo-se o valor de Π_2 dado por (20) na expressão acima resulta:

$$\tilde{\delta}_1 = \begin{bmatrix} Y_1 ' X (X' X)^{-1} X' Y_1 & Y_1 ' X_1 \\ X_1 ' Y_1 & X_1 ' X_1 \end{bmatrix}^{-1} \begin{bmatrix} Y_1 ' X (X' X)^{-1} X' \\ X_1 ' \end{bmatrix} y_1$$

Alternativamente, essa expressão pode ser escrita de maneira mais compacta do seguinte modo:

$$\tilde{\delta}_1 = \left[Z_1 ' X (X' X)^{-1} X' Z_1\right]^{-1} Z_1 ' X (X' X)^{-1} X' y_1$$

O enfoque que adotamos na derivação do estimador de mínimos quadrados em duas etapas não mostra o porque do seu nome. Todavia, não é difícil entender qual a origem da palavra duas etapas no nome do estimador. Com efeito, observe que o valor previsto de Y_1 que denominamos por $\hat{Y}_1$, é igual ao produto da matriz X vezes a estimativa de mínimos quadrados ordinários, $\hat{\Pi}_2$, da forma reduzida:

$$\hat{Y}_1 = X \hat{\Pi}_2 = X (X' X)^{-1} X' Y_1$$

Como,

$$\hat{Y}_1 ' \hat{Y}_1 = Y_1 ' X (X' X)^{-1} X' X (X' X)^{-1} X' Y_1 = Y_1 ' X (X' X)^{-1} X' Y_1$$

e:

$$\hat{Y}_1 ' X_1 = Y_1 ' X (X' X)^{-1} X' X_1 = Y_1 ' X (X' X)^{-1} X' X D = Y_1 ' X D = Y_1 ' X_1$$

Segue-se que:

$$\tilde{\delta}_1 = \begin{bmatrix} \hat{Y}_1'\hat{Y}_1 \hat{Y}_1' X_1 \\ \\ X_1'\hat{Y}_1 X_1' X_1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{Y}_1' \\ \\ X_1' \end{bmatrix} y_1$$

Logo, a estimativa $\tilde{\delta}_1$ é também obtida pelo seguinte procedimento em duas etapas:

<u>1ª etapa</u>: a partir da regressão de Y_1
contra X obtém-se $\hat{Y}_1$;

<u>2º etapa</u>: faz-se, então, a regressão de
y_1 contra $\hat{Y}_1$ e X_1 .

Pode-se provar, de maneira análoga ao que foi feito para o estimador indireto generalizado de mínimos quadrados, que o estimador de mínimos quadrados em duas estapas é consistente, limp $\tilde{\delta}_i = \delta_1$, e que $\sqrt{T}\ (\tilde{\delta}_1 - \delta_1)$ tem uma distribuição asintótica normal com média zero e matriz de variância-covariância igual a:

$$\sigma_{11} \text{limp} \left[\left(\frac{Z_1' X}{T} \right) \left(\frac{X' X}{T} \right)^{-1} \left(\frac{X' Z_1}{T} \right) \right]^{-1}$$

Portanto, os erros padrões da estimativa dos elementos do vetor $\tilde{\delta}_1$ são dados pelas raízes quadradas dos elementos da diagonal principal da matriz:

$$s_{11} \left[Z_1' X\ (X' X)^{-1}\ X' Z_1 \right]^{-1}$$

A interpretação que acabamos de descrever do estimador de mínimos quadrados em duas etapas é bastante sugestiva, mas não deve ser levada ao pé da letra, sob pena de se cometer um erro importante no cálculo dos erros padrões das estimativas dos parâmetros do modelo. Com efeito, suponhamos que um pesquisador não dispondo de um pacote de computação que calculasse diretamente a estimativa de mínimos quadrados em duas etapas resolvesse aplicar, em duas etapas, mínimos quadrados ordinários. Nestas circunstâncias a estimativa da variância σ_{11} seria fornecida pela seguinte expressão:

$$\tilde{\sigma}_{11} = \frac{\tilde{u}_1' \tilde{u}_1}{T}$$

onde $\tilde{u}_1 = y_1 - \hat{Y}_1\ \tilde{\gamma}_1 - X_1\ \tilde{\beta}_1$.

Como $Y_1 = \hat{Y}_1 + \hat{V}_1$, segue-se então que:

$$\tilde{u}_1 = \hat{u}_1 + \hat{V}_1\ \tilde{\gamma}_1$$

onde $\hat{u}_1 = y_1 - Y_1\ \tilde{\gamma}_1 - X_1\ \tilde{\beta}_1$.

Com a finalidade de examinarmos se o estimador $\tilde{\sigma}_{11}$ é consistente procedermos do seguinte modo. O limite em probabilidade de $\tilde{\sigma}_{11}$ é igual a:

$$\text{limp}\ \tilde{\sigma}_{11} = \text{limp}\ \frac{\tilde{u}_1'\ \tilde{u}_1}{T} = \lim \frac{(\hat{u}_1 + \hat{V}_1\ \tilde{\gamma}_1)'\ (\hat{u}_1 + \hat{V}_1\ \tilde{\gamma}_1)}{T}$$

Alternativamente:

$$(27)\qquad \text{limp}\ \tilde{\sigma}_{11} = \text{limp}\frac{\hat{u}_1'\ \hat{u}_1}{T} + \text{limp}\ 2\ \frac{\hat{u}_1'\hat{v}_1\ \tilde{\gamma}_1}{T} + \text{limp}\ \frac{\tilde{\gamma}_1'\hat{V}_1'\hat{V}_1\ \tilde{\gamma}_1}{T}$$

Como

$$\text{limp}\ \frac{\hat{u}_1'\ \hat{u}_1}{T} = \sigma_{11}$$

e provaremos em seguida que:

$$\text{limp}\ 2\ \frac{\hat{u}_1'\hat{V}_1\ \tilde{\gamma}_1}{T} = (\text{limp}\ 2\ \frac{u_1'V_1}{T})\ \gamma_1 = \text{limp}\ \frac{\tilde{\gamma}_1'\hat{V}_1'\hat{V}_1\ \tilde{\gamma}_1}{T} = \gamma_1'\ \text{limp}\ \frac{V_1'\ V_1}{T}\ \gamma_1 =$$

segue-se, então, que em geral:

$$\text{limp}\ \tilde{\sigma}_{11} \neq \sigma_{11}$$

A matriz de resíduos $\hat{V}_1$ é igual a:

$$\hat{V}_1 = Y_1 - \hat{Y}_1 = Y_1 - X\ (X'X)^{-1}\ X'\ Y_1 = \left[I - X(X'X)^{-1}\ X'\right] Y_1$$

Como $Y_1 = X\ \Pi_2 + V_1$, a matriz $\hat{V}_1$ pode ser expressa por:

$$\hat{V}_1 = \left[I - X\ (X'X)^{-1}\ X\right] V_1 = M\ V_1$$

Logo,

$$\text{limp}\ \frac{\hat{V}_1'\hat{V}_1}{T} = \text{limp}\ \frac{V_1\ '\ M'\ M\ V_1}{T} = \text{limp}\ \frac{V'\ M\ V_1}{T}$$

ou ainda:

$$\text{limp}\frac{\hat{V}_1'\hat{V}_1}{T} = \text{limp}\frac{V_1'V_1}{T} - \text{limp}\frac{V_1'X}{T}\left(\frac{X'X}{T}\right)^{-1}\frac{X'V_1}{T}$$

Em virtude de

$$\operatorname{limp} \frac{V_1' X}{T} = 0$$

pois VΓ = U e limp $\frac{U' X}{T}$ = 0, segue-se que:

$$\operatorname{limp} \frac{\hat{V}_1' \hat{V}_1}{T} = \operatorname{limp} \frac{\mathrm{V}_1' V_1}{T}$$

Por outro lado,

$$\operatorname{limp} \frac{\hat{u}_1' \hat{V}_1}{T} = \operatorname{limp} \frac{U_1' N' M V_1}{T}$$

onde $\hat{u}_1 = y_1 - Z_1 \hat{\delta}_1 = N u_1$, e

$$N = I - Z_1 \left[Z_1' X (X' X)^{-1} X' Z_1 \right]^{-1} Z_1' X (X' X)^{-1} X'$$

Com um pouco de paciência e álgebra pode-se, então, mostrar que:

$$\operatorname{limp} \frac{u_1' N' M V_1}{T} = \operatorname{limp} \frac{u_1' V_1}{T}$$

Utilizando-se a partição indicada antes da expressão (11), obtém-se:

$$v_1 - V_1 \gamma_1 = u_1$$

Consequentemente:

$$\operatorname{limp} \frac{u_1' V_1}{T} = \operatorname{limp} \frac{(\mathrm{v}_1 - V_1 \gamma_1)' V_1}{T} = \operatorname{limp} \frac{v_1' V_1}{T} - \operatorname{limp} \frac{\gamma_1' V_1' V_1}{T}$$

Substituindo-se este resultado em (27), conclui-se que:

$$\operatorname{limp} \tilde{\sigma}_{11} = \sigma_{11} + 2 \left(\operatorname{limp} \frac{v_1' V_1}{T}\right) \gamma_1 - \gamma_1' \left(\operatorname{limp} \frac{V_1' V_1}{T}\right) \gamma_1$$

Admitindo-se que no modelo de equações simultâneas a matriz de variância-covariância dos termos estocásticos da matriz U da forma estrutural é dada por:

$$E\, U' U = \Sigma$$

a matriz de variância-covariância da forma reduzida é igual a:

$$E\, V' V = E (\Gamma^{-1})' U' U\, \Gamma^{-1} = (\Gamma^{-1})' \Sigma\, \Gamma^{-1} = \Omega$$

Logo, se a matriz for particionada como em (11) segue-se que:

$$E\begin{bmatrix} v_1' \\ V_1' \\ V_2' \end{bmatrix}[v_1 \vdots V_2 \vdots V_2] = E\begin{bmatrix} v_1'v_1 & v_1'V_1 & v_1'V_2 \\ V_1'v_1 & V_1'V_1 & V_1'V_2 \\ V_2'v_1 & V_2'V_1 & V_2'V_2 \end{bmatrix}$$

que é igual a:

$$EV'V = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix}$$

Portanto:

$$\text{limp}\ \frac{v_1'V_1}{T} = \Omega_{12}$$

e

$$\text{limp}\ \frac{V_1'V_1}{T} = \Omega_{22}$$

e a inconsistência do estimador $\hat{\sigma}_{11}$ está demonstrada. Isto é:

$$\text{limp}\ \tilde{\sigma}_{11} = \tilde{\sigma}_{11} + 2\ \Omega_{12}\ \gamma_1 - \gamma_1'\ \Omega_{22}\ \gamma_1$$

<u>Máxima Verossimilhança: Informação Limitada</u>

O estimador de máxima verossimilhança de informação limitada dos parâmetros de uma equação estrutural do modelo de equações simultâneas na sua concepção é bastante simples, pois consiste na aplicação direta do método de máxima verossimilhança. Entretanto, sua obtenção requer um pouco de álgebra. Com a finalidade de tornar a dedução desse estimador menos complicada comecemos por introduzir uma notação que facilite as manipulações algébricas. A equação estrutural (10), que deseja estimar, passa a ser então escrita como:

$$[y_1 \vdots Y_1]\begin{bmatrix} 1 \\ -\gamma_1 \end{bmatrix} = X_1\beta_1 + u_1$$

ou:

(28) $\quad Y_c\ \gamma = X_1\ \beta_1 + u_1$

onde:

$$Y_c = [y_1 \vdots Y_1]\ \text{e}\ \gamma = \begin{bmatrix} 1 \\ -\gamma_1 \end{bmatrix}$$

A equação (11), da forma reduzida do modelo de equações simultâneas,permite escrever[7]

(29) $Y_c = X_1 \Pi_1 + X_2 \Pi_2 + V_c$

onde:

$$\Pi_1 = [\Pi_{11} \; \Pi_{21}]$$

$$\Pi_2 = [\Pi_{12} \; \Pi_{22}]$$

$$V_c = [v_1 \;\; V_1]$$

As restrições (14) repetidas aqui por conveniência,

$$\Pi_{11} = \Pi_{21} \; \gamma_1 + \beta_1$$
$$\Pi_{12} = \Pi_{22} \; \gamma_1$$

podem ser reescritas na nova notação do seguinte modo:

(30) $\Pi_1 \; \gamma = \beta_1$

(31) $\Pi_2 \; \gamma = 0$

A matriz V_C dos termos estocásticos da forma reduzida segue uma distribuição normal com média igual a 0 e variância-covariância igual a Ω_C, cuja função de densidade de probabilidade é igual a:

$$p\,(V_c) = (2\Pi)^{-\frac{L_1 T}{2}} \left|\Omega_c\right|^{-T/2} \;\; \exp\,(-\frac{1}{2}\, tr\, V_c\,'V_c\; \Omega_c^{-1})$$

onde exp( ) indica o número natural e elevado ao termo entre parênteses e o símbolo tr representa o traço da matriz.

Como o Jacobiano da transformação de V_C para Y_C é igual a 1, o logaritmo da função de densidade de probabilidade de Y_C é igual a:

(32)
$$\log \text{p}\,(Y_c \,/\, \Pi_1, \Pi_2, \Omega_c, X) = -\frac{L_1 \; T}{2} \log 2\Pi - \frac{T}{2} \log \left|\Omega_c\right|$$
$$-\frac{1}{2}\, tr\, (Y_c - X_1 \,\Pi_1 - X_2 \,\Pi_2)'\,(Y_c - X_1 \,\Pi_1 - X_2 \,\Pi_2)\, \Omega_c^{-1}$$

Uma vez conhecidos os valores de Y_C esta função se transforma no logaritmo de função de verossimilhança, isto é:

[7]Mudamos, por conveniência, a notação adotada até aqui: o símbolo Π_2 representa agora outra matriz.

(33) $\ell = \log\, p\,(\Pi_1, \Pi_2, \Omega_c \,/\, Y_c, X)$

A equação (30) não implica em nenhuma restrição sobre os elementos da matriz Π_1. Ela indica apenas como calcular o vetor β_1 quando Π_1 e γ forem conhecidos. O mesmo não ocorre com a equação (31), pois ela impõe restrições sobre os elementos da matriz Π_2. Consequentemente, os parâmetros do modelo devem ser estimados levando-se em conta a restrição (31). O problema consiste, então, em maximizar a função

$$\log\, p\,(\Pi_1, \Pi_2, \Omega_c \,/\, Y_c,\, X)$$

sujeito a restrição:

$$\Pi_2\ \gamma = 0$$

Antes de escrever a expressão de Lagrange deste problema de maximização condicionada é interessante observar que a restrição não involve as matrizes Π_1 e Ω_c. Este fato simplifica um pouco a solução do problema pois podemos derivar parcialmente o logaritmo da função de verossimilhança em relação a Π_1 e Ω_c^{-1}, igualá-las a zero, e obter os valores de Π_1 e Ω_c em função de Π_2, Y_c e X. Isto é:[8]

$$\frac{\partial \ell}{\partial \Pi_1} = -\frac{1}{2}\,(-2\,\Omega_c^{-1})\left[(Y_c - X_2\ \Pi_2)'\,X_1 - \Pi_1{}'Y_1{}'\,X_1\right] = 0$$

e

$$\frac{\partial \ell}{\partial \Omega_c^{-1}} = \frac{T}{2}\,\Omega_c - \frac{1}{2} V_c{}'V_c = 0$$

Portanto,

$$\tilde{\Pi}_1 = (X_1{}'\,X_1)^{-1}\ X_1{}'(Y_c - X_2\,\Pi_2)$$

e

$$\Omega_c = \frac{V_c{}'V_c}{T}$$

---

[8]Aplicamos aqui os seguintes resultados:

$$\frac{\partial tr}{\partial \Pi}(Y - X\ \Pi)'(Y - X\ \Pi)\ \Omega^{-1} = -2\ \Omega^{-1}\ (Y'\,X - \Pi'\,X'\,X)$$

$$\frac{\partial \log}{\partial \Omega_c{}^{-1}}\left|\Omega_c\right| = -\ \Omega_c$$

$$\frac{\partial tr}{\partial \Omega^{-1}} V_c{}'V_c\ \Omega^{-1} = V_c{}'V_c$$

e derivamos em relação a $\Omega_c{}^{-1}$, ao invés de Ω_c, pois a derivada é mais fácil e a propriedade de invariância do método de máxima verossimilhança permite que isto seja feito.

Como

$$V_c = Y_c - X_1\,\Pi_1 - X_2\,\Pi_2 = Y_c - X_1\,(X_1'\,X_1)^{-1}\,X_1'(Y_c - X_2\,\Pi_2) - X_2\,\Pi_2$$

$$= \left[I - X_1\,(X_1'\,X_1)^{-1}\;X_1' \right] (Y_c - X_2\,\Pi_2) = M_1\,(Y_c - X_c\,\Pi_2)$$

onde $M_1 = I - X_1\,(X_1'\,X_1)^{-1}\,X_1'$, é uma matriz idempotente, segue-se que:

$$\Omega_c = \frac{V_c' V_c}{T} = \frac{(Y_c - X_2\,\Pi_2)'\,M_1\,(Y_c - X_2\,\Pi_2)}{T}$$

Substituindo-se esta expressão na função de verossimilhança obtém-se a função de verossimilhança concentrada:

$$\log \text{p}\,(\Pi_2\,/\,Y_c\,,X) = -\frac{L_1 T}{2}\log 2\Pi - \frac{T}{2}\log\left|\frac{(Y_c - X_2\,\Pi_2)'\,M_1\,(Y_c - X_2\,\Pi_2}{T}\right| - \frac{1}{2}\,tr\;I$$

Desprezando-se as constantes que aparecem nesta função, a maximização do log p $(\Pi_2/Y_c,X)$, sujeito à restrição $\Pi_2\,\gamma=0$, é equivalente ao seguinte problema:

minimizar

$$\log\left|(Y_c - X_2\,\Pi_2)'\,M_1\,(Y_c - X_2\,\Pi_2)\right|$$

com relação Π_2, sujeito à restrição:

$$\Pi_2\,\gamma = 0$$

A expressão de Lagrange deste problema é igual a:

(34) $$L = \log\left|(Y_c - X_2\,\Pi_2)'\,M_1(Y_c - X_2\,\Pi_2)\right| - 2\,\lambda'\,\Pi_2\,\gamma$$

onde λ é o vetor formado pelos multiplicadores de Lagrange. As derivadas parciais de L com relação a Π_2, λ e γ são dadas por:[9]

(35) $$\frac{\partial L}{\partial \Pi_2} = 2(X_2'\,M_1\,Y_c - X_2'\,M_1 X_2\,\Pi_2)W^{-1} - 2\,\lambda\,\gamma' = 0$$

(36) $$\frac{\partial L}{\partial \lambda} = 2\,\Pi_2\;\gamma = 0$$

[9]Aplicamos aqui os seguintes resultados:

$$\frac{\partial \log}{\partial X}\left|X'BX\right| = 2BX\,(X'BX)^{-1}\;,\;B = B'$$

$$\frac{\partial a'Xb}{\partial X} = a\,b' \quad e \quad \frac{\partial a'x}{\partial x} = a$$

(37) $$\frac{\partial L}{\partial \gamma} = -2\,\Pi_2\,'\,\lambda = 0$$

onde $W = (Y_c - X_2\,\Pi_2)\,'\,M_1\,(Y_c - X_2\,\Pi_2)$.

A solução do sistema de equações formado por (35), (36) e (37) é um pouco intrincada e requer algumas manipulações algébricas. A partir da equação (35) temos que:

(38) $$\Pi_2 = (X_2\,'\,M_1\;X_2)^{-1}\,(X_2\,'\,M_1 Y_c{}^{-\lambda}\;\gamma' W)$$

Como pela equação (36), $\Pi_2\ \lambda = 0$, resulta que:

$$(X_2\,'\,M_1 X_2)^{-1}\,(X_2\;M_1\;Y_c - \lambda\,\gamma'\,W)\;\gamma = 0$$

Logo:

$$\lambda = \frac{X_2\,'\,M_1\,Y_c\;\gamma}{\gamma'\,W\,\gamma}$$

Substituindo-se este valor de λ em (38) obtém-se:

(39) $$\Pi_2 = \hat{\Pi}_2 - \frac{\hat{\Pi}_2\;\gamma\gamma' W}{\gamma' W\,\gamma}$$

onde:

$$\hat{\Pi}_2 = (X_2\,'\,M_1\;X_2)^{-1}\;X_2\,'\;M_1\;Y_c$$

A equação (37) implica em que $\Pi_2\ \lambda = 0$. Substituindo-se os valores de Π_2 e λ obtidos anteriormente nesta expressão, resulta que:

$$\left(\hat{\Pi}_2 - \frac{\hat{\Pi}_2 \gamma\gamma' W}{\gamma' W \gamma}\right)' \frac{X_2\,'M_1 Y_c}{\gamma' W \gamma}\gamma = 0$$

ou alternativamente:

(40) $$\left(\hat{\Pi}_2\,'X_c\,'M_1 Y_c - \frac{W\gamma\gamma'\hat{\Pi}_2\,'X_2\,'M_1 Y_c}{\gamma' W \gamma}\right)\gamma = 0$$

É fácil verificar-se que

$$\hat{\Pi}_2\,'\,X_2\,'\,M_1\;Y_c = Y_c\,'\;M_1\;Y_c - \hat{W}$$

onde:

$$\hat{W} = (Y_c - X_2\ \hat{\Pi}_2)'\ M_1\ (Y_c - X_2\ \hat{\Pi}_2)$$

e denominando-se por:

$$\theta = \gamma'\ W\ \gamma$$

$$\theta_1 = \gamma'\,\hat{\Pi}_2'\ X_2'\ M_1\ Y_c\ \gamma$$

a equação (40) transforma-se em:

(41) $$\left(Y_c' M_1 Y_c - \hat{W} - \frac{\theta_1 W}{\theta} \right) \gamma = 0$$

Levando-se o valor de Π_2, da equação (39), na expressão de W, obtém-se:

(42) $$W = \hat{W} + \theta^{-2}\ \theta_1\ W\ \gamma\gamma'\ W$$

Pós-multiplicando-se ambos os lados desta equação pelo vetor γ resulta:

$$W\gamma = \hat{W}\gamma + \theta^{-2}\ \theta_1\ W\gamma\ \gamma'\ W\gamma$$

e daí obtém-se:

(43) $$W\ \gamma = \frac{1}{1-\phi}\ \hat{W}\ \gamma$$

onde $\phi = \theta_1/\theta$

Substituindo-se (43) em (41) chega-se a:

$$(Y_c'\ M_1\ Y_c - \hat{W} - \frac{\phi}{1-\phi}\hat{W})\ \gamma = 0$$

Alternativamente:

$$(Y_c'\ M_1\ Y_c - \mu \hat{W})\ \gamma = 0$$

ou ainda:[10]

---

[10]A matriz $\hat{W}$ é igual a $Y_c'\ M\ Y_c$, onde, $M = M_1 - M_1\ X_2\ (X_2'\ M_1\ X_2)^{-1}\ X_2'\ M_1 = I - X\,(X'\,X)^{-1}\ X'$. Para provar a última igualdade basta desenvolver:

$$I - [X_1\ X_2] \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1} \begin{bmatrix} X_1' \\ X_2' \end{bmatrix}$$

(44) $(Y_c\ M_1\ Y_c - \mu\, Y_c'\ M\ Y_c)\ \gamma = 0$

onde $\mu = 1/1-\phi$. Para que este sistema de equações tenha uma solução diferente da solução trivial $\gamma = 0$ é necessário que o valor de μ seja tal que:

$$\left| Y_c\ M_1\ Y_c - \mu\, Y_c'\ M\ Y_c \right| = 0$$

Esta equação fornece um polinômio do grau L_1 na variável μ, e portanto teremos L_1 raízes, que são denominadas de raízes características. Resta, então, saber qual dessas raízes minimiza a função de verossimilhança condicionada. A seguir demonstraremos que o mínimo da função ocorre para a menor raiz característica do polinômio. Para tal finalidade comecemos por substituir (43) em (42):

$$W = \hat{W} + \theta^{-2}\ \theta_1\ \frac{\hat{W}\ \gamma \gamma'\ \hat{W}}{(1-\phi)^2}$$

Associado a cada raiz característica μ existe um vetor característico γ que, entretanto, não é único, pois se γ for uma solução é fácil verificar-se que $k\gamma$ também é uma solução para qualquer valor de $k \neq 0$. Consequentemente, para se ter valores únicos para γ é preciso adotar-se uma regra de normalização. Adotaremos aqui a seguinte normalização:

$$\gamma'\ \hat{W}\ \gamma = 1$$

O valor de θ passa, então, a ser dado por:

$$\theta = \gamma'\, W\, \gamma = \frac{\gamma'\, \hat{W}\, \gamma}{1-\phi} = \frac{1}{1-\phi}$$

Com esses dois últimos resultados a expressão anterior de W transforma-se em:

(45) $W = \hat{W} + (\mu - 1)\ \hat{W}\ \gamma\ \gamma'\ \hat{W}$

pois $\theta_1 = \phi\ \theta = \mu$-1.

Seja C a matriz formada de acordo com:

$$C = \left[\, \gamma_{(1)}\ ,\ \gamma_{(2)}\, ,\ldots,\, \gamma_{(L_1)} \,\right]$$

onde $\gamma_{(1)}\ ,\ \gamma_{(2)}\, ,\ldots,\, \gamma_{(L_1)}$ são os vetores característicos correspondentes às raízes características $\mu_1\ ,\ \mu_2\, ,\ldots,\mu_{L_1}$. Em virtude da normalização adotada temos que:

$$C'\ \hat{W}\ C = I$$

Utilizando-se uma propriedade de determinantes resulta que:

---

usando-se resultados da inversa de uma matriz particionada 2x2.

(46) $\left|C'\hat{W}C\right| = \left|C\right|^2 \left|\hat{W}\right| = 1$

Por outro lado, pré-multiplicando-se (45) por C' e pós-multiplicando-a por C obtemos:

$$C'\,W\,C = C'\,\hat{W}C + (\mu_k - 1)\,C'\,\hat{W}\,\gamma_{(k)}\,\gamma'_{(k)}\,\hat{W}C$$

que é igual a:

$$C'\,W\,C = I + (\mu_k - 1)\,e_k\,e_k^{'}$$

em virtude da regra de normalização e do fato que C' $\hat{W}$ $\gamma_{(k)} = e_k$, onde $e_k\,' = (0,\ldots 1,\ldots,0)$, com o valor 1 na késima posição. Segue-se, então, que:

$$\left|C'\,W\,C\right| = \mu_k$$

Como $\left|C'\hat{W}C\right| = \left|C\right|^2 \left|\hat{W}\right|$ e levando-se em conta a expressão (46), conclui-se que:

$$\left|W\right| = \mu_k\,\left|\hat{W}\right|$$

Logo, a menor raiz característica fornecerá o valor mínimo da função de verossimilhança condicionada.

Uma vez determinado o valor de γ os demais parâmetros do modelo, W, Π_2, Π_1 e β_1, são facilmente calculados. Cabe ainda observar que o método de máxima verossimilhança de informação limitada independe de qual é a variável escolhida para ser colocada no lado esquerdo na equação estrutural, fato este que não ocorre com os demais métodos de estimação até aqui apresentados. Esta propriedade decorre do fato já assinalado anteriormente que a estimativa do parâmetro γ é determinada a menos de uma constante de proporcionalidade.

Estimador de Máxima Verossimilhança de Informação Limitada: Enfoque de Minimização da Razão da Variância

A equação (44) possibilita escrever μ do seguinte modo:

$$\mu = \frac{\gamma' Y_c\,'\,M_1\,Y_c\,\gamma}{\gamma' Y_c\,'\,M\,Y_c\,\gamma}$$

Esta expressão pode ser interpretada de uma maneira bastante interessante. Com efeito, como

(47) $Y_c\,\gamma = X_1\,\beta_1 + \mu_1$

Se γ fosse conhecido, a soma dos quadrados dos resíduos desta regressão seria igual a :

$$SQR_1 = \gamma'\,Y_c\,'\,M_1\,Y_c\,\gamma$$

Por outro lado, se acrescentássemos a matriz X_2 nos regressores da equação (47),

$$Y_c \ \gamma = Y_1 \ \beta_1 \ + \ X_2 \ \beta_2 \ + u_1 \, ,$$

a soma dos quadrados dos resíduos, na suposição de um valor conhecido de γ, desta regressão seria igual a:

$$SQR_2 \ = \gamma' \ Y_c \ ' \ M \ Y_c \ \gamma$$

O valor de μ é justamente igual à razão dessas duas somas de quadrados de resíduos:

$$\mu = \frac{SQR_1}{SQR_2} = \frac{\gamma' Y_c \ ' \ M_1 \ Y_c \ \gamma}{\gamma' \ Y_c \ ' \ M \ Y_c \ \gamma}$$

Como o vetor de coeficientes β_2 é igual a zero, uma idéia de certo modo intuitiva é obter o valor de γ que minimiza a razão das somas dos quadrados dos resíduos, de sorte a se obter o menor valor para μ. Consequentemente, derivando-se μ em relação a γ,

$$\frac{\partial \mu}{\partial \gamma} = \frac{(\ \gamma' \ Y_c \ ' \ M \ Y_c \ \gamma) \ 2 \ \gamma_c \ ' \ M_1 \ Y_c \ \gamma - (\ \gamma' \ Y_c \ ' \ M_1 \ Y_c \ \gamma) \ 2 \ Y_c \ ' \ M \ Y_c \ \gamma}{(\ \gamma' Y_c \ ' \ M \ Y_c \ \gamma)^2}$$

e igualando-se este resultado a zero, obtém-se:

$$(Y_c \ ' \ M_1 \ Y_c \ - \mu \, Y_c \ ' \ M \ Y_c \ ) \ \gamma = 0$$

que nada mais é do que a equação (44). Uma vez determinado o valor de γ a estimativa de β_1 é facilmente calculada através de:

$$\hat{\beta}_1 = (\ X_1 \ ' \ X_1 \ )^{-1} \ X_1 \ ' Y_c \ \gamma$$