一般化積率法(GMM)を用いた分析

講師: 長倉大輔

GMMを用いた分析

GMMによる推定および検定の体系は、GMMをマスターすればほとんどの分析に対応できるといっていいほどまでに発展している。

そのため、GMMについてこれらの体系をすべてカバーするのは1回の講義では不可能である。

この講義では、あまり数学的厳密性にこだわらないで、GMMの直観的な考え方の解説と、最近の話題のいくつかをとりあげるにとどめる。

GMMを用いた分析

- Generalized Method of Moment (GMM)

GMM (一般化モーメント法）は、1982年に(2013年にノーベル経済学賞を受賞した）シカゴ大学のハンセン教授によって提案された推定法である(Hansen, 1982)。

GMMはその適用範囲が広く、様々なモデルの推定に用いることができるため計量経済分析において非常によく用いられている。

GMMを用いた分析

- モーメント法

モーメント法は**モーメント条件**（**直交条件**とも呼ばれる）を用いることにより推定を行う。

モーメント条件とは、モデルより導かれる確率変数とパラメーターのある関数の期待値が 0 となるという条件のことである。

GMMを用いた分析

■ モーメント法の例

例えば $x_i \sim \mathrm{i.i.d}\ N(\mu, \sigma^2)$ において μ と σ^2 を モーメント法で推定するとしよう。

推定するパラメータベクトルを $\boldsymbol{\theta} = [\mu, \sigma^2]^{\mathrm{T}}$ とする (ここで上付き文字 T は行列の転置を表す）。

このとき、

$$g_1(x_i\,;\boldsymbol{\theta}) = x_i - \mu, \quad g_2(x_i\,;\boldsymbol{\theta}) = (x_i - \mu)^2 - \sigma^2$$

とすると

GMMを用いた分析

$$E[g_1(x_i\,;\boldsymbol{\theta})] = E(x_i - \mu) = E(x_i) - \mu = 0$$

$$E[g_2(x_i\,;\boldsymbol{\theta})] = E[(x_i - \mu)^2 - \sigma^2] = E[(x_i - \mu)^2] - \sigma^2 = 0$$

の2つの式が成り立つ。モーメント条件とは上記 2 つの式、すなわち、

$$E[g_1(x_i\,;\boldsymbol{\theta})] = 0 \quad と \quad E[g_2(x_i\,;\boldsymbol{\theta})] = 0$$

である。モーメント法ではこのモーメント条件の**標本による類似**をパラメータについて解くことによって推定量を求める。

GMMを用いた分析

具体的には、このモーメント条件における**期待値を標本平均で置き換え**

$$g_{1,n}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g_1(x_i, \boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} x_i - \mu = 0$$

$$g_{2,n}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g_2(x_i; \boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2 - \sigma^2 = 0$$

という 2 つの式を得る。これを μ と σ^2 について解くと

GMMを用いた分析

最初の式より

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

を得る。これを2つ目の式に代入して

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \hat{\mu})^2$$

を得る。これがモーメント法による、μ と σ^2 の推定量である。

GMMを用いた分析

モーメント条件は必ずしもパラメーターの数と一致するわけではない。

例えば、先ほどの例においては

$$E[(x_i - \mu)^4] = 3\sigma^4$$

も成り立つので $g_3(x_i\,;\boldsymbol{\theta}) = (x_i - \mu)^4 - 3\sigma^4$ とすると、

$$E[\,g_3(x_i\,;\boldsymbol{\theta})\,] = 0$$

は新たなモーメント条件である。

GMMを用いた分析

ここで、この期待値を標本平均で置き換えたものを

$$g_{3,n}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} g_3(x_i;\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} (x_n - \mu)^4 - 3\sigma^4$$

とし、これを 0 とおいた $g_{3,n}(\boldsymbol{\theta}) = 0$ という式も $\boldsymbol{\theta}$ の推定に使用することができる。

GMMを用いた分析

この時、$\boldsymbol{\theta} = [\mu, \sigma^2]^T$ という **2 つ**のパラメーターの推定に使用可能な式は

$$g_{1,n}(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^{n} x_i - \mu = 0$$

$$g_{2,n}(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^{n} (x_i - \mu)^2 - \sigma^2 = 0$$

$$g_{3,n}(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^{n} (x_i - \mu)^4 - 3\sigma^4 = 0$$

の **3つ**あることになる。

GMMを用いた分析

しかしながら、2 つの変数 μ と σ^2 でこの3 つの方程式を**同時に満たす**ような μ と σ^2 は**存在しない**。

このような場合、通常のモーメント法ではこれらのモーメント条件のうち、いずれか 2つの式だけを満たすように推定量を決定する。

GMMを用いた分析

例えば、最初と2番目の式を用いれば先ほどの推定量が得られ、最初と最後の式を用いれば、μ と σ^2 の推定量は

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

および

$$\hat{\sigma}^2 = \sqrt{\frac{1}{3n}\sum_{i=1}^{n}(x_i - \hat{\mu})^4}$$

となる。

GMMを用いた分析

このように モーメント条件の数がパラメーターの数より大きいときはいくつかのモーメント条件を**使用しない**ことによりパラメーターの推定を行うことはできる。

しかしながら、せっかくなので、できるならば、**すべてのモーメント条件を使用して**、推定の効率をあげたい。

そのような要望に応えたものがGMMである。

GMM は**モーメント条件の数がパラメータ数より多い場合**に対してモーメント法を一般化したものである。

GMMを用いた分析

- GMMの目的関数

先ほどの 2×1のパラメーターベクトルに対して、モーメント条件、$E[\, g_j\,(x_i\,;\,\boldsymbol{\theta})\,] = 0$, j =1, …,3 における期待値を標本平均で置き換えたものを $g_{j,n}(\boldsymbol{\theta})$ とし、それらを並べたベクトルを

$$\mathbf{g}_n(\boldsymbol{\theta}) = \begin{bmatrix} g_{1,n}(\boldsymbol{\theta}) \\ g_{2,n}(\boldsymbol{\theta}) \\ g_{3,n}(\boldsymbol{\theta}) \end{bmatrix}$$

と定義する。

GMMを用いた分析

このとき、GMMでは以下のような目的関数を定義する。

$$Q_n(\boldsymbol{\theta}, \mathbf{W}) = \mathbf{g}_n(\boldsymbol{\theta})^{\mathrm{T}}\,\mathbf{W}\,\mathbf{g}_n(\boldsymbol{\theta})$$

ここで $\mathbf{W}$ は 正値定符号行列とする。この $\mathbf{W}$ を**重み行列(Weighting matrix)** と呼ぶ。この目的関数を最小化する $\boldsymbol{\theta}$ の値が、$\boldsymbol{\theta}$ の **GMM 推定値**である。

後にみるように $\boldsymbol{\theta}$ の推定量の分散は $\mathbf{W}$ に依存し、$\mathbf{W}$ を最適に選ぶことにより、$\boldsymbol{\theta}$ の推定量の分散を可能な限り小さくすることができる。

GMMを用いた分析

- GMMの原理の直観的な説明

例えば **W** として 単位行列を選べば目的関数は

$$Q_n(\boldsymbol{\theta}, \mathbf{W}) = \mathbf{g}_n(\boldsymbol{\theta})^{\mathrm{T}} \mathbf{I}_3 \, \mathbf{g}_n(\boldsymbol{\theta})$$
$$= g_{1,n}(\boldsymbol{\theta})^2 + g_{2,n}(\boldsymbol{\theta})^2 + g_{3,n}(\boldsymbol{\theta})^3$$

となる。これを最小にするということは、$g_{j,n}(\boldsymbol{\theta})$ のそれぞれをできるかぎり 0 に近づけるということである。

W が正値定符号行列の場合も同様である。
GMMの目的関数はある意味で $\mathbf{g}_n(\boldsymbol{\theta})$ と **0** との**距離**を表すものと解釈できる。それを最小にするように選ぶという事である。

GMMを用いた分析

本来なら $g_{j,n}(\boldsymbol{\theta})$ のそれぞれをすべて同時に0にするような $\boldsymbol{\theta}$ があれば、それが一番良いが、そのような$\boldsymbol{\theta}$は一般的には必ずしも存在しない。ならばすべての $g_{j,n}(\boldsymbol{\theta})$を全体的に0に近くするような $\boldsymbol{\theta}$ を求めようということである。

GMMを用いた分析

一般的にモーメント条件は観測値ベクトル $\mathbf{w}_i$ と $h\times 1$未知パラメーターベクトル $\boldsymbol{\theta}$ の関数 $g_j(\mathbf{w}_i;\boldsymbol{\theta})$ に対して

$$E[g_j(\mathbf{w}_i;\boldsymbol{\theta})]=0,\quad j=1,..,r,\ \ i=1,...,n$$

で与えられる。この場合 r がモーメント条件の数である。先ほどの正規分布の例では $\mathbf{w}_i = x_i$ 、$\boldsymbol{\theta} = [\mu, \sigma^2]^{\mathrm{T}}$とし、
$g_1(\mathbf{w}_i;\ \boldsymbol{\theta}) = \mathbf{w}_i - \mu,\ \ g_2(\mathbf{w}_i;\ \boldsymbol{\theta}) = (\mathbf{w}_i - \mu)^2 - \sigma^2,$
$g_3(\mathbf{w}_i;\ \boldsymbol{\theta}) = (\mathbf{w}_i - \mu)^3 - 3\sigma^4$ としたものである。

GMM推定量が定義されるためには**モーメント条件の数は 未知パラメータの数($\boldsymbol{\theta}$ の次元) 以上でないといけない。**

GMMを用いた分析

モーメント条件の期待値を標本平均で置き換えたものを

$$g_{j,n}(\boldsymbol{\theta}) = n^{-1}\sum_{i=1}^{n} g_j(\mathbf{w}_i;\boldsymbol{\theta}),$$

とすると、GMM推定の目的関数は

$$Q_n(\boldsymbol{\theta}, \mathbf{W}) = \mathbf{g}_n(\boldsymbol{\theta})^{\mathrm{T}}\,\mathbf{W}\,\mathbf{g}_n(\boldsymbol{\theta})$$

と定義される。ここで $\mathbf{g}_n(\boldsymbol{\theta})$ は $r\times 1$ ベクトル:

$$\mathbf{g}_n(\boldsymbol{\theta}) = [g_{1,n}(\boldsymbol{\theta})\ \ g_{2,n}(\boldsymbol{\theta})\ \ \ldots\ \ g_{r,n}(\boldsymbol{\theta})]^{\mathrm{T}}$$

で、$\mathbf{W}$ は $r\times r$ の正値定符号行列である。これを最小化する $\boldsymbol{\theta}$ の値が **GMM 推定量**である。

GMMを用いた分析

- GMMとモーメント法による推定量

モーメント法はすべてのモーメント条件

$$\mathbf{g}_n(\boldsymbol{\theta}) = \mathbf{0}$$

を満たすようなパラメーターの値を推定値とした。これに対してGMMは目的関数

$$\mathbf{g}_n(\boldsymbol{\theta})^{\mathrm{T}}\mathbf{W}\mathbf{g}_n(\boldsymbol{\theta})$$

を最小にするパラメーターの値を推定値とした。
一般に、モーメント条件の数とパラメーターの数が等しい時(若干の追加条件のもとで)、$\mathbf{g}_n(\boldsymbol{\theta}) = \mathbf{0}$ の解が存在する。よってこの時、GMM目的関数は($\mathbf{W}$が正値定符号行列なので)$\mathbf{g}_n(\boldsymbol{\theta}) = \mathbf{0}$ で最小値 0 を取る。つまりこの時、GMM推定量とモーメント法による推定量は一致する。

GMMを用いた分析

- GMMの例(最小二乗法)

GMMは既存の様々な推定量を特別な場合として含んでいる。最小二乗法も GMM の特別な場合としてみなすことができる。

以下の標準的な回帰モデルを考える。

$$Y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_K X_{Ki} + u_i, \quad u_i \sim \text{i.i.d. } (0, \sigma^2)$$

また X_{kj} と u_i はすべての j と i で独立とする。**未知係数は $K+1$ 個あることに注意**。

GMMを用いた分析

この時、X_{ki} と u_i は独立なので

$$E(X_{ki}\, u_i) = 0,\ \ k = 1,\ldots,K \quad \text{および} \quad E(u_i) = 0$$

が成り立つ。よって $\mathbf{w}_i = [Y_i\,, X_{1i}\,,\ldots, X_{Ki}]^{\mathrm{T}}$,
$\boldsymbol{\theta} = [\beta_0, \beta_1, \ldots, \beta_K]^{\mathrm{T}}$ とし

$$g_j(\mathbf{w}_i;\boldsymbol{\theta}) = X_{ji}u_i = X_{ji}(Y_i - \beta_0 - \beta_1 X_{1i} - \ldots - \beta_K X_{Ki})$$

および $g_{K+1}(\mathbf{w}_i;\, \boldsymbol{\theta}) = Y_i - \beta_0 - \beta_1 X_{1i} - \ldots - \beta_K X_{Ki}$ とすれば

$$E[g_j(\mathbf{w}_i\,;\,\boldsymbol{\theta})] = 0\ ,\ \ j = 1,\ldots, K+1$$

という $K+1$ 個のモーメント条件を得る。

GMMを用いた分析

この期待値を標本平均で置き換えたものは

$$g_{j,n}(\mathbf{w}_i\,;\boldsymbol{\theta}) = n^{-1}\sum_{t=1}^{n} X_{ji}(Y_i - \beta_0 - \beta_1 X_{1i} - \ldots - \beta_K X_{Ki}) = 0$$

$$\Leftrightarrow$$

$$\sum_{t=1}^{n} X_{ji}(Y_i - \beta_0 - \beta_1 X_{1i} - \ldots - \beta_K X_{Ki}) = 0$$

という条件を得る。ここで $j = 1, \ldots, K+1$, $X_{K+1,i} = 1$ とする。

これは最小二乗法を得る際に導出した正規方定式と同じである。よってこの最小二乗法は前述のモーメント条件を用いたモーメント法とみることができる。よってGMMの特殊ケースとして含まれる。

GMMを用いた分析

- 重み行列Wの選択

GMM推定量の**推定効率(漸近分散)は重み行列 W に依存している**。推定効率をもっともよくする W は

$$\mathbf{W} = \mathbf{S}^{-1} = (E[\mathbf{g}(\mathbf{w}_i;\boldsymbol{\theta})\mathbf{g}(\mathbf{w}_i;\boldsymbol{\theta})^{\mathrm{T}}])^{-1}$$

であることを示すことができる。ここでg(wi; θ) は

$$\mathbf{g}(\mathbf{w}_i;\,\boldsymbol{\theta}) = [\,g_1(\mathbf{w}_i;\,\boldsymbol{\theta}),\,g_2(\mathbf{w}_i;\,\boldsymbol{\theta})\,,\,\ldots,\,g_r(\mathbf{w}_i;\,\boldsymbol{\theta})]^{\mathrm{T}}$$

である。ただし、この S は直接は観測できないのでその推定値を用いる。この S の推定にはいろいろな方法が提案されているが、よく用いられるのは Andrews (1991) が提案したような カーネルベースの推定量である。

GMMを用いた分析

- Efficient GMM 推定量

W として $\mathbf{S}^{-1}$ の一致推定量を用いたGMM 目的関数

$$Q_n(\boldsymbol{\theta}, \hat{\mathbf{S}}_n^{-1}) = \mathbf{g}_n(\boldsymbol{\theta})^{\mathrm{T}} \hat{\mathbf{S}}_n^{-1} \mathbf{g}_n(\boldsymbol{\theta})$$

を最小にする $\boldsymbol{\theta}$ は **Efficient GMM 推定量**と呼ばれ、漸近的に正規分布に従い、その漸近分散は他のどのような重み行列を用いた時よりも小さくなる。

以下では GMM に関連したいくつかの統計量および問題点を挙げておく。

GMMを用いた分析

- J 統計量

GMM目的関数を Efficient GMM 推定量 で評価したものに n (観測数) をかけたもの、すなわち

$$J_n = nQ_n(\hat{\boldsymbol{\theta}}_{egmm}, \hat{\mathbf{S}}_n^{-1})$$

は **J 統計量**と呼ばれる。ここで $\hat{\boldsymbol{\theta}}_{egmm}$ は Efficient GMM 推定量である。J 統計量は**モデルの全ての仮定が正しいとき**、自由度 $r-h$ のカイ二乗分布に従う。ここで r はモーメント条件の数、hはパラメータベクトルの次元である。

GMMを用いた分析

よってこの統計量が、このカイ二乗分布からの標本の値として大きすぎるときには、(モーメント条件などの)モデルの仮定のどれかが満たされていないと考えられるので、改めてモデルを見直す必要が出てくる。

このように J 統計量を用いて、モデルの定式化の誤りがあるかどうかを検定することを **J 検定**という

J 検定によって　帰無仮説 : 定式化の誤りがない、が棄却された場合でも、どの仮定が満たされていないかについては何も示唆してくれないので注意が必要である。

GMMを用いた分析

- モーメント(条件)選択基準

実証分析ではしばしば J 統計量によってモデルの仮定が棄却される。この理由として、使用したモーメント条件のうち正しくないものが含まれている可能性がある。

そのような場合、どのモーメント条件が正しくてどのモーメント条件が正しくないかを推定するような方法があれば有用であろう。

ここではそのような方法のうち Andrews (1999) と Hall, Inoue, Jana, and Shin (2007) による**モーメント(条件)選択基準（Moment Selection Criteria; MSC)** と呼ばれるものを紹介する。

GMMを用いた分析

- Andrews の MSC

Andrews (1999) は以下の 2つのMSCを提案した

$$\text{SIC} - \text{based:} \quad J_n - (r - h)\log(n)$$
$$\text{HQIC} - \text{based:} \quad J_n - 2.01(r - h)\log(\log(n))$$

ここで J_T は J 統計量、r は J 統計量の計算(パラメーター の推定)に使用したモーメントの数、h は推定したパラメータの数である。SIC – based は情報量基準のSIC(BIC)のGMMに対する類似の統計量、HQIC – based は情報量基準の HQIC(Hannan-Quinn 情報量基準) のGMMに対する類似の統計量である。

GMMを用いた分析

SIC (HQIC) – based を可能なモーメント条件の組み合わせに対して計算し、その値が一番小さくなるモーメント条件の組み合わせが “正しい” モーメント条件として選ばれる。

Simulation の結果、SIC –based のパフォーマンスが一番良いことが分かっている。

GMMを用いた分析

- Hall, Inoue, Jana, and Shin の MSC

Hall, Inoue, Jana, and Shin (2007) は 線形回帰モデルに対して、以下の Relevant MSCを提案した

$$\text{Relevant MSC} = \log(|\hat{\mathbf{V}}_\theta|) + \frac{(r-h)}{\sqrt{\tau}}\log(\sqrt{\tau})$$

ここで $\hat{\mathbf{V}}_\theta$ は Efficient GMM の漸近分散の推定値、τは重み行列の推定の仕方に応じて取る値が変わる値である。

RMSCもそれぞれのモーメント条件の組み合わせに対して計算し、最も小さい値のモーメント条件を選ぶ。

演習問題

問題1 p.13 で求めた σ^2 の推定量が一致推定量であることを証明しなさい。以下の大数の法則が成り立つとして証明してよい。

$$\frac{1}{n}\sum_{i=1}^{n} x_i \to_p \mu \quad , \quad \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^4 \to_p 3\sigma^4$$

問題2 γ 分布(密度関数 $f(x_i) = x_i^{\alpha-1} e^{-x/\theta} / \Gamma(\alpha)\theta^{\alpha}$ $(x_i > 0)$)に対して $E(x_i) = \alpha\theta$、$\mathrm{var}(x_i) = \alpha\theta^2$ より α と θ のモーメント推定量を求めなさい。

参考文献

Andrews, D.W.K. (1991) “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation,” ***Econometrica***, 59, 817-858.

Andrews, D.W.K. (1999) “Consistent Moment Selection Procedures for Generalized Method of Moments Estimation,” ***Econometrica***, 67(3), 543-564.

Hall, A. R., A. Inoue, K.Jana, and C. Shin (2007) “Information in Generalized Method of Moments Estimation and Entropy-based Moments Selection,” ***Journal of Econometrics***, 38, 488-512.

Hansen, L., (1982) “Large Sample Properties of Generalized Method of Moments Estimators,” ***Econometrica***, 50, 1029-1054.