

ファクター分析[†]

講師: 長倉大輔

[†] この資料は私のゼミにおいて使用するために作成した資料です。ゼミのWEBページ上で公開しており、自由に参照して頂いて構いません。ただし、内容について、一応検証してありますが、もし間違いがあった場合でもそれによって生じるいかなる損害、不利益について責任を負いかねますのでご了承ください。間違いは発見次第、継続的に直していますが、まだ存在する可能性があります。間違いを見つけた場合 nagakura@z7.keio.jp までご連絡いただくと大変嬉しく思います。

内容

1. ファクターモデルについて
2. ファクターの推定法について
3. ファクター数の決定について
4. 実証例

1. ファクターモデルについて

■ ファクター分析とは？

ファクターモデルでは、多数の経済変数を少数の共通するファクター(因子)の関数として表す。

ファクター分析ではそれらのファクターを推定し、またこの少数のファクターが多数の経済変数にどのような影響を与えているのかも推定する。

推定した共通ファクターは、インフレ率などの経済変数の予測に有用なことが、多くの実証分析で示されている。

1. ファクターモデルについて

■ ファクターモデル

観測されるデータを x_{it} , $i = 1, \dots, N$, $t = 1, \dots, T$ とする。 x_{it} のファクターモデルは

$$x_{it} = \mathbf{f}_t^T \boldsymbol{\lambda}_i + \varepsilon_{it},$$

と表される。ここで $\mathbf{f}_t = [f_{1t}, f_{2t}, \dots, f_{rt}]^T$ は $r \times 1$ のファクターベクトル、 $\boldsymbol{\lambda}_i = [\lambda_{1i}, \lambda_{2i}, \dots, \lambda_{ri}]^T$ は各ファクター f_{kt} の係数のベクトルであり**ファクター一負荷**と呼ばれる。 $\boldsymbol{\lambda}_i$ も \mathbf{f}_t も**観測されない**とする。

1. ファクターモデルについて

- \mathbf{f}_t はクロスセクション方向の単位 i に依存せず、各ファクター x_{it} に**共通である**。
- λ_i は各 i に依存して異なる、つまり各 x_{it} において共通ファクター \mathbf{f}_t の**係数は異なっている**。
- ファクターの個数 r は通常 変数の個数 N よりかなり小さい (通常 $N = 50 \sim 100$, $r = 3 \sim 5$ である)。

1. ファクターモデルについて

(ファクターモデルのベクトル表記)

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ x_{4t} \\ \vdots \\ \vdots \\ \vdots \\ x_{N-2,t} \\ x_{N-1,t} \\ x_{N,t} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{r1} \\ \lambda_{12} & \cdots & \lambda_{r2} \\ \lambda_{13} & \cdots & \lambda_{r3} \\ \lambda_{14} & \cdots & \lambda_{r4} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \lambda_{1,N-2} & \cdots & \lambda_{r,N-2} \\ \lambda_{1,N-1} & \cdots & \lambda_{r,N-1} \\ \lambda_{1,N} & \cdots & \lambda_{r,N} \end{bmatrix} \begin{bmatrix} f_{1t} \\ \vdots \\ f_{rt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_{N-2,t} \\ \varepsilon_{N-1,t} \\ \varepsilon_{N,t} \end{bmatrix}$$

r は小さい

N は大きい

2. ファクターの推定法について

以下ではこの共通ファクター f_t とファクター負荷 λ_i の推定方法、およびファクター数 r の決定方法について述べるが、その前にファクターモデルによる予測について少し述べておく。

2. ファクターの推定法について

■ ファクターモデルによる予測

ファクターモデルの推定の前に、推定したファクターをどのように予測に用いることができるかを簡単に述べておく。

ここでは Stock and Watson (2002) の方法を述べる。

2. ファクターの推定法について

まず、先ほどのファクターモデルにおいて x_{it} を $i = 1, \dots, N$ に横に並べると

$$\begin{aligned} [x_{1t}, x_{2t}, \dots, x_{Nt}] &= [\mathbf{f}_t^T \boldsymbol{\lambda}_1, \mathbf{f}_t^T \boldsymbol{\lambda}_2, \dots, \mathbf{f}_t^T \boldsymbol{\lambda}_N] + [\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}] \\ &= \mathbf{f}_t^T [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_N] + [\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}] \end{aligned}$$

である。ここで $\mathbf{x}_t = [x_{1t}, \dots, x_{Nt}]^T$, $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t}, \dots, \varepsilon_{Nt}]^T$
 $\boldsymbol{\Lambda} = [\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \dots, \boldsymbol{\lambda}_N]^T$ とすると、上記はさらに

$$\mathbf{x}_t^T = \mathbf{f}_t^T \boldsymbol{\Lambda}^T + \boldsymbol{\varepsilon}_t^T, \quad t = 1, \dots, T$$

と表せる。 $\boldsymbol{\Lambda}$ は $N \times r$ 行列であることに注意。

2. ファクターの推定法について

次に、予測する(スカラー)変数を y_t とする。
 y_t の**予測モデル**として

$$y_{t+h} = \mathbf{f}_t^T \boldsymbol{\beta} + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h}$$

というモデルを考える。ここで $\boldsymbol{\beta}$ は $r \times 1$ 係数ベクトル、 \mathbf{w}_t は観測される $p \times 1$ の変数のベクトルで \mathbf{f}_t 以外で y_{t+h} の予測に寄与する変数とし、 $\boldsymbol{\gamma}$ は \mathbf{w}_t の $p \times 1$ 係数ベクトルとする。

\mathbf{w}_t としては y_t の過去の値がよく用いられる。

2. ファクターの推定法について

よって予測モデルは

$$\begin{aligned}\mathbf{x}_t^T &= \mathbf{f}_t^T \Lambda^T + \boldsymbol{\varepsilon}_t^T, \\ y_{t+h} &= \mathbf{f}_t^T \boldsymbol{\beta} + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h}\end{aligned}$$

という2つの式からなる。

2つ目の式は \mathbf{f}_t が観測できないため、直接は推定できない。よって1つ目の式より \mathbf{f}_t を推定し、それを用いて2つ目の式の $\boldsymbol{\beta}$ と $\boldsymbol{\gamma}$ を推定する。

2. ファクターの推定法について

最後に y_{t+h} の予測値 \hat{y}_{t+h} を

$$\hat{y}_{t+h} = \hat{\mathbf{f}}_t^T \hat{\boldsymbol{\beta}} + \mathbf{w}_t^T \hat{\boldsymbol{\gamma}}$$

によって計算する。ここでハットがついたものはそれぞれハットなしのものの推定値である。

2. ファクターの推定法について

■ 予測モデルの解釈

先ほどの予測モデルでは \mathbf{x}_t の**共通ファクター**を y_t の予測に用いていた。これは \mathbf{x}_t を予測に用いた式、つまり

$$\begin{aligned} y_{t+h} &= \delta_1 x_{1t} + \delta_2 x_{2t} + \dots + \delta_N x_{Nt} + \mathbf{w}_t^T \boldsymbol{\gamma} + e_{t+h} \\ &= \mathbf{x}_t^T \boldsymbol{\delta} + \mathbf{w}_t^T \boldsymbol{\gamma} + e_{t+h} \end{aligned}$$

とどのような関係があるだろうか？

2. ファクターの推定法について

共通ファクターを用いたモデルは、全て x_{it} を用いた予測モデルの**簡易的な表現**と考えることができる。なぜならこの式に先ほどのファクターモデルを代入すると

$$\begin{aligned} y_{t+h} &= (\mathbf{f}_t^T \Lambda^T + \boldsymbol{\varepsilon}_t^T) \boldsymbol{\delta} + \mathbf{w}_t^T \boldsymbol{\gamma} + e_{t+h} \\ &= \mathbf{f}_t^T \Lambda^T \boldsymbol{\delta} + \mathbf{w}_t^T \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t^T \boldsymbol{\delta} + e_{t+h} \end{aligned}$$

であり、ここで $\boldsymbol{\beta} = \Lambda^T \boldsymbol{\delta}$ 、 $u_t = \boldsymbol{\varepsilon}_t^T \boldsymbol{\delta} + e_t$ とおけば、先ほどの予測モデルを得ることになる。

2. ファクターの推定法について

これは N が大きいときはもともとの \mathbf{x}_t が直接入った予測モデルを精度よく推定するのは困難なので、 \mathbf{x}_t から取り出した少数の共通因子によって予測モデルの推定精度を上げていると考えることができる。

多数の変数からの情報を共通ファクターという少数の情報に集約することができるのがファクターモデルの利点の一つである。

2. ファクターの推定法について

■ 識別の議論の準備

$\mathbf{x}_t, t = 1, \dots, T$ を転置し、並べて

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_T^T \end{bmatrix} = \begin{bmatrix} x_{11} & \dots & x_{N1} \\ \vdots & \dots & \vdots \\ x_{1T} & \dots & x_{NT} \end{bmatrix}$$

という $T \times N$ 行列 \mathbf{X} を定義する。

2. ファクターの推定法について

同様に、 \mathbf{f}_t を転置し並べた $T \times r$ 行列と $\boldsymbol{\varepsilon}_t$ を転置して並べた $T \times N$ 行列をそれぞれ

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}_1^T \\ \vdots \\ \mathbf{f}_T^T \end{bmatrix} = \begin{bmatrix} f_{11} & \cdots & f_{r1} \\ \vdots & \cdots & \vdots \\ f_{1T} & \cdots & f_{rT} \end{bmatrix}$$

$$\mathbf{E} = \begin{bmatrix} \boldsymbol{\varepsilon}_1^T \\ \vdots \\ \boldsymbol{\varepsilon}_T^T \end{bmatrix} = \begin{bmatrix} \varepsilon_{11} & \cdots & \varepsilon_{N1} \\ \vdots & \cdots & \vdots \\ \varepsilon_{1T} & \cdots & \varepsilon_{NT} \end{bmatrix}$$

と定義する。

2. ファクターの推定法について

これらを用いると先ほどのファクターモデルは

$$\mathbf{X} = \mathbf{F}\mathbf{\Lambda}^T + \mathbf{E}$$

$(T \times N) \quad (T \times r)(r \times N) \quad (T \times N)$

と簡潔に表現できる。

以下ではまずこの表現を用いてファクターとファクター負荷の識別の問題を考える。

2. ファクターの推定法について

■ 識別の問題

識別の問題とは観測された X の値から F と Λ を1つだけ決めることができるかという問題である。

言い換えると、1つの X の値に対しては一つの F と Λ が対応していないと識別できない。

以下に見るように、ファクターモデルでは F と Λ は**識別できない**。

2. ファクターの推定法について

\mathbf{A} を $r \times r$ の**任意の**正則行列とすると、先ほどのモデルは

$$\mathbf{X} = \mathbf{F}\mathbf{\Lambda}^T + \mathbf{E} = \mathbf{F}\mathbf{A}\mathbf{A}^{-1}\mathbf{\Lambda}^T + \mathbf{E} = \mathbf{F}^*\mathbf{\Lambda}^{*T} + \mathbf{E}$$

と表すことができる。ここで $\mathbf{F}^* = \mathbf{F}\mathbf{A}$ および $\mathbf{\Lambda}^* = \mathbf{\Lambda}\mathbf{A}^{-T}$ である。

2. ファクターの推定法について

上記の議論より、 \mathbf{X} と観測上同じになるファクターとファクター負荷の値 \mathbf{F} と $\mathbf{\Lambda}$ の組み合わせが**無数にある**ことがわかる。

よって \mathbf{F} と $\mathbf{\Lambda}$ に何らかの制約を置かないと \mathbf{X} から \mathbf{F} と $\mathbf{\Lambda}$ を識別できないことがわかる。

よく用いられる制約は

「 $\mathbf{F}^T \mathbf{F} / T = \mathbf{I}_r$ かつ $\mathbf{\Lambda}^T \mathbf{\Lambda}$ が対角行列」

というもの。ここで \mathbf{I}_r は $r \times r$ の単位行列。

2. ファクターの推定法について

これらの識別条件によって、 $F\Lambda^T$ の値に対して一意的に F と Λ が決まる (識別条件は $F\Lambda^T$ を一意的に F と Λ^T にわけるための制約)。

この時、 F と Λ の推定値は識別条件によって変わる。言い換えると真の F と Λ は推定できない

それでは F と Λ の推定値はいったい何を推定しているのだろうか？

2. ファクターの推定法について

結論を言うと、(後述する) \mathbf{F} と $\mathbf{\Lambda}$ の推定値は \mathbf{F} と $\mathbf{\Lambda}$ が張る空間を推定している。

つまり 真の \mathbf{F} と $\mathbf{\Lambda}$ の値を \mathbf{F}_{true} , $\mathbf{\Lambda}_{\text{true}}$ とすると、推定しているのはそれらに任意の正則行列 \mathbf{A} を掛けたもの

$$\mathbf{F}_{\text{true}}\mathbf{A}, \quad \mathbf{\Lambda}_{\text{true}}\mathbf{A}^{-T}$$

である(ここで \mathbf{A} は同じ行列)。

2. ファクターの推定法について

■ 識別の予測値への影響

これは前述した予測値にどのような影響を与えるであろうか？

結論をいうと、**識別条件は予測値には影響を与えない。**

これは以下のように考えるとわかる。

2. ファクターの推定法について

F の真の値は推定できないが、

$$\mathbf{F}_{\text{true}} \mathbf{A} = \begin{bmatrix} \mathbf{f}_{1,\text{true}}^T \mathbf{A} \\ \vdots \\ \mathbf{f}_{T,\text{true}}^T \mathbf{A} \end{bmatrix}$$

が推定できるのであれば、予測モデルの第2式

$$y_{t+h} = \mathbf{f}_{t,\text{true}}^T \boldsymbol{\beta} + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h}$$

は

2. ファクターの推定法について

$$\begin{aligned}y_{t+h} &= \mathbf{f}_{t,\text{true}}^T \boldsymbol{\beta} + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h} \\ &= \mathbf{f}_{t,\text{true}}^T \mathbf{A} \mathbf{A}^{-1} \boldsymbol{\beta} + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h} \\ &= \mathbf{f}_{t,\text{true}}^T \mathbf{A} \boldsymbol{\beta}^* + \mathbf{w}_t^T \boldsymbol{\gamma} + u_{t+h}\end{aligned}$$

と書き直すことができる。 $\mathbf{f}_{t,\text{true}}^T \mathbf{A}$ は正しく推定できるので、この場合推定する係数が $\boldsymbol{\beta}$ から $\boldsymbol{\beta}^*$ になるだけで、予測値には影響を与えない。

2. ファクターの推定法について

■ F と Λ の推定

現在、計量経済学の分野で、 F と Λ^T の推定に最もよく用いられる推定方法は**主成分分析** (Principal component analysis) である。

主成分分析は統計学で長く用いられてきた方法であるが、近年、計量経済学の分野で、ファクターモデルの推定に応用するための理論的な正当化がなされ、その推定量としての理論的な性質が明らかになってきている。

2. ファクターの推定法について

- 主成分分析 (手法が同じだけで解釈は全く異なる)

主成分分析ではファクター \mathbf{F} (主成分分析では主成分負荷量ベクトルに相当) を $\mathbf{X}\mathbf{X}^T$ の固有値の大きいものから r 番目 (r は事前に設定するファクターの数) までの固有ベクトルに $T^{1/2}$ を掛けたものとして推定する。この \mathbf{F} の推定値を $\hat{\mathbf{F}}$ とすると、 Λ (主成分分析では主成分に対応) の推定値は

$$\hat{\Lambda} = \mathbf{X}^T \hat{\mathbf{F}} / T$$

によって得られる。

2. ファクターの推定法について

■ ファクター分析による推定の解釈

ファクター分析による推定は**最小二乗法**として解釈できる。目的関数を

$$\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \mathbf{f}_t^T \boldsymbol{\lambda}_i)^2$$

とした場合、前述の推定量はこの目的関数を **$\mathbf{F}^T \mathbf{F} = \mathbf{I}$** という制約のもとで $\boldsymbol{\Lambda}$ と \mathbf{F} について最小化した解と同じになる。

2. ファクターの推定法について

主成分分析による F と Λ の推定量の一致性は Stock and Watson (2002)、その漸近分布の導出は Bai (2003) による。

一致性、漸近正規性は N と T の両方が ∞ に行く漸近理論のもとで証明される。

(詳しい条件は省略するが) 漸近分布は

2. ファクターの推定法について

$\hat{\mathbf{f}}_t$ については

$$\sqrt{N}(\hat{\mathbf{f}}_t - \mathbf{A}\mathbf{f}_t) \rightarrow_d N(0, \mathbf{V}_f)$$

で与えられ、 $\hat{\lambda}_i$ については

$$\sqrt{T}(\hat{\lambda}_i - \mathbf{A}^{-T}\lambda_i) \rightarrow_d N(0, \mathbf{V}_\lambda)$$

によって与えられる。ここで \mathbf{A} は両方に共通のある正則行列である。

3. ファクター数の決定について

■ ファクターの数の選択

実際の分析においてファクターの数 r がわかっていることはまずないのでデータから推定する必要がある。

よく用いられる方法に Bai and Ng (2002) による方法と Onatski (2010) の方法があるのでこれらを紹介する。

3. ファクター数の決定について

■ Bai and Ng (2002) の方法

Bai and Ng(2002) は情報量基準を用いて、 r を選択する方法を提案した。彼らが提案した情報量基準は

$$PC(r) = \min_{\substack{\lambda_1, \dots, \lambda_N \\ \mathbf{f}_1, \dots, \mathbf{f}_T}} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \mathbf{f}_t^T \boldsymbol{\lambda}_i)^2 + rg(N, T)$$

と定義される。ここで $g(N, T)$ はペナルティ関数とよばれるもので、

3. ファクター数の決定について

例えば、

$$g(N, T) = \hat{\sigma}^2 \left(\frac{N + T}{NT} \right) \ln \left(\frac{NT}{N + T} \right)$$

などが提案されている。ここで $\hat{\sigma}^2$ は ε_{it} の分散の一致推定量で例えば

$$\hat{\sigma}^2 = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \hat{\mathbf{f}}_t^T \hat{\boldsymbol{\lambda}}_i)^2$$

を用いることができる。 $r = 1, 2, 3, \dots$ に対してこのような情報量基準を計算し、最も小さい値に対応する r を選ぶ。

3. ファクター数の決定について

■ Onatski (2010) の方法

Onatski (2010) は XX^T の固有値の減退具合を利用することにより、ファクターの数を選ぶ方法を提案している。

$XX^T/(NT)$ の固有値で大きい者から順に d_1, d_2, \dots とすると、共通因子の個数 r を超えたところから (d_{r+1} から) 固有値の値が急激に減退することが知られている。

3. ファクター数の決定について

Onatski (2010) はこの性質を利用し、固有値の大きい方から順に差

$$w_k = d_k - d_{k+1},$$

を取り、 w_k がある基準値より大きくなった時の k の値をファクター数 r とする方法を提案している (基準値としてどのような値を選ぶかは Onatski (2010) を参照)

4. 実証例

■ 実証例

ここではファクターモデルを用いた実証例として Stock and Watson (2002) と Shintani (2005) を簡単に紹介しておく。

4. 実証例

■ Stock and Watson (2002, JBES)

Stock and Watson (2002) は (total industrial production や real manufacturing などの) US の 8 つの月次のマクロ経済変数について、ファクターモデル、ARモデル、VARモデルを含むいくつかの予測モデルの比較を、サンプル外予測比較によって行っている。そこではファクターの推定に最大で215個の時系列を用いている(データ期間は1959:1から1998:12)。

4. 実証例

彼らはファクター以外の変数として、その系列自体の過去の値を使用している(次数はBICで選択)ファクターの数は1, 2, 3, 4のいずれかで固定してる(情報量基準などは用いていない)。

結果は全ての場合において、ファクターモデルはARやVARなどのベンチマークモデルよりもよいパフォーマンスを示し(ただしファクターの数にもよる)、またファクターの数を多くしても必ずしも予測精度は改善しないという結果も得た。さらに215個の変数のほとんどが6つほどのファクターでほぼ説明できることも発見している。

4. 実証例

■ Shintani (2005, JMCB)

Shintani (2005)はStock and Watson (2002) 達の結果を用い、ファクターが非線形なARモデルに従う場合にモデルを拡張している。提案した手法を日本のデータに応用し(ファクターの抽出には235個の系列を用いている)、日本のデータに対してもファクターモデルによる予測が有用であると結論している。また推定したファクターを用いて、ファクターモデルの非線形性も検定し、いくつかのファクターについては非線形性が認められるとしている。

まとめ

1. ファクターモデルの推定手法を紹介した
2. ファクターの数の選択は情報量基準や行列の固有値を用いる方法があることを紹介した

参考文献

Shintani, M. (2005) “Nonlinear Forecasting Analysis Using Diffusion Indexes: An Application to Japan,” *Journal of Money, Credit and banking*, 37(3), 517-538.

Stock, J.H. and Watson, M. W. (2002) “Macroeconomic Forecasting Using Diffusion Indexes”, *Journal of Business and Economic Statistics*, 20(2), 47-162.