UNIVERSIDADE FEDERAL DE JUIZ DE FORA - *CAMPUS* Governador Valadares
ECONOMETRIA ESPACIAL
Prof. Vinícius de Azevedo Couto Firme

Econometria Espacial:

Capítulo 6 – Estimando a Dependência Espacial

Estrutura da Apresentação:

1. Introdução
2. Os problemas do EMQO em modelos espaciais
3. Estimando o Modelo SAR (MV ou VI)
4. Estimando o Modelo SEM (MV ou MGM)
5. Estimando o Modelo SMA (MV ou MGM)
6. Estimando o Modelo SAC (MV ou MQ2EE)
7. Estimando o Modelo SDM (MV ou VI)
8. Estimando outros Modelos Espaciais
9. Medidas de Qualidade do Ajuste

1. Introdução

- No capítulo anterior vimos que ignorar a dependência espacial pode gerar sérios problema às estimações (ineficiência, viés e inconsistência).

- Assumindo que selecionamos o modelo espacial correto (SAR, SEM, etc) veremos quais estimadores devem ser utilizados em cada caso.

Nota 1: A seleção do modelo adequado envolve: a) a análise do *I de Moran* e dos multiplicadores de *Lagrange*, para ver se $\rho \neq 0$ e/ou $\lambda \neq 0$; b) testes de significância sobre WX, para ver se $\tau \neq 0$.
Nota 2: Geralmente, os modelos são estimados com diferentes matrizes espaciais (W), usando os critérios AIC ou SC para definir o mais adequado.

- Como será visto, o EMQO não é uma boa opção para as estimações espaciais que contenham $\rho \neq 0$, $\lambda \neq 0$ ou $\gamma \neq 0$. Nestes casos, deve-se optar pelo:

a) Estimador de Máxima Verossimilhança (MV): útil apenas quando há normalidade dos resíduos (característica incomum em pequenas amostras).

b) Método Generalizado dos Momentos (MGM): mais leve computacionalmente que o MV, é útil quando não há normalidade dos resíduos.

2. Os problemas de estimar por MQO

- Dentre os problemas associados ao MQO, pode-se destacar:

a) O EMQO tende a gerar estimativas enviesadas do ρ.

Considere o modelo SAR puro: $y = \rho \mathrm{W}\mathrm{y} + \varepsilon$ (1)

Como o EMQO para o β da função $y = \mathrm{X}\beta + \varepsilon$ será: $\hat{\beta} = (X'X)^{-1}X'y$ (2)

A estimativa de ρ, via EMQO, da Eq. 1 é: $\hat{\rho} = [(Wy)'(\mathrm{Wy})]^{-1}(Wy)'y$ (3)

Substituindo (1) em (3): $\hat{\rho} = \{[(Wy)'(\mathrm{Wy})]^{-1}(Wy)'\}\rho \mathrm{Wy} + \varepsilon$ (4)

Aplicando distributiva:
$\hat{\rho} = [(Wy)'(\mathrm{Wy})]^{-1}[(Wy)'\mathrm{Wy}]\rho + [(Wy)'(\mathrm{Wy})]^{-1}[(Wy)'\varepsilon]$ (5)

Logo: $\hat{\rho} = \rho + [(Wy)'(\mathrm{Wy})]^{-1}[(Wy)'\varepsilon]$ (6)

2. Os problemas de estimar por MQO

Relembrando as 2 propriedades dos operadores de esperança abaixo:

$$E(bX|X) = bE(X|X) = bX \quad (7)$$

Nota: Com o valor de X é possível saber o valor de bX [Ex: $E(2X|X = 3) = 6$].

Além disso, se Y= $f(X)$ é possível mostrar que (*lei das expectativas iteradas*):

$$E[E(Y|X)] = E(Y) \quad (8)$$

Nota: Se $Y = bX$, tem-se que: $\mathrm{E}[E(Y|X)] = E[E(bX|X)] = E(bX) = E(Y)$.

É possível mostrar que a esperança, condicional em Wy, da expressão (6) é:

$$E(\hat{\rho}|Wy) = \mathrm{E}(\rho|Wy) + E\{[(Wy)'(\mathrm{Wy})]^{-1}[(Wy)'\varepsilon]|Wy\} \quad (9)$$

Aplicando a propriedade (7) em (9):

$$\hat{\rho} = \rho + [(Wy)'(\mathrm{Wy})]^{-1}Wy'E(\varepsilon|Wy) \quad (10)$$

2. Os problemas de estimar por MQO

- Com base em (10) nota-se que $\hat{\rho} = \rho$ (sem viés) apenas se $E(\varepsilon|Wy) = 0$.

Nota: $E(\varepsilon|Wy) = 0$ implica não haver nenhuma associação linear ou não linear entre ε e Wy (hipótese mais forte que a correlação – apenas linear).

- Como a forma reduzida da Eq. (1) é:

$$y = (I - \rho \mathrm{W})^{-1}\varepsilon \qquad (11)$$

- O $E(\varepsilon|Wy) \neq 0$. Logo, o EMQO produzirá estimativas enviesadas de ρ.

Nota: também haveria viés nos $\hat{\beta}$s em um modelo SAR misto (com $X\beta$). Neste caso, se $E(\varepsilon|Wy) \neq 0$, haverá viés no $\hat{\rho}$ e, portanto, em todos os parâmetros, pois: $\frac{\partial y}{\partial X_k} = (I - \rho W)^{-1}\beta_k$.

b) <u>O EMQO tende a gerar estimativas inconsistentes para ρ</u>.

- As propriedades assimptóticas do EMQO garantem que a distribuição dos $\hat{\beta}$ e das estatísticas t e F convergem para uma distribuição normal quando a amostra tende ao infinito (*lei dos grandes números*).

2. Os problemas de estimar por MQO

- Para que o EMQO a-espacial não seja enviesado, é preciso que $E(\varepsilon_i|X) = 0$ (Hip. 3 de Gauss-Markov – ver Slide 12 do Cap. 1).

- Contudo, quando a amostra de um EMQO tende ao infinito, basta que a $\text{cov}(\varepsilon_i, X) = 0$ ou a $\text{corr}(\varepsilon_i, X) = 0$ para que não exista viés (consistente).

- Formalmente, valendo-se das propriedades assintóticas do EMQO, podemos usar os *limites da probabilidade* para achar o $\hat{\beta}$ a-espacial, quando $n \to \infty$:

$$plim(\hat{\beta}) = plim[(X'X)^{-1}X'y] = plim[(X'X)^{-1}X'(X\beta + \varepsilon)] \quad (12)$$

Logo: $plim(\hat{\beta}) = \beta + plim[(X'X)^{-1}X'\varepsilon)] = \beta + cov(X, \varepsilon)/var(X, X)$ (13)

Caso a $cov(X, \varepsilon) = 0$, $plim(\hat{\beta}) = \beta$ (consistente quando $n \to \infty$).

Nota: A hip. de $E(\varepsilon_i|X) = 0$ é mais forte, pois qualquer associação (linear ou não) causaria viés. Na $\text{cov}(\varepsilon_i, X) = 0$, só associações lineares causariam viés.

2. Os problemas de estimar por MQO

- Voltando ao modelo SAR, como Wy é endógeno e afeta ε via multiplicador espacial (ver Eq. 11) a $\text{cov}(\varepsilon_i, Wy) \neq 0$. Tornando o EMQO inconsistente.

Nota: O parâmetro de uma variável específica k, isto é $\hat{\beta}_k$, pode até ser consistente se $\text{cov}(\varepsilon_i, X_k) = 0$. Todavia, o cálculo do seu efeito total ficaria comprometido, pois: $\frac{\partial y}{\partial X_k} = (I_n - \rho W)^{-1} I_n \beta_k$.

c) <u>O EMQO será ineficiente na presença de ACS associada aos resíduos</u>.

- O EMQO será não enviesado porém ineficiente nos Modelos SEM ($W\xi$) ou SMA ($W\varepsilon$). Nestes modelos, a matriz Var-Cov do erro (ξ) é:

SEM (ver Eq. 32 dos Slides – Cap. 5)

$$E(\xi'\xi) = \sigma^2[(I_n - \lambda W)'(I_n - \lambda W)]^{-1} \quad (14a)$$

SMA (ver Eq. 48 dos Slides – Cap. 5):

$$E(\xi'\xi) = \sigma^2[(I_n - \gamma W)'(I_n - \gamma W)] = \sigma^2[I_n + \gamma(W + W') + \gamma^2 W'W] \quad (14b)$$

- Em ambos os casos, há dependência espacial e heterocedasticidade associada aos resíduos. Portanto o EMQO seria ineficiente.

3. Estimando o Modelo SAR (Wy)

- No SAR, os resíduos obtidos via EMQO apresentariam matriz Var-Cov plena [efeito global: $(I_n - \rho W)^{-1}\varepsilon$]. Neste caso, recomenda-se a Estimação via Máxima Verossimilhança (MV) ou Variáveis Instrumentais (VI).

3.1 O Estimador de Máxima Verossimilhança (MV)

- O EMV permite estimar modelos com respostas lineares e não lineares (Ex.: Logit e Probit). Além disso, a heterocedasticidade é incluída automaticamente no EMV, possibilitando ganhos de eficiência.

- De modo geral, o EMV é o mais assintoticamente eficiente quando o modelo populacional é corretamente especificado (ver Eq. 27). Para tanto, é preciso definir uma distribuição particular para as variáveis aleatórias (Ex.: Normal, Poisson, etc).

Nota: o EMV não é robusto à falhas nas hipóteses de distribuição das variáveis.

3. Estimando o Modelo SAR (Wy)

- Diferentemente do MQO, onde o $\hat{\beta}$ minimiza a soma dos quadrados dos resíduos de uma regressão (Figura 1), o EMV maximiza a probabilidade de uma observação ocorrer (Ex.: qual altura tem maior probabilidade de ocorrer na Fig. 2?).

- Ex. (Modelo incondicional): busca-se estimar a altura média (β) de pessoas com base numa amostra (Y), assumindo que $\mathrm{Y} \sim normal$ (neste caso, β e σ^2 tornam-se parâmetros do modelo). Assim, o EMV estimará o valor de β (ou da σ^2) que maximiza a probabilidade de uma altura específica ocorrer.

Figura 1. Valores previstos ($\hat{y}_i$) e resíduos ($\hat{u}_i$)

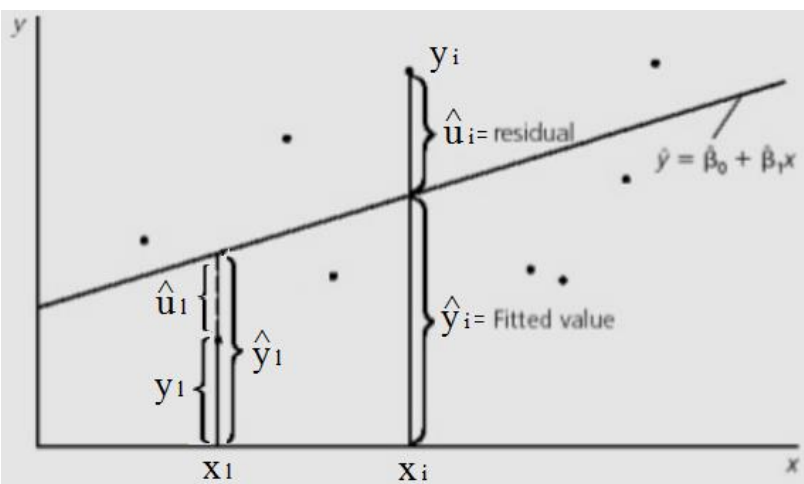
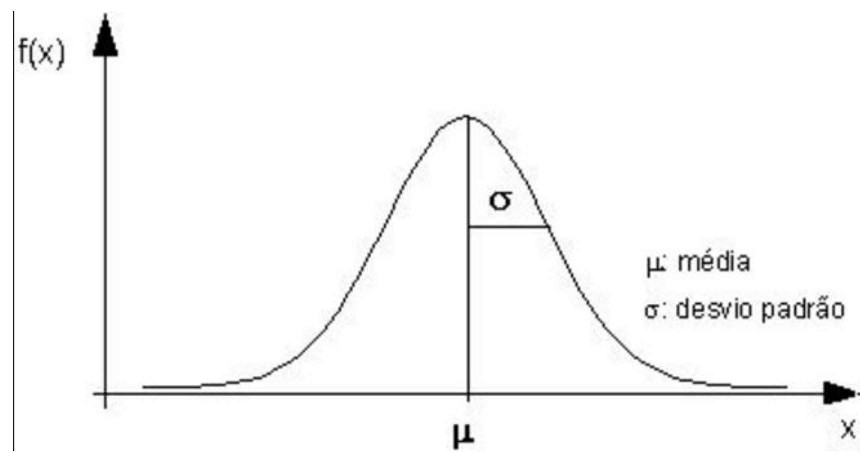


Figura 2. Distribuição Normal ($\beta = \mu$)



3. Estimando o Modelo SAR (Wy)

Nota 1: O fato de $\mathrm{Y}{\sim}normal$ implica na determinação conjunta de β e σ^2 via EMV (ver Eq. 15). Contudo, se $Y{\sim}Bernoulli$ (p/exemplo), apenas β seria necessário.

Nota 2: Caso $\varepsilon{\sim}\mathrm{Normal}(0, \sigma^2 I_n)$ e $\mathrm{Y}{\sim}\mathrm{Normal}(\mathrm{X}\beta, \sigma^2 I_n)$, ver Hip. 6 (Cap. 1), o EMQO e o EMV serão não enviesados, porém o MQO será mais eficiente.

- Assumindo que $\mathrm{Y}{\sim}Normal(\beta, \sigma^2)$ e é *iid**, podemos estabelecer a seguinte <u>função densidade de probabilidade (FDP)</u> para cada variável y_i:

$$f(\beta, \sigma^2, y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y_i - \beta}{\sigma}\right)^2} \quad (15)$$

<u>Onde:</u> $y_1, \ldots, y_n = Y$; Neste caso, β é o escalar que mede a média de Y (é o coeficiente que queremos obter!).

Nota 3: A Eq. 15 representa a FDP de um modelo incondicional ($\cong$ ao MQO apenas com a constante). Em um modelo condicional $f(\beta, \sigma^2, y_i|X)$, o expoente da Eq. 15 seria $-\frac{1}{2}\left(\frac{y_i - X\beta}{\sigma}\right)^2$.

* $X_i, \ldots, X_n$ são independentes e identicamente distribuídos (*iid*) – veremos que esta hipótese não é válida no modelo SAR e em outros modelos espaciais.

3. Estimando o Modelo SAR (Wy)

Lembrando que $\sqrt{1/a} = (1/a)^{1/2} = a^{-1/2}$ e $e = exp$ (ver nota), temos:

$$f(\beta, \sigma^2, y_i) = (2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}} exp\left[-\frac{1}{2}\left(\frac{y_i - \beta}{\sigma}\right)^2\right] \qquad (16)$$

- Fazendo o produtório da função (16) para todas as n variáveis $(y_i, \ldots, y_n = Y)$, encontramos a <u>Função de Verossimilhança</u>:

$$V(\beta, \sigma^2, Y) = \prod_{i=1}^{n}\left\{(2\pi)^{-\frac{1}{2}}(\sigma^2)^{-\frac{1}{2}} exp\left[-\frac{1}{2}\left(\frac{y_i - \beta}{\sigma}\right)^2\right]\right\} \qquad (17)$$

Lembrando que: $\left(a^b * a^b = a^{2b}\right)$ e que $\left(a^b * a^b * \cdots * a^b = a^{nb}\right)$, temos que:

$$V(\beta, \sigma^2, Y) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^{n}\left\{exp\left[-\frac{1}{2}\left(\frac{y_i - \beta}{\sigma}\right)^2\right]\right\} \qquad (18)$$

Lembrando que $e = exp$ e que $e^{2a} * e^{2a} * \cdots * e^{2a} = e^{2\sum_{i=1}^{n} a} = e^{2na}$, temos:

$$V(\beta, \sigma^2, Y) = (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} \left\{exp\left[-\frac{1}{2}\sum_{i}^{n}\left(\frac{y_i - \beta}{\sigma}\right)^2\right]\right\} \qquad (19)$$

Nota: e^x ou $\exp(x)$ é a *função exponencial natural*, cuja base é o nº. de Euler (≅ 2,718281828). Adicionalmente, tem-se que $\frac{de^x}{dx} = e^x$ e $ln(e^x) = x$.

3. Estimando o Modelo SAR (Wy)

Aplicando logaritmo em (19) para obter a função de Log-Verossimilhança:

$$ln[V(\beta,\sigma^2,Y)] = ln\left\{(2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}}exp\left[-\frac{1}{2}\sum_i^n\left(\frac{y_i-\beta}{\sigma}\right)^2\right]\right\} \qquad (20)$$

Lembrando que $\ln\left(a^{-b} * c^d\right) = -bln(a) + \mathrm{dln(c)}$ e $\ln[exp(a)] = a$, temos:

$$ln[V(\beta,\sigma^2,Y)] = \left[-\frac{n}{2}ln(2\pi)\right] + \left[-\frac{n}{2}\ln(\sigma^2)\ \right] + \left[-\frac{1}{2\sigma^2}\sum_i^n(y_i-\beta)^2\right] \qquad (21)$$

Nota: $\left[-\frac{1}{2\sigma^2}\sum_i^n(y_i-\beta)^2\right] = \left[-\frac{1}{2\sigma^2}\sum_i^n(\varepsilon_i)^2\right] = \left[-\frac{\varepsilon\prime\varepsilon}{2\sigma^2}\right]$, onde ε é o termo de erro do EMV.

- A Eq. 21 é a função de Log-Verossimilhança do EMV. Para obtermos $\hat{\beta}$ é necessário maximizá-la (isto é maximizar a probabilidade de $\hat{\beta}$ ocorrer). Para tanto, basta fazer $\partial\ln(V)/\partial\beta = 0$.

Usando a regra da cadeia ,[lembrete: $\frac{\partial(a+bx)^2}{\partial x} = 2(a+bx)b$], temos:

$$\frac{\partial\ \ln(V)}{\partial\beta} = -\frac{1}{2\sigma^2}2\sum_i^n(y_i-\beta)\,(-1) = \frac{\sum_i^n(y_i-\beta)}{\sigma^2} \qquad (22)$$

3. Estimando o Modelo SAR (Wy)

Igualando a Eq. (22) a zero e assumindo que β não varia, temos:

$$\sum_i^n (y_i - \beta) = 0 \quad \rightarrow \quad n\beta = \sum_i^n y_i \quad \rightarrow \quad \hat{\beta} = \frac{\sum_i^n y_i}{n} \tag{23}$$

- Logo, o $\hat{\beta}$ estimado via EMV equivale à média de Y (voltar à Figura 2) e, portanto, é igual ao EMQO (apenas com constante).

Nota: a derivada de segunda ordem da Eq. (22) é: $\frac{d^2 Ln(V)}{d\beta^2} = \frac{d^2[(\sum_i^n y_i/\sigma^2) - (n\beta/\sigma^2)]}{d\beta^2} = -\frac{n}{\sigma^2}$. Portanto, a função $\ln(V)$, Eq. (21), é negativa (*i.e.*: tem formato de "U" invertido) e tem ponto de máximo.

- Para obter $\hat{\sigma}^2$ basta fazer a $\partial \ln(V)/\partial \sigma^2 = 0$. Onde $\ln(V) = \text{Eq. } 21$.

Dica: a) a derivada está em função de σ^2. Trate $\sigma^2 = k$ e faça $\partial \ln(V)/\partial k$); b) Lembre-se: $\frac{\partial \ln(k)}{\partial k} = \frac{1}{k}$.

$$\frac{\partial \ln(V)}{\partial(\sigma^2)} = -\frac{n}{2(\sigma^2)} - \frac{\sum_i^n (y_i - \beta)^2}{2} \left\{ \frac{\partial\left[(\sigma^2)^{-1}\right]}{\partial(\sigma^2)} \right\} = -\frac{n}{2\sigma^2} + \frac{\sum_i^n (y_i - \beta)^2}{2\sigma^4} \tag{24}$$

Igualando a Eq. (24) a zero: $\frac{n}{2\sigma^2} = \frac{\sum_i^n (y_i - \beta)^2}{2\sigma^4} \quad \rightarrow \quad \hat{\sigma}^2 = \frac{\sum_i^n (y_i - \beta)^2}{n}$ (25)

3. Estimando o Modelo SAR (Wy)

- A demonstração até aqui (Eq. 15 a 25) focou em modelos incondicionais $[E(Y)]$. Contudo, os $\hat{\beta}$ condicionais $[E(Y|X)]$, gerados via EMV e EMQO, continuarão idênticos se ambos os estimadores forem não viesados.

- Os $\hat{\beta}$ e a $\hat{\sigma}^2$ oriundo do EMV condicional, $f(\beta, \sigma^2, y_i|X)$, serão:

$$\hat{\beta}_{MV} = (X'X)^{-1}X'y \quad (26)$$

$$\hat{\sigma}^2_{MV} = \frac{\sum_i^n (y_i - X\beta)^2}{n} \quad \text{ou} \quad \hat{\sigma}^2_{MV} = \frac{\varepsilon'\varepsilon}{n} \quad (27)$$

Dica: para achar Eq. 26 ($\hat{\beta}_{MV}$ do modelo condicional), basta fazer $\partial \ln(V)/\partial\beta = 0$, usando o $\ln(V)$ da Eq. 21 com o último termo sendo: $\left[-\frac{1}{2\sigma^2}\sum_i^n (y_i - X\beta)^2\right]$.

- Perceba que, sob hipóteses de não enviesamento, $\hat{\beta}_{MV} = \hat{\beta}_{MQO}$. Além disso, $\hat{\sigma}^2_{MV} \to 0$ quando $n \to \infty$. Logo, o EMV é assintoticamente eficiente.

- No contexto espacial, o EMV se torna complexo pois a dependência espacial faz com que os dados não sejam *iid* (ver Eq. 15). Logo, não é possível realizar o produtório das FDPs tradicionalmente (Eq. 17).

3. Estimando o Modelo SAR (Wy)

- Neste caso, é necessário maximizar diferentes funções de Log-Verossimilhança (Eq. 21) – via integração multi-dimensional (mais complexo!).

- Ainda assim, o EMV espacial continua assintoticamente consistente, eficiente e com normalidade assintótica (importante para estatísticas t e F) se:

a) A Função de Verossimilhança existe - $V(\beta, \sigma^2, y_i|X)$.

b) A derivadas de 1ª, 2ª e 3ª ordem de $V(\beta, \sigma^2, y_i|X)$ em relação à β existem.

c) A derivada de 1ª ordem de $V(\beta, \sigma^2, y_i|X)$ em relação à β é: $\frac{\partial v\left(\beta, \sigma^2, y_i \middle| X\right)}{\partial \beta} < \infty.$

d) A matriz var-cov do Modelo Gerador de Dados é não singular (tem inversa!).

Nota: o EMV requer que $n \to \infty$ (amostra grande) para que as propriedades assintóticas de normalidade sejam válidas.

3. Estimando o Modelo SAR (Wy)

- Neste momento, estamos aptos a analisar o EMV para o modelo SAR:

$$y = \rho Wy + X\beta + \varepsilon \qquad\qquad \varepsilon \sim normal(0, \sigma^2 I_n) \qquad (28)$$

- Assumindo $\varepsilon \sim normal$, a função de log-verossimilhança da Eq. (28) será:

$$ln[V(\beta, \sigma^2, Y)] = \left[-\frac{n}{2} ln(2\pi)\right] + \left[-\frac{n}{2} \ln(\sigma^2)\ \right] + \left[\sum_i^n ln\,|I_n - \rho W| - \frac{\varepsilon\prime\varepsilon}{2\sigma^2}\right] \qquad (29)$$

Onde: $\varepsilon = (y - \rho Wy - X\beta)$

- A diferença fundamental entre Eq. (29) e Eq. (21), modelo SAR e a-espacial (incondicional), respectivamente, é que a Eq. (29) envolve o cálculo do determinante da matriz Jacobiana $|I_n - \rho W|$, de dimensão nxn.

Como o Jacobiano contém as derivadas entre $\frac{\partial \varepsilon_i}{\partial y_i}$ (para $i = j$ e $i \neq j$), tem-se:

$$\det\left(\frac{\partial \varepsilon_i}{\partial y_i}\right) = det\left(\frac{\partial(y - \rho Wy - X\beta)}{\partial y}\right) = det\left(\frac{\partial y(I - \rho W)}{\partial y} - \frac{\partial X\beta}{\partial y}\right) = det(I - \rho W) \qquad (30)$$

Nota: assumindo exogeneidade de X, tem-se que $(\partial X\beta/\partial y) = 0$.

3. Estimando o Modelo SAR (Wy)

- Ao ignorar o Jacobiano, o MQO se torna inadequado para estimar $\hat{\beta}$ em alguns modelos espaciais. Em síntese, o EMQO minimizaria a soma dos quadrados dos resíduos do SAR abaixo:

$$\varepsilon'\varepsilon = (y - \rho Wy - X\beta)'(y - \rho Wy - X\beta) \qquad (31)$$

- Contudo, a Eq. (31), via EMQO, contém apenas parte da Eq. (29), via EMV.

- A maximização do Log de verossimilhança do SAR, Eq. (29), para obter $\hat{\beta}$, envolve o seguinte cálculo:

$$\frac{\partial \ln(V)}{\partial \beta} = \frac{\partial\left[\sum_i^n ln|I_n - \rho W| - \frac{(y-\rho Wy - X\beta)'(y-\rho Wy-X\beta)}{2\sigma^2}\right]}{\partial \beta} = 0 \qquad (32)$$

Assumindo que $(y - \rho Wy - X\beta)'(y - \rho Wy - X\beta) = (y - \rho Wy - X\beta)^2$:

$$\frac{\partial \ln(V)}{\partial \beta} = -2\frac{(y-\rho Wy - X\beta)}{2\sigma^2}(-X) = \frac{X(y-\rho Wy-X\beta)}{\sigma^2} = 0 \qquad (33)$$

3. Estimando o Modelo SAR (Wy)

Multiplicando ambos os lados por σ^2 e aplicando distributiva:

$$X'y - \rho X'Wy - X'X\beta = 0 \quad \rightarrow \quad \hat{\beta}_{MV} = (X'X)^{-1} X'y - \rho(X'X)^{-1}X'Wy \quad (34)$$

Colocando $(X'X)^{-1} X'$ e depois y em evidência, a Eq. (34) se torna:

$$\hat{\beta}_{MV} = [(X'X)^{-1} X'](y - \rho Wy) \quad \rightarrow \quad \hat{\beta}_{MV} = [(X'X)^{-1} X'](I - \rho W)y \quad (35)$$

- O 1º termo da Eq. (34) é o EMQO tradicional (o 2º não faz parte do EMQO e mede o viés, que pode ser - ou +, visto que $-1 < \rho < 1$). A Eq. (35) mostra que o viés do EMQO é causado justamente pelo termo Jacobiano $(I - \rho W)$.

- Analogamente, a estimativa de $\hat{\sigma}^2$ implica em (rever Eq. 27):

$$\frac{\partial \log(V)}{\partial(\sigma^2)} = 0 \quad \rightarrow \quad \hat{\sigma}^2_{MV} = \frac{\varepsilon'\varepsilon}{n} \quad \rightarrow \quad \hat{\sigma}^2_{MV} = \frac{(y - \rho Wy - X\beta)'(y - \rho Wy - X\beta)}{n} \quad (36)$$

- A $\hat{\sigma}^2_{MV}$ é enviesada em pequenas amostras (denominador não considera número de parâmetros estimados, k) mas é assintoticamente eficiente.

3. Estimando o Modelo SAR (Wy)

- Como a matriz Jacobiana $|I_n - \rho W|$ é plena (não há elementos nulos), a maximização necessária ao EMV exige grande capacidade computacional.
- Logo, Ord (1975) sugere substituir W, na Eq. (29) por seus autovalores (ω). Assim, não seria necessário calcular diversos determinantes. Contudo a precisão desta técnica diminui a medida que a dimensão de W aumenta.

Dica: para facilitar computacionalmente a estimação via MV recomenda-se o uso de matrizes espaciais esparsas (contendo muitos zeros). Ex: Rainha, Torre, K vizinhos mais próximos...

Nota: Qualquer autovalor (ω_{1x1}) e autovetor (V_{nx1}) da matriz W_{nxn} permite que façamos uma aproximação da matriz W (mantendo-se a proporcionalidade) via $WV = \omega V$. Para tanto:

Passo 1: calcula-se o $DET(W - \omega I) = 0$ e resolve-se o polinômio resultante p/ obter ω (**autovalor**).
Passo 2: resolve-se a expressão $(W - \omega I)V = 0$ para obter V (**autovetor**).

Exemplo: Seja $W = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$. Logo, $DET\left(\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \omega \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = 0$ será $\omega^2 - 4\omega - 5 = 0$. $\begin{cases} \omega_1 = 5 \\ \omega_2 = -1 \end{cases}$

Logo, o $V_1 = \begin{bmatrix} x \\ y \end{bmatrix}$, de ω_1, é: $\left(\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}\right) V_1 = 0$. Resolvendo o sistema, tem-se que $x = y/2$.

Assim, se $x = 1$, por exemplo, $V_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ e $WV = \omega V = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$. O mesmo pode ser feito c/ $\omega_2 \rightarrow V_2$.

3. Estimando o Modelo SAR (Wy)

- A maximização da função de Log-Verossimilhança do SAR é complexa e não linear. Logo, o EMV costuma ser realizado em 5 passos (Arbia, 2006):

a) Regride-se (via MQO) y contra o conjunto de variáveis explicativas (X):

$$y = X\beta_o + \varepsilon_o \tag{37}$$

b) Regride-se (via MQO) Wy contra o conjunto de variáveis explicativas (X):

$$Wy = X\beta_L + \varepsilon_L \tag{38}$$

c) Obtém-se os valores de y e Wy que não estão associados à X (isto é: $\hat{\varepsilon}_o$ e $\hat{\varepsilon}_L$):

$$\hat{\varepsilon}_o = y - X\hat{\beta}_o \quad \text{e} \quad \hat{\varepsilon}_L = Wy - X\hat{\beta}_L \tag{39}$$

d) Substitui-se $(y - X\beta) \rightarrow \hat{\varepsilon}_o$ e $Wy \rightarrow \hat{\varepsilon}_L$ na função de log-verossimilhança (Eq. 29) e maximiza-se a expressão resultante em função de ρ para obter $\hat{\rho}$:

$$ln[V] = \left[-\frac{n}{2} ln(2\pi)\right] + \left[-\frac{n}{2}\ln(\sigma^2)\ \right] + \left[\sum_i^n ln\,|I_n - \rho W| - \frac{(\hat{\varepsilon}_o - \rho\hat{\varepsilon}_L)'(\hat{\varepsilon}_o - \rho\hat{\varepsilon}_L)}{2\sigma^2}\right] \tag{40}$$

Nota: Na Eq. 29, $\varepsilon = (y - \rho Wy - X\beta) \rightarrow \varepsilon = (y - X\beta - \rho Wy) \rightarrow \varepsilon = (\hat{\varepsilon}_o - \rho\hat{\varepsilon}_L)$.

e) Usa-se $\hat{\rho}$ para obter as seguintes estimativas de σ^2 e β para o EMV:

$$\hat{\sigma}^2_{MV} = \frac{1}{n}(\hat{\varepsilon}_o - \rho\hat{\varepsilon}_L)'(\hat{\varepsilon}_o - \rho\hat{\varepsilon}_L) \quad \text{e} \quad \hat{\beta}_{MV} = \hat{\beta}_o - \hat{\rho}\hat{\beta}_L \tag{41}$$

3. Estimando o Modelo SAR (Wy)

Nota: a $\hat{\sigma}^2_{MV}$ e o $\hat{\beta}_{MV}$ da Eq. (41) originam-se das Eq. (36) e Eq. (34), respectivamente.

- O EMV é recomendado quando $\varepsilon \sim normal$. Contudo, caso $n \to \infty$, é possível estimar via EMV mesmo sem garantir a normalidade de ε. Neste caso o estimador é denominado *Quase Maxima verossimilhança* (EQMV).

- O EQMV é assintoticamente consistente, eficiente e seus resultados convergem para uma distribuição normal (útil para estatísticas t e F). Contudo, estas garantias não se aplicam à pequenas amostras.

3.2 O Método das Variáveis Instrumentais (VI)

- A Hip. 3 de Gauss-Markov diz que $E(\varepsilon_i|X) = 0$, caso contrário o EMQO será enviesado (Slide 12 – Cap. 1).

- Contudo, esta hipótese costuma falhar devido a: **a)** Omissão de variáveis relevantes; **b)** Erro de Mensuração; **c)** Endogeneidade (ou simultaneidade).

3. Estimando o Modelo SAR (Wy)

- Como a variável Wy do modelo SAR afeta e é afetada por y, temos um problema de endogeneidade (*simultaneidade*).

Ex.: Meus vizinhos me afetam. Mas eu sou vizinho dos meus vizinhos. Logo, eu afeto meus vizinhos.

- Tal problema é semelhante a um modelo de Equações simultâneas, conforme o exemplo abaixo:

$$y = \rho Wy + X\beta + \varepsilon \tag{41}$$
$$Wy = \alpha y + WX\varphi + \mu \tag{42}$$

- Como nosso interesse é estimar o y da Eq. (41), devemos expurgar a endogeneidade existente em Wy.

- Mesmo se $E(\varepsilon_i|X) = 0$, o $E(\varepsilon_i|Wy) \neq 0$ (parte da Eq. 42, que afeta Wy e não consta na Eq. 41, fará parte de ε). Logo, é necessário utilizar um "instrumento" para a variável Wy, denominado z_i, a fim de assegurar o verdadeiro impacto de Wy sobre y.

3. Estimando o Modelo SAR (Wy)

- De modo geral, um bom instrumento precisa ter as seguintes características:

a) $Cov(z, \varepsilon) = 0$ (Propriedade da Exogeneidade) (43)

b) $Cov(z, Wy) \neq 0$ (Propriedade da Relevância) (44)

- Infelizmente, apenas a Eq. 44 (relevância) é testável. Para tanto, é necessário que $\pi_1 \neq 0$ na regressão abaixo:

$$Wy = \pi_0 + \pi_1 z + \epsilon \quad (45)$$

Nota: quanto à exogeneidade, deve-se buscar justificativas teóricas para a $Cov(z, \varepsilon) = 0$.

- Caso z seja considerado um bom instrumento para x, realiza-se a estimação por VI de forma análoga ao EMQO. Isto é, supondo que $X = x$ (vetor nx1):

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{MQO}} = (x'x)^{-1}x'y \quad \text{e} \quad \widehat{\boldsymbol{\beta}}_{\boldsymbol{VI}} = (z'x)^{-1}z'y \quad (46)$$

- Onde o $\hat{\beta}_{VI}$ é um estimador consistente do verdadeiro β.

3. Estimando o Modelo SAR (Wy)

- A $\hat{\sigma}^2_{VI}$ será assintoticamente normal e eficiente. Porém, sempre será maior que a $\hat{\sigma}^2_{MQO}$. Isto ocorre porque a medida de ajuste entre x e z é $R^2_{x,z} < 1$:

$$\hat{\sigma}^2_{MQO} = \sigma^2/(SQT_x) \quad \text{e} \quad \hat{\sigma}^2_{VI} = \sigma^2/(SQT_x * R^2_{x,z}) \tag{47}$$

Onde: SQT_x é a soma dos quadrados totais de x: $SQT_x = \sum_{i=1}^{n}(x_i - \bar{x})^2$. Logo, a $\hat{\sigma}^2_{VI}$ se aproximará da $\hat{\sigma}^2_{MQO}$ quando $R^2_{x,z} \rightarrow 1$ (se $z = x$, $R^2_{x,z} = 1$ e $\hat{\sigma}^2_{MQO} = \hat{\sigma}^2_{VI}$).

Nota: o R^2_{MQO} é, sempre, maior que o R^2_{VI} (pois o EMQO minimiza a soma dos quadrados dos resíduos). O importante é verificar se o $\beta_x^{MQO} \neq \beta_x^{VI}$ (Teste Hausman). Caso sejam diferentes, haverá indícios de que x é, de fato, endógena.

- Os instrumentos recomendados para Wy no modelo SAR são: **a)** Wx ou W^2x (Kelejian e Robinso, 1992; Kelejian e Prucha, 1998); **b)** $W(I - \hat{\rho}W)^{-1}X\hat{\beta}$ (Lee, 2003).

3. Estimando o Modelo SAR (Wy)

- Justificativa para usar WX como instrumento de Wy: Se X for uma matriz de variáveis exógenas, isto é $Cov(X, \varepsilon) = 0$, é razoável supor que WX também é: $Cov(WX, \varepsilon) = 0$. Assim, se $Cov(WX, Wy) \neq 0$, o EVI será consistente pois:

$$\rho \lim(\hat{\rho}_{VI}) = \rho + \left(\frac{Corr(WX,\varepsilon)}{Corr(WX,Wy)} * \frac{\sigma_\varepsilon}{\sigma_{Wy}}\right) \quad \rightarrow \quad \rho \lim(\hat{\rho}_{VI}) = \rho \tag{48}$$

- Se o instrumento $z = WX$ tiver $Cov(z, \varepsilon) \neq 0$, o EVI é inconsistente. Contudo, se $Cov(z, \varepsilon) > Cov(z, Wy)$ o viés assintótico do EVI aumentará a medida que $n \rightarrow \infty$. Logo, assegure-se de que z é bastante correlacionado com Wy.

- O procedimento do EVI espacial envolve 2 passos (Nota: quando usamos mais de uma variável como instrumento de WY, como é caso da Eq. 49, o EVI se torna um EMQ2E):

a) Expurgar a endogeneidade de Wy usando o(s) instrumento(s) (ex: WX) e as variáveis explicativas exógenas da Eq. 41. Para tanto, basta regredir:

$$Wy = X\varphi + \mathrm{WX}\pi + \zeta \tag{49}$$

Nota: como $Cov(X, \varepsilon) = Cov(WX, \varepsilon) = 0$ e o $\widehat{Wy}$, estimado na Eq. (49), é uma combinação linear de X e WX, espera-se que $Cov(\widehat{Wy}, \varepsilon) = 0$.

3. Estimando o Modelo SAR (Wy)

b) Substituir Wy por $\widehat{Wy}$ na Eq. 41 e estimar por MQO normalmente:

$$y = \rho\widehat{Wy} + X\beta + \varepsilon \quad (50)$$

Questão: Se Wx for uma variável explicativa do modelo, posso usá-la como instrumento de Wy?
Resp.: A inclusão é permitida apenas se houver um outro instrumento além de Wx (EX.: W^2x). Caso contrário, haveria perfeita colinearidade entre Wx e $\widehat{Wy}$.

O EVI possui as seguintes vantagens:

a) Prescinde da hipótese de normalidade dos erros (necessária ao EMV);
b) Computacionalmente leve [não contém matrizes Jacobianas, $(I - \rho W)$];
c) Procedimento do EVI está implementado em diversos *softwares*;
d) Respeitando as hipóteses das Eq. 43 e 44, é um estimador consistente.

Dentre as desvantagens do EVI, destaca-se:

a) Ineficiente em pequenas amostras (eficiência assintótica);
b) Muito sensível à escolha do(s) instrumento(s).
c) A estimação via EVI pode gerar $|\hat{\rho}| > 1$. Neste caso, deve-se rever a especificação do modelo e/ou a escolha da matriz W.

4. Estimando o Modelo SEM ($W\xi$)

4.1 Máxima Verossimilhança (MV)

- Considere o seguinte modelo SEM, com $\varepsilon \sim normal(0, \sigma^2 I_n)$:

$$y = X\beta + \xi \qquad \text{onde} \qquad \xi = \lambda W\xi + \varepsilon \tag{51}$$

- Como $\varepsilon \sim normal$ podemos estabelecer a seguinte função de Log-Verossimilhança:

$$ln[V] = \left[-\frac{n}{2} ln(2\pi)\right] + \left[-\frac{n}{2}\ln(\sigma^2)\ \right] + \left[ln|I_n - \lambda W| - \frac{[(y - X\beta)'(I - \lambda W)]'[(I - \lambda W)(y - X\beta)]}{2\sigma^2}\right] \tag{52}$$

- A origem do último termo da Eq. 52 está na expressão $-\frac{\varepsilon'\varepsilon}{2\sigma^2}$ da Eq. 21 ou 29:

A forma reduzida da Eq. 51 é: $y = X\beta + (I - \lambda W)^{-1}\varepsilon$ (53)

Multiplicando ambos os lados da Eq. 52 por $(I - \lambda W)$ e aplicando distributiva:

$$(I - \lambda W)y = (I - \lambda W)X\beta + \varepsilon \quad \rightarrow \quad y - \lambda Wy = X\beta - \lambda W X\beta + \varepsilon \tag{54}$$

4. Estimando o Modelo SEM ($W\xi$)

Colocando a Eq. 54 em função de ε e colocando y e $X\beta$ em evidência:

$$\varepsilon = y - \lambda Wy - X\beta + \lambda WX\beta \quad \rightarrow \quad \varepsilon = (y - X\beta)'(I - \lambda W) \tag{55}$$

- Fazendo $\varepsilon'\varepsilon$ na Eq. 55 teremos o numerador do último termo da Eq. (52).
- O $ln[V]$ do modelo SEM, Eq. (52), também apresenta o Jacobiano $(I - \lambda W)$. Logo, o EMQO não é apropriado pois consideraria apenas parte da Eq. (52).
- Assim como no EMV para o SAR, Ord (1975) propõe que se trabalhe com os autovalores de W, isto é ω, a facilitar a estimação computacional.
- A obtenção do $\hat{\beta}_{MV}$ e da $\hat{\sigma}^2_{MV}$ para o SEM requer a derivada parcial da Eq. (52):

$$\frac{\partial \ln(V)}{\partial \beta} = 0 \rightarrow \hat{\beta}_{MV} = [(X - \lambda WX)'(X - \lambda WX)]^{-1}(X - \lambda WX)'(y - \lambda Wy) \tag{56}$$

$$\frac{\partial \log(V)}{\partial(\sigma^2)} = 0 \rightarrow \hat{\sigma}^2_{MV} = \frac{\varepsilon'\varepsilon}{n} \rightarrow \hat{\sigma}^2_{MV} = \frac{(\xi - \lambda W\xi)'(\xi - \lambda W\xi)}{n} \tag{57}$$

Nota: Eq. 56 baseada na 54 e Eq. 57 baseada na 27.

4. Estimando o Modelo SEM ($W\xi$)

- Caso $\varepsilon \neq normal$, na Eq. 51, é possível estimar a expressão consistentemente e com eficiência assintótica por QMV (Keleijian e Prucha, 1999). Contudo o EQMV, pode ser tão (ou mais!) pesado, computacionalmente, que o EMV.

4.2 Método Generalizado dos Momentos de Keleijian e Prucha (MGM)

- Devido à dificuldade computacional, o EMV e o EQMV podem não ser exequíveis em amostras muito grandes (computador não suporta estimação).

- Pensado nisso, Keleijian e Prucha (1999) desenvolveram um estimador via MGM espacial que, além de mais leve computacionalmente, não requer que os erros apresentem distribuição normal.

- Na prática, o MGM espacial se reduz a um Estimador MQGE*, onde a variância do erro ($\hat{\sigma}^2$) e o efeito causado pelo erro dos vizinhos ($\hat{\lambda}$) são estimados e reutilizados a fim de gerar estimativas consistentes e eficientes.

* Mínimos Quadrados Generalizados Exequíveis – MQGE.

4. Estimando o Modelo SEM ($W\xi$)

- Baseados na Eq. 51, os autores derivam 3 condições de momento de ε e Wε:

$$a)\ \ E\left(\frac{\varepsilon\prime\varepsilon}{n}\right)=\sigma^2\,; \qquad \text{b) } E\left(\frac{\varepsilon' W' W\varepsilon}{n}\right)=\frac{\sigma^2 tr(W'W)}{n}\,; \qquad \text{c) } E\left(\frac{\varepsilon' W\varepsilon}{n}\right)=0 \qquad (58)$$

Onde: $tr = traço$ (soma dos n elementos da diagonal principal).

- A operacionalização do EMGM envolve estimar a Eq. 51 por MQO para obter $\hat{\xi}$ e, assim, W$\hat{\xi}$ e WW$\hat{\xi}$. Logo, com base nas Eq. (58), é possível montar o seguinte sistema de Equações para obter $\hat{\sigma}^2$ e $\hat{\lambda}$:

$$\begin{bmatrix} (2/n)\hat{\xi}'W\hat{\xi} & (-1/n)W\hat{\xi}'W\hat{\xi} & 1 \\ (2/n)WW\hat{\xi}'W\hat{\xi} & (-1/n)WW\hat{\xi}'WW\hat{\xi} & (1/n)tr(W'W) \\ (1/n)(\hat{\xi}'WW\hat{\xi}+W\hat{\xi}'W\hat{\xi}) & (-1/n)W\hat{\xi}'WW\hat{\xi} & 0 \end{bmatrix}\begin{bmatrix}\lambda \\ \lambda^2 \\ \sigma^2\end{bmatrix}=\begin{bmatrix}(1/n)\hat{\xi}'\hat{\xi} \\ (1/n)W\hat{\xi}'W\hat{\xi} \\ (1/n)\hat{\xi}'W\hat{\xi}\end{bmatrix} \qquad (59)$$

- Fazendo-se a transformação das variáveis da Eq. 51, conforme proposto por Cochrane-Orcutt, é possível filtrar (eliminar) o efeito espacial causado por λ:

4. Estimando o Modelo SEM ($W\xi$)

$$y^* = y - \hat{\lambda}Wy \quad \text{e} \quad X^* = X - \hat{\lambda}WX \tag{60}$$

Nota: observando a Eq. 54 é possível notar que o procedimento (60) elimina o efeito de λ sobre ε.

- O $\hat{\beta}_{MQGE}$ (ou MGM) é obtido ao regredir y^* contra X^* via MQO. Ou seja:

$$\hat{\beta}_{MQGE} = (X^{*\prime}X^*)^{-1}\, X^{*\prime}y^* \tag{61}$$

- Resumo do procedimento MGM (ou MQGE) Espacial:

a) Estimar Eq. 51 via MQO para obter $\hat{\xi} \rightarrow \hat{\xi} = y - X\hat{\beta}$;
b) Usar $\hat{\xi}$, W$\hat{\xi}$ e WW$\hat{\xi}$ no sistema (59) para obter $\hat{\sigma}^2$ e $\hat{\lambda}$;
c) Filtrar efeito de λ via transformação de Cochrane-Orcutt (Eq. 60);
d) Estimar, via MQO, o modelo com variáveis filtradas (Eq. 61).

- As estimativas dos erros padrões via MGM são robustas contra a dependência espacial ($W\xi$). Em pequenas amostras, o EMGM é tão eficiente quanto o EMV.

Nota: o procedimento do MGM trata $\hat{\lambda}$ como um parâmetro de distúrbio (*nuisance parameter*). Logo, a significância estatística associada ao $\hat{\lambda}$ é obtida de forma indireta.

5. Estimando o Modelo SMA ($W\varepsilon$)

5.1 Máxima Verossimilhança (MV)

- Considere o seguinte modelo SMA, com $\varepsilon \sim normal(0, \sigma^2 I_n)$:

$$y = X\beta + \xi \qquad \text{onde} \qquad \xi = \gamma W\varepsilon + \varepsilon \tag{62}$$

A forma reduzida da Eq. 62 é: $y = X\beta + (I - \gamma W)\varepsilon$ (63)

Colocando em função de ε, temos:

$$\varepsilon = (y - X\beta)'(I - \gamma W)^{-1} \tag{64}$$

Lembrando que o último termo da Eq. (29) é $-\frac{\varepsilon'\varepsilon}{2\sigma^2}$, podemos estabelecer a seguinte função de Log-Verossimilhança para o Modelo SMA:

$$ln[V] = \left[-\frac{n}{2} ln(2\pi)\right] + \left[-\frac{n}{2}\ln(\sigma^2)\ \right] + \left[ln|I_n - \gamma W| - \frac{[(y - X\beta)'(I - \gamma W)^{-1}]'[(I - \gamma W)^{-1}(y - X\beta)]}{2\sigma^2}\right] \tag{65}$$

5. Estimando o Modelo SMA ($W\varepsilon$)

- O $ln[V]$ do modelo SMA, Eq. (65), também apresenta o Jacobiano $(I - \gamma W)$. Logo, o EMQO não é apropriado pois consideraria apenas parte da Eq. (65).

- Assim como os demais EMV, a obtenção do $\hat{\beta}_{MV}$ e da $\hat{\sigma}^2_{MV}$ para o SMA requer a derivada parcial da Eq. (65) em relação ao β e à σ^2, respectivamente.

- Caso $\varepsilon \neq normal$, na Eq. 65, é possível estimar a expressão consistentemente e com eficiência assintótica por QMV (Keleijian e Prucha, 1999). Contudo o EQMV, pode ser tão (ou mais!) pesado, computacionalmente, que o EMV.

5.2 Método Generalizado dos Momentos de Fingleton

- Inspirado no MGM espacial de Keleijian e Prucha (1999), para o modelo SEM, esta versão para o SMA segue os mesmos passos de seu antecessor.

5. Estimando o Modelo SMA ($W\varepsilon$)

- Passos do MGM de Fingleton para o Modelo SMA:

a) Definir os momentos associados ao SMA, análogo ao sistema (59);

b) Estimar Eq. 62 via MQO para obter $\hat{\xi} \rightarrow \hat{\xi} = y - X\hat{\beta}$;

c) Usar $\hat{\xi}$, W$\hat{\xi}$ e WW$\hat{\xi}$ no sistema gerado pelo passo (a) para obter $\hat{\sigma}^2$ e $\hat{\gamma}$;

d) Filtrar efeito de γ via transformação de Cochrane-Orcutt (análogo à Eq. 60);

e) Estimar, via MQO, o modelo com as variáveis filtradas (análogo à Eq. 61).

Nota: este procedimento também trata $\hat{\gamma}$ como um parâmetro de distúrbio (*nuisance parameter*). Logo, a significância estatística de $\hat{\gamma}$ é obtida de forma indireta.

6. Estimando o Modelo SAC (Wy e Wξ)

6.1 Máxima Verossimilhança (MV)

- Considere o seguinte modelo SAC, com $\varepsilon \sim normal(0, \sigma^2 I_n)$:

$$y = \rho W_1 y + X\beta + \xi \qquad \text{onde} \qquad \xi = \lambda W_2 \xi + \varepsilon \tag{66}$$

Colocando Eq. (66) em função de ξ e ε e assumindo que $A = (I - \rho W_1)$ e $B = (I - \lambda W_2)$, temos:

$$\xi = y - \rho W_1 y - X\beta \quad \rightarrow \quad \xi = (I - \rho W_1)y - X\beta \quad \rightarrow \quad \xi = Ay - X\beta \tag{67}$$

$$\varepsilon = \xi - \lambda W_2 \xi \quad \rightarrow \quad \varepsilon = (I - \lambda W_2)\xi \quad \rightarrow \quad \varepsilon = B\xi \tag{68}$$

Substituindo Eq. (67) em (68), temos:

$$\varepsilon = (Ay - X\beta)' B \tag{69}$$

- Como o último termo da Eq. (29) é $-\frac{\varepsilon'\varepsilon}{2\sigma^2}$, podemos estabelecer a seguinte função de Log-Verossimilhança para o Modelo SAC:

$$ln[V] = \left[-\frac{n}{2} ln(2\pi)\right] + \left[-\frac{n}{2} \ln(\sigma^2)\ \right] + \left[ln|A| + Ln|B| - \frac{[(Ay - X\beta)'B]'[B'(Ay - X\beta)]}{2\sigma^2}\right] \tag{70}$$

6. Estimando o Modelo SAC (Wy e $W\xi$)

- Note que a Eq. (70) possui 2 Jacobinos, $|A|$e $|B|$. Como a estimação via MQO ignora ambos, seus resultados seriam enviesados e inconsistentes.

Nota: Os 2 Jacobianos da Eq. (70) tornam o EMV do SAC mais complexo que os anteriores (e pesado computacionalmente!).

- Assim como os demais EMV, a obtenção do $\hat{\beta}_{MV}$ e da $\hat{\sigma}^2_{MV}$ para o SAC requer a derivada parcial da Eq. (70) em relação ao β e à σ^2, respectivamente.

- Segundo Anselin (1988) é necessário impor algumas restrições aos parâmetros do SAC para que o EMV seja assintoticamente consistente:

$$|I - \rho W_1| > 0 \quad \text{e} \quad |I - \lambda W_2| > 0 \tag{71}$$

- Se W_1 e W_2 estiverem normalizadas nas linhas, a restrição (71) se torna:

$$|\rho| < 1 \quad \text{e} \quad |\lambda| < 1 \tag{72}$$

- Logo, ρ e λ não devem ter caráter explosivo.

6. Estimando o Modelo SAC (Wy e $W\xi$)

6.2 Método MQ2E Espacial Generalizado de Keleijian e Prucha

- Dada a complexidade do EMV ou EQMV para o SAC, Keleijian e Prucha (1998) propuseram um estimador (MQ2EE) que, além de muito mais leve computacionalmente, não requer erros com distribuição normal.

- A Eq. (66) pode ser reescrita como [$\varepsilon \sim qualquer(0, \sigma^2 I_n)$]:

$y = \rho W_1 y + X\beta + \xi \quad \rightarrow \quad y = Z\delta + \xi$ (73a)

Onde: $\xi = \lambda W_2 \xi + \varepsilon$ (73b)

Sendo: Z uma matriz contendo as variáveis X e $W_1 y$, ou seja: $Z = [X, W_1 y]$. δ é um vetor contendo os coeficientes das k variáveis de Z, isto é: $\delta = [\rho, \beta]$.

- O procedimento do MQ2EE inclui 4 passos:

a) Estimar SAR - Eq. (73a), por MQ2E usando $W_1 X$ e/ou $W_1^2 X$ de instrumento(s).
b) Usar os resíduos do passo (a) no sistema (59) para obter $\hat{\lambda}$.
c) Usar $\hat{\lambda}$ para filtrar y e Z via Cochrane-Orcutt: $y^* = y - \hat{\lambda} W_2 y$; $Z^* = Z - \hat{\lambda} W_2 Z$
d) Usar o EMQO na Eq. 73a (com y^* e Z^*) para obter $\hat{\delta}$, isto é: $\hat{\delta} = (Z^{*\prime} Z^*)^{-1} Z^* y^*$

6. Estimando o Modelo SAC (Wy e Wξ)

- Os instrumento sugeridos para o MQ2EE por Kelejian e Prucha (1998) são:

$$X, W_1X, W_1^2X \dots W_2X, W_2^2X \dots W_1W_2X, W_1^2W_2X, W_1W_2^2X \dots \quad (74)$$

- Já Lee (2003) sugere o seguinte instrumento para o MQ2EE:

$$(I - \hat{\lambda}W_2)^{-1}\left[W_1(I - \hat{\rho}W_1)^{-1}X\hat{\beta}\right] \quad (75)$$

- O vetor $\hat{\delta}$, estimado via MQ2EE é consistente. Contudo, a significância associada ao $\hat{\gamma}$ precisa ser obtida de forma indireta (*nuisance parameter*).

- Dukker *et al* (2010) propuseram uma forma de estimar o SAC com mais de 1 variável endógena (além de W_1y). Os autores derivaram as condições de momentos (rever Eq. 58) para um SAC contendo W_1y, X e E, do lado direito. Onde E é uma matriz de variáveis endógenas.

- O 1º passo consiste em estimar o SAC por MQ2E, usando instrumentos para W_1y e E, e usar os resíduos para obter $\hat{\lambda}$ no sistema oriundo das condições de momentos (semelhante ao sistema 59). Feito isto, basta seguir os passos (c) e (d) do MQ2EE de Kelejian e Prucha (1998).

7. Estimando o Modelo SDM (Wy e WX)

- Assumindo que X é exógeno, WX também será. Logo, o modelo SLX, quando $y = f(X, WX)$, poderia ser estimado por MQO.

- Sendo assim, a estimação do SDM (Wy e WX) sofre apenas dos problemas do SAR (Wy) e pode ser estimado via: a) EMV (erro normal); b) EQMV (erro não normal); c) EVI (mais leve e sem pressuposto de normalidade dos erros).

Nota 1: como WX já faz parte do SDM, ela não deve ser o único instrumento para Wy no EMQ2E (como fizemos no SAR). Neste caso, inclua outro instrumento (Ex.: W^2X).

Nota 2: o SDM estimado por MQ2E tende a apresentar multicolinearidade entre X e WX (e W^2X , caso esta última seja usada como instrumento).

<u>Solução</u>: Eliminar variáveis muito correlacionadas da matriz X.

8. Estimando Outros Modelos Espaciais

- Todos os demais modelos espaciais podem ser estimados com algum método mencionado até aqui. Por exemplo:

- O SLX (WX) pode ser estimado via MV ou MQO. Porém, se $\mathrm{Corr}(\mathrm{X}, \mathrm{WX}) \neq 0$, o SLX será enviesado devido à omissão de Wy. Neste caso, o SDM (Wy e WX) é o adequado (ver Cap. 5: Eq. 53 + Notas 1 e 2 do Slide 25 + Slides 30 e 31).

- O SDEM (WX e $\mathrm{W}\xi$) pode ser estimado via MV ou MGM de Keleijian e Prucha (1998) – ver seção 4.2.

Nota: a multicolinearidade é comum onde há X e WX. Isto é: SLX, SDM, SDEM.

- Para o SARMA (Wy e $\mathrm{W}\varepsilon$), usa-se a seguinte adaptação do MGM de Keleijian e Prucha (1998), proposta por Fingleton (2008) – 4 passos:

a) Estima-se o SAR por EVI (ou MQ2E) para obter os resíduos - $\hat{\xi}$;
b) Usa-se $\hat{\xi}$ no sistema gerado via condição de momentos ($\cong$ Eq. 59) para obter $\hat{\gamma}$;
c) Usa-se $\hat{\gamma}$ para filtrar y, Wy e X via Cochrane-Orcutt e obter y^*, Wy^* e X^*;
d) Regride-se y^*contra Wy^* e X^* por MQO.

8. Estimando Outros Modelos Espaciais

- A **técnica "*Coding*"** é uma alternativa que permite obter $\hat{\beta}$ não enviesados via EMV tradicional (Eq.26) ou EMQO (Besag, 1972, 1974; Besag e Moran, 1975).
- O método parte do princípio de que observações não vizinhas são independentes. Logo, como as matrizes W definem os vizinhos, é possível obter sub-amostras aleatórias contendo apenas observações independentes.
- Neste caso, não há necessidade de incluir nenhum termo de dependência espacial na regressão (amostra conterá apenas observações aleatórias e independentes) e é possível obter $\hat{\beta}$ facilmente via EMQO ou EMV.
- Desvantagens do método "*Coding*":

a) Como sub-amostras são consideradas, deve-se trabalhar com grandes amostras. Caso contrário, haverá perda de eficiência.
b) Resultados extremamente sensíveis à escolha da matriz W.
c) Impossibilita obtenção e análise dos parâmetros que acompanham Wy, WX, $W\xi$ e $W\varepsilon$.

9. Medidas de Qualidade do Ajuste

- Na presença de erros não esféricos (heterocedásticos e com ACS), a análise do R^2 e do R^2 ajustado não é apropriada.

- No EMV, deve-se observar o valor da função de verossimilhança (LIK) – quanto maior melhor.

- Complementarmente, usa-se os critérios de informação de Akaike (AIC) e Schwartz (SC) – quanto menor melhor.

- AIC e SC usam o LIK como parâmetro e punem o excesso de variáveis incluídas no modelo – k (Nota: o SC tende a penalizar mais que o AIC). Formalmente:

$$AIC = -2LIK + 2K \quad \text{e} \quad SC = -2LIK + kln(n) \qquad (76)$$

Nota: alguns *softwares* (Ex: *SpaceStat* e *OpenGeoda*) desconsideram o λ no cálculo do AIC e SC (ambos ficam com k-1 variáveis). Isto tende a melhorar os critérios dos modelos SEM em relação ao SAR.
Solução: recalcular com base nas Eq. 76.

- No EVI, usa-se o $Pseudo\ R^2$ como critério de qualidade. Formalmente:

$$Pseudo\ R^2 = Var(\hat{y})/Var(y) \qquad (77)$$

Quadro 1. Resumo dos Métodos de Estimação dos Modelos Espaciais

Modelo	Normalidade	Método
SAR	Sim	MV
	Não	VI ou QMV
SEM	Sim	MV
	Não	MGM ou QMV
SAC	Sim	MV
	Não	MQ2EE ou QMV
SDM	Sim	MV
	Não	VI ou QMV
SLX	Sim	MQO ou MV
	Não	QMV
SDEM	Sim	MV
	Não	MGM ou QMV
SLXMA	Sim	MV
	Não	MGM ou QMV
SMA	Sim	MV
	Não	MGM ou QMV

Referência

1. ALMEIDA, E. *Econometria Espacial Aplicada.* 1ª ed. Alínea, 2012.