Econometria e Séries Temporais
- Aula 13 -

Prof. Mestre. Omar Barroso

Instituto Brasileiro de Educação, Pesquisa e Desenvolvimento

GMM (Métodos Generalizado de Momentos)

- A teoria por trás do Método Generalizado dos Momentos (GMM) está enraizada na ideia de usar condições de momento – relações derivadas de modelos econômicos – para estimar parâmetros de interesse. O GMM fornece uma estrutura flexível e poderosa, particularmente na presença de endogeneidade, heterocedasticidade e autocorrelação, onde métodos tradicionais como OLS ou MLE (Verosimilhança máxima) podem falhar.

Condições de Momentos

- Na econometria, um modelo especifica certas expectativas ou momentos que se mantêm na população. Esses momentos relacionam os parâmetros do modelo aos dados. Matematicamente, as condições de momento podem ser escritas como:

$$\mathbb{E}[g(X_t, \theta)] = 0$$

- No qual,
- X_t : É um vetor de uma data 't'.
- θ : É o vetor dos parâmetros estimados.
- $g(X_t, \theta)$: É um vetor de condições de momentos que dependem de uma data e parâmetros. A expectativa (média) dessas funções devem ser zero na população.

Ortogonalidade no GMM

- No contexto do GMM, ortogonalidade significa que **não deve haver correlação (ou relacionamento)** entre o **termo de erro** do modelo e **os instrumentos** (variáveis observáveis usadas para estimar parâmetros). Matematicamente, isso é expresso por meio de condições de momento. A condição de momento garante que:

$$\mathbb{E}[g(X_t, \theta)] = \mathbb{E}[Z_t(y_t - X_t'\theta)] = 0$$

- Z_t:é um vetor de instrumentos (variáveis exógenas ou instrumentos válidos para variáveis endógenas),
- y_t: A variável dependente,
- X_t: é o vetor de variáveis explicativas (que podem incluir as endógenas),
- θ: é o vetor de parâmetros a serem estimados,
- $y_t - x'_t\theta$: representa o termo de erro (ou residual) do modelo.

Ortogonalidade no GMM

- A condição implica que os instrumentos Zt devem ser não correlacionados com os resíduos ou erros do modelo, ou seja, $E[z_t u_t] = 0$. Isso é significado por "ortogonalidade": os instrumentos Zt são ortogonais (não correlacionados) aos resíduos, garantindo que as condições de momento sejam mantidas.
- Essa configuração é essencial para garantir que o GMM produza estimativas de parâmetros consistentes, mesmo em endogeneidade. **Os instrumentos** fornecem informações externas e exógenas que ajudam a isolar o verdadeiro efeito das variáveis explicativas na variável dependente.

Amostra de Momentos e Momentos da População

- A ideia básica do GMM é combinar momentos amostrais com seus momentos populacionais correspondentes implícitos pelo modelo. Se os momentos populacionais forem zero (como esperado), então os momentos amostrais também devem ser próximos de zero.
- Sendo $\hat{g}T(\theta)$ representando a contrapartida amostral da condição do momento populacional, que é:

$$\hat{g}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(X_t, \theta)$$

superidentificados vs. exatamente identificados

- **Modelo exatamente identificado:** O número de condições de momento é igual ao número de parâmetros. Neste caso, o GMM simplifica para resolver um sistema de equações onde os momentos da amostra são iguais a zero.
- **Modelo superidentificado:** O número de condições de momento excede o número de parâmetros. Aqui, o GMM minimiza uma soma ponderada dos desvios quadrados dos momentos da amostra de zero.

A Função Objeto do GMM

- Em casos superidentificados, o GMM minimiza a seguinte função objetivo quadrática:

$$Q_T(\theta) = \hat{g}_T(\theta)' W_T \hat{g}_T(\theta)$$

- $\hat{g}T(\theta)$: Vetor de momentos amostrais.
- wT: A matriz de pesos, que pode ser escolhido para melhorar a eficiência da estimativa.
- O objetivo é encontrar θ que minimize esta função, que representa o quão distantes os momentos da amostra estão de suas expectativas teóricas.

Matriz de ponderação ótima

- A escolha ótima para a matriz de ponderação wT é a inversa da matriz de covariância das condições de momento. Isso torna o estimador eficiente, alcançando a menor variância possível entre todos os estimadores consistentes.
- Na prática, o GMM geralmente começa com uma matriz de ponderação inicial (geralmente a matriz identidade), estima os parâmetros e, em seguida, calcula a matriz de ponderação ótima em uma segunda etapa (GMM de duas etapas).

Consistência e Normalidade Assintótica

- As estimativas do GMM são consistentes e assintoticamente normais sob condições gerais, o que significa que, à medida que o tamanho da amostra cresce:
- O estimador do **GMM converge** em probabilidade para o valor do parâmetro verdadeiro ($\theta 0$).
- A distribuição do estimador do GMM se aproxima de uma **distribuição normal**, permitindo inferência (intervalos de confiança e testes de hipóteses).
- A consistência requer que as condições de momento sejam mantidas na população e que os instrumentos usados (no caso de endogeneidade) sejam válidos.

Instrumentos e Endogeneidade

- Em muitos modelos econométricos, alguns dos regressores são **endógenos**, o que significa que **são correlacionados com o termo de erro**, levando a estimativas tendenciosas e inconsistentes no MQO. O GMM aborda isso usando **variáveis instrumentais (IVs)** que são **correlacionadas com os regressores endógenos**, mas não correlacionadas com o termo de erro.
- Por exemplo, em um modelo: $y_t = X_t' \beta + u_t$
- Se os erros são correlacionados com o termo de erro, ou seja, presença de endógeneidade. Então temos que encontrar uma IV para correlacionar com Xt.

Variáveis Instrumentais (muito simplificado)

- Utilizamos uma variável instrumental no caso se houver algum tipo de correlação entre uma variável independente e o termo de erro, ou seja, $E[\hat{x}\widehat{\epsilon]} \neq 0$.
- Em outras palavras, não temeos presença de ortogonalidade.
- Desta maneira, temos que encontrar uma variável que se não esteja correlacionada com o termo de erro.
- $\hat{x} = w(w'w)^{-1}$

Aplicações em Séries Temporais

- Em econometria de séries temporais, o GMM é particularmente útil para estimar modelos dinâmicos (com variáveis defasadas) e modelos com variáveis endógenas. As condições de momento são frequentemente derivadas de modelos teóricos ou baseadas em valores defasados das variáveis. O GMM pode lidar com problemas como:
- **Correlação serial:** quando os erros são correlacionados ao longo do tempo, o GMM pode acomodar a autocorrelação nos resíduos.
- **Endogeneidade:** o GMM usa valores defasados das variáveis endógenas ou outros instrumentos para obter estimativas consistentes.

Vamos praticar...

Exemplo com o desemprego “Genosha”

- Agora, vamos configurar um modelo AR(1), onde a taxa de desemprego no tempo “t” depende da taxa de desemprego no tempo t-1.

$$\text{unemployment}_t = \alpha + \beta \cdot \text{unemployment}_{t-1} + \varepsilon_t$$

- O termo de erro é assumido seguindo um ruído branco.
- Desta forma, vamos utilizar o GMM para estimar os parâmetros de α e β.

Definindo o Momento e Condições

- No GMM, precisamos definir as condições de momento. Para um modelo AR(1), a condição de ortogonalidade é que o termo de erro (residual) deve ser não correlacionado com os valores defasados da taxa de desemprego:
- $E[(unemployment_t - \alpha - \beta.\boldsymbol{unemployment_{t-1}}) = 0$
- Usaremos a taxa de desemprego defasada em 1 período como instrumento.

Entendendo nosso modelo

- Modelo AR(1)
- Um modelo AR(1) (AutoRegressivo de ordem 1) assume que o valor atual de uma série temporal y t depende de seu valor anterior yt−1, mais algum ruído aleatório ϵt. Matematicamente, é expresso como:

$$y_t = \alpha + \beta y_{t-1} + \epsilon_t$$

- α: Sendo o intercepto.
- β: O coeficiente do valor defasado yt-1.
- ϵ: O termo de erro ou ruído, seguindo uma distribuição normal.

GMM em nosso modelo

- No contexto do modelo AR(1), a condição de momento vem da suposição de que os resíduos (a diferença entre o valor observado yt e o valor previsto α+βy t−1) devem ter uma média zero quando multiplicados pelos valores defasados de yt. Isso leva à seguinte condição de momento:

$$E[(y_t - \alpha - \beta y_{t-1})y_{t-1}] = 0$$

- Esta condição de momento afirma que os **resíduos**, quando ponderados pelos valores defasados, devem ser **não correlacionados com os instrumentos** (neste caso, os valores defasados yt−1).
- No GMM, estimamos os parâmetros (α e β) encontrando os valores que tornam essas condições de momento o mais próximo possível de zero.

Análise dos Resultados

- Coeficientes:
- O valor estimado para α (Theta[1]) é 0,1619 (intercepto).
- O valor estimado para β (Theta[2]) é 0,7120.
- Essas estimativas sugerem que o valor atual de yt depende de cerca de 71,2% do valor anterior yt−1.
- Ou seja, 71,2% do valor do desemprego atual de Genosha depende do valor anterior.

```
Call:
gmm(g = ar1_moments, x = gmm_data, t0 = initial_params)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  1.43371 )

Coefficients:
          Estimate   Std. Error   t value   Pr(>|t|)
Theta[1]  0.16191        Inf      0.00000   1.00000
Theta[2]  0.71203        Inf      0.00000   1.00000

J-Test: degrees of freedom is -1
                J-test                P-value
Test E(g)=0:    2.43689758263322e-06  *******

Initial values of the coefficients
 Theta[1]  Theta[2]
0.1619110 0.7120279

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  39
Gradian eval. =  NA
>
```

Análise dos Resultados

- Os erros padrão são exibidos como Inf (infinito), indicando que a estimativa pode ter enfrentado problemas, possivelmente devido à singularidade na matriz de covariância. Esse problema pode surgir da **multicolinearidade** ou do processo de geração de dados não corresponder totalmente às suposições do modelo.

```
Call:
gmm(g = ar1_moments, x = gmm_data, t0 = initial_params)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  1.43371 )

Coefficients:
          Estimate   Std. Error   t value   Pr(>|t|)
Theta[1]  0.16191        Inf      0.00000   1.00000
Theta[2]  0.71203        Inf      0.00000   1.00000

J-Test: degrees of freedom is -1
                J-test                P-value
Test E(g)=0:    2.43689758263322e-06  *******

Initial values of the coefficients
 Theta[1]  Theta[2]
0.1619110 0.7120279

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  39
Gradian eval. =  NA
>
```

Análise dos Resultados

- O teste J avalia se as condições de momento são satisfeitas. Neste caso, os graus de liberdade são negativos (-1), o que é um sinal de que o modelo pode estar superidentificado ou as condições de momento não são válidas.
- A estatística J é extremamente pequena, mas o valor P não é fornecido, marcado como "*******". Este é outro sinal de problemas potenciais com o ajuste do modelo ou a estrutura de dados.

```
Call:
gmm(g = ar1_moments, x = gmm_data, t0 = initial_params)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  1.43371 )

Coefficients:
          Estimate  Std. Error  t value  Pr(>|t|)
Theta[1]  0.16191       Inf     0.00000  1.00000
Theta[2]  0.71203       Inf     0.00000  1.00000

J-Test: degrees of freedom is -1
                J-test                P-value
Test E(g)=0:    2.43689758263322e-06  *******

Initial values of the coefficients
 Theta[1]  Theta[2]
0.1619110 0.7120279

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  39
Gradian eval. =  NA
>
```

Análise dos Resultados

- O código de convergência 0 indica que o algoritmo de otimização convergiu com sucesso.
- O número de avaliações de função foi 39, o que sugere que o algoritmo de otimização iterou várias vezes para encontrar uma solução.

```
Call:
gmm(g = ar1_moments, x = gmm_data, t0 = initial_params)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  1.43371 )

Coefficients:
          Estimate  Std. Error  t value  Pr(>|t|)
Theta[1]  0.16191       Inf     0.00000  1.00000
Theta[2]  0.71203       Inf     0.00000  1.00000

J-Test: degrees of freedom is -1
                J-test               P-value
Test E(g)=0:    2.43689758263322e-06  *******

Initial values of the coefficients
 Theta[1]  Theta[2]
0.1619110 0.7120279

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  39
Gradian eval. =  NA
>
```

Análise dos Resultados

- O código de convergência 0 indica que o algoritmo de otimização convergiu com sucesso.
- O número de avaliações de função foi 39, o que sugere que o algoritmo de otimização iterou várias vezes para encontrar uma solução.

```
Call:
gmm(g = ar1_moments, x = gmm_data, t0 = initial_params)


Method:  twoStep

Kernel:  Quadratic Spectral(with bw =  1.43371 )

Coefficients:
          Estimate  Std. Error  t value  Pr(>|t|)
Theta[1]  0.16191       Inf     0.00000  1.00000
Theta[2]  0.71203       Inf     0.00000  1.00000

J-Test: degrees of freedom is -1
                J-test                P-value
Test E(g)=0:    2.43689758263322e-06  *******

Initial values of the coefficients
 Theta[1]  Theta[2]
0.1619110 0.7120279

#############
Information related to the numerical optimization
Convergence code =  0
Function eval. =  39
Gradian eval. =  NA
>
```

Recomendações Para Melhorar o Modelo

- **Não embaralhe os dados:** como os modelos AR(1) dependem da ordenação temporal das observações, evite embaralhar os dados.
- **Verifique a estacionariedade:** certifique-se de que a série temporal seja estacionária, pois a não estacionariedade pode levar a estimativas tendenciosas nos modelos AR. Você pode verificar isso usando um teste de raiz unitária, como o teste Augmented Dickey-Fuller (ADF).

Recomendações Para Melhorar o Modelo

- **Instrumentos adicionais:** Podemos adicionar mais instrumentos, como diferenças defasadas de yt, para fortalecer as condições de momento.
- **Revise a especificação do modelo:** os problemas com erros padrão infinitos e matriz de covariância singular também podem indicar especificação incorreta do modelo. Devemos verificar novamente as suposições por trás do modelo AR(1) e considere especificações alternativas, como incluir defasagens adicionais ou considerar uma estrutura de erro diferente.

Referências

- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115-143.
- Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica, 50(5), 1269-1286.
- Kleiber, Christian, and Achim Zeileis. 2008. *Applied Econometrics with R*. New York: Springer-Verlag. https://CRAN.R-project.org/package=AER.
- Stock, J. H., & Wright, J. H. (2000). GMM with weak identification. Econometrica, 68(5), 1055-1096.