（5）「多変量解析と時系列解析のハザマ」に関する研究報告

大規模マルチファクターモデルの推定について

統計数理研究所　川崎能典

1　はじめに

株価収益率を個別企業の財務データやテクニカル指標 (以下，ファクター) で説明しようとする試みは 1960 年代から一般的になっている．その捉え方や解釈は一様ではないが，いわゆるアクティブ運用の立場からすれば，マルチファクターモデルは予測モデルである．最もよく用いられる手法は，各時点で個別に OLS 回帰して求めた係数 (リターン係数) を，適当な個数だけとって片側後方移動平均して予測に用いるというものである．このような事後的な平滑化に対し，本報告では平滑化事前分布を利用したモデリングを試みる．

2　OLS アプローチ

ある月 t における投資対象ユニバース (例えば東証一部上場銘柄すべて) の収益率を r_t，説明変数である前月の個別企業のファクターを F_{t-1} とする．ファクター行列には定数項も含まれるが，定数項は推定しても予測には用いない．定数項はいわばある時点における十分に分散化されたポートフォリオの収益率であり，その時々の相場平均と考えられる．従ってこの部分は予測不能である．(可能なら先物取引に専念すればよいのであって，以下のようなファクターモデルを考える必要がない．)

さて，機関投資家などで一般に行われているのは，OLS 回帰で $\hat{\beta}_t = (F'_{t-1}F_{t-1})^{-1}F'_{t-1}r_t$ を求め，(例えば)60ヶ月片側後方移動平均によって平滑化したリターン係数ベクトル $\tilde{\beta}_t$ を使って翌月の収益を予測するというものである．(r_t, F_t が揃った時点で上の回帰を行い，$F_t\tilde{\beta}_t$ をもとにポートフォリオの組み替えを行う．) 事後的に平滑化を行うのは，単に $\hat{\beta}_t$ を使って一期先予測をすることがよくないことが経験的に知られているからであるが，ここでの目標はそのメカニズムを明らかにしながら，平滑化事前分布によって予測の意味での最適な平滑化をデータから決める仕組みを与えることである．

この問題は時変係数重回帰モデルと捉えることができるので，状態空間モデルの極めて標準的な適用例に帰着する．とりあえず線形ガウスから分析を始めるとすれば，カルマンフィルタ等の逐次公式の恩恵に与ることができる．しかし，OLS 回帰の枠組みを維持したままで係数に平滑化事前分布を導入して解いても，結果は OLS 回帰の係数そのものに殆ど一致する．これは，観測次元が事前情報の次元を圧倒していることも一因であろうが，回帰モデルの共分散行列構造のクラスが狭いためにカルマンゲインのクラスが限られ，結果的に予測誤差の殆どが状態モデルの変動と見なされてしまうことに大きな原因があると考えられる．

3　Temporal Effect Model

そこで本稿では，標準的なモデルを以下のように拡張し，これを temporal effect (TE) model と呼ぶ．

$$\beta_t = \beta_{t-1} + v_t, \qquad v_t \sim \mathrm{NID}(0, D_1)$$

$$\begin{aligned} r_t &= F_{t-1}(\beta_t + w_t) + u_t \\ w_t &\sim \mathrm{NID}(0, D_2) \\ u_t &\sim \mathrm{NID}(0, \sigma^2 I_N) \end{aligned}$$

ここで $D_1 = \mathrm{diag}(\tau_1^2, \ldots, \tau_K^2)$，$D_2 = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_K^2)$ である．簡単な変形によって明らかなように，これは OLS 型の特定化を少数のパラメタによって定義される GLS 型に拡張していることに等しい．TOPIX 採用銘柄を対象にした解析では，TE モデルは TE がないモデル (事後的平滑化のない OLS にほぼ対応) に比べて尤度にして約 64 改善しており，パラメタ数を考慮しても TE モデルがよいと判断される．

4 運用パフォーマンスによる評価

OLS と移動平均を組み合わせた従来型の方法も俎上に乗せて比較するには，尤度以外に別種の規準が必要である．ここでは，予測収益率をもとに市場中立型のロングショート運用を毎月行って，そのパフォーマンスを年率換算のシャープレシオで表示して比較することとした．この評価規準の下では売り買いの方向性 (符号) が大切で，必ずしも最尤法と合致するものではないが，実際上の観点からは重要な指標である．データは 1985 年 1 月から 1997 年 10 月までの TOPIX 採用銘柄を，ファクターは，一ヶ月騰落率，キャッシュフロー株価比，配当利回りを用いている．OLS+移動平均のシャープレシオは 2.85%，TE モデルは 2.73%，TE なしのモデルは 1.91%であった．なお，モデルのパラメタを再推定しながらアウトサンプルで予測したときは 2.96%で，インサンプルの場合と遜色ない結果である．少なくともここでの分析からは，状態空間モデルで得られたリターン係数に従った運用は，OLS+移動平均と同程度のシャープレシオであっても，平均リターンがやや小さい分リスクも小さくなっている傾向が認められる．詳しくは Kawasaki et al (1998) 参照．

5 高階差分，初期化，計算コスト

平滑化事前分布を用いる方法の実用上の難点は，膨大な計算時間がかかることである．これは，銘柄数を N としたときに，カルマンゲインの計算において $N \times N$ 行列の逆行列の評価を行わなければならないことに起因する．特に TOPIX 採用銘柄を投資ユニバースと考えたときには $N = 1341$ であり，最も効率がよい処理方法でも，ひとつのモデル推定を終えるのに 30 時間を要する．(IBM RS/6000 を 11 並列用いて，FORTRAN でプログラミングした場合である．) 最適化コストを削減するため，逐次フィルタの初期分布の与え方と平滑化事前分布の差分の次元をさまざまに組み合わせてモデル推定を行った．日経平均採用銘柄を元にした分析から得た結論としては，差分の次元は 1 次で十分であり，初期状態の与え方は後ろ向きフィルタの併用か，diffuse prior によるのが妥当であろう．

参考文献

Kawasaki, Y., Sato, S. and Tachiki, S. (1998) Smoothness Prior Approach to Estimate Large Scale Multifactor Models, ISM Research Memorandum No.714.

多変量時系列解析によるシステム解析について

NTT コミュニケーション科学研究所　加藤比呂子

1 はじめに

これまで経済や生体システムの動的解析に多変量時系列モデルを適用し、データ間の関係についての議論をおこなってきた (Kato et al (1994, 1996)、加藤、石黒 (1997)、 Kato and Kawahara (1998)). それらの統計的システム解析の部分では、中川、赤池 (1972) による統計的フィードバックシステム解析を適用した. 線形システム解析をベースとしているので、モデルの扱いや解釈が簡単であり、スペクトル、応答関数により、変数間の因果関係を把握することができ、物価の変動に関する経済システムや、聴覚システムと音声生成機構との関係について調べることができた.

筆者が聴覚フィードバック解析に携わり実感したのは、実際収録される生体時系列データの多くは、非定常性はもちろん、そのダイナミクスに、単なる線形のシステム解析では見えない非線形的な要素が含まれているということである. 自然界におけるこのような非線形ダイナミクスについては、すでに物理の分野でカオス的アプローチの適用により多く研究されているが、観測データからシステムの特性を直接引き出す方法とは言い難い. そのようなデータに対して、非線形時系列解析は必要であることは以前から注目されていて、様々なモデルが研究されている. 中でも、システムにおける非線形ダイナミクスの再構築を目的としたモデルとしては、 Generalized ExpAR (GExpAR) モデル (Ozaki et al. (1997)) が新しく提案されている. これは、もともとリミットサイクルの非線形振動機構を AR モデルに組み込み、非線形時系列モデルとして良いパフォーマンスを示した Exponential AR (ExpAR) モデル (Ozaki (1985)) に対し、状態依存モデルと非線形関数の RBF 近似を併せて一般化したものである.

また、非線形システム解析への一考察としては、 Kato 等 (1998) がある時系列の特性に他の関連した信号が関与した場合を想定して Coefficient-Modulated AR (CMAR) モデル を考案している. このモデルは、 Functional-Coefficient AR モデル (Chen and Tsay (1993)) の係数部分に、他信号が multiplicative に入った拡張的なモデルの形という見方もできる.

本報告では、上記に記した解析例をまとめてシステム解析に役に立ちそうな多変量時系列モデルと解析例を紹介した. 用いたデータは、音声の基本周波数データと、その揺らぎに関係していると報告されている (Orlikoff and Baekn (1989)) 耳脈のデータから得られた心拍のデータ、それを用いて生成したシミュレーションデータである.

2 モデル例

2.1 VAR モデルによるシステム解析

Kato and Kawahara(1998) を参照.

2.2 Radial Basis Function-AR 型モデル

観測された 1 次元のデータ系列を $x_n (n = 1, \cdots, N)$ とする. ExpAR モデルにおいて、一期前の状態に依存する AR の係数部分を RBF を用いて近似する.

$$
\begin{aligned}
x_n &= \phi_0(X_{t-1}) + \sum_{i=1}^{p} \phi_i(X_{n-1}) x_{n-i} + e_n \quad e_n \sim N(0, \sigma^2) \\
\phi_i(X_{n-1}) &= c_{i,0} + \sum_{k=1}^{m} c_{i,k} \prod_{j=1}^{d} \exp(-\frac{(x_{n-j} - z_{k,j})^2}{h_k})
\end{aligned}
$$

このモデルは基本周波数と同時に測定された心拍のデータに適用した. モデルの係数部分は全て状態の関数であり、それらを用いて時間推移に伴った instantaneous な固有値の変化を確かめることができた. また普通の線形 AR モデルと周期的データに適用する subset AR モデルとの比較をおこなうと、AIC の意味では、 RBF-AR モデルがよくフィットしていることがわかる. また、基本周波数との関係を調べるための instantaneous VAR モデルについての解析例を示した.

2.3 Coefficient-Modulated AR モデル

観測された1次元のデータ系列 $y_n(n = 1, \cdots, N)$ の特性に対して、関連したもうひとつのデータ系列 $x_n(n = 1, \cdots, N)$ が関与するとする.

$$y_n = \sum_{m=1}^{M} a_m(x|\theta)y_{n-m} + \epsilon_n, \qquad \epsilon_n \sim N(0, \sigma^2)$$

$a_m(x|\theta)$ は x の関数である AR の係数を表す.

RBF-AR 型モデルと同様に推定された $a_m(x|\theta)$ から instantaneous な固有根を計算することができる. もし根が単位円内の場合、時間推移に伴う instantaneous なパワースペクトルのをみることができる. 心拍のデータを用いて作成したシミュレーションデータに適用した.

謝辞

本研究の動機付けは、ATR 人間情報通信研究所との共同研究にあります. また非線形時系列解析に関して議論させて頂いている統計数理研究所 尾崎統教授、データ解析を手伝って頂いた施招雲博士に感謝します.

参考文献

1. Kato, H., Kawahara, H. (1998). An Application of the Bayesian Time Series Model and Statistical System Analysis for F0 Control, *Journal of Speech Communication*, Vol. 24(4), pp.325-339.
2. 加藤比呂子, 石黒真木夫 (1997). 多変量時系列モデルによる経済システムの動的解析, 統計数理, 45 巻, 2 号, 303 - 321.
3. Kato, H., Naniwa, S., Ishiguro, M. (1996). A Bayesian Multivariate Nonstationary Time Series Model for Estimating Mutual Relationships among Variables, *Journal of Econometrics*, Vol. 75, pp.147-161.
4. Kato, H., Wada, T., Ishiguro, M. (1994). A Study of Human Body Balance by New Multivariate Feedback Models with Common Low Frequency Components, *Japanese Journal of Biometrics*, Vol. 15, pp. 41-57.
5. Akaike, H. and Nakagawa, T. (1988), Statistical analysis and control of dynamic systems, Kluwer Scientific Publishers, Tokyo.
6. Ozaki, T. (1985), Non-linear time series models and dynamical systems, Handbook of Statistics, 5, E.J. Hannan, P.R. Krishnaiah and M. M. Rao Eds., North-Holland, Amsterdam, 25-83.
7. Ozaki, T., Sosa, P.V., and Haggan-Ozaki, V. (1997), Reconstructing Nonlinear Dynamics from Time Series : With Application to Epilepsy Data Analysis, Manchester Centre for Statistical Science Technical Report 1997/02.
8. Chen, R. and Tsay, R. S. (1993), Functional-Coefficient Autoregressive Models, Journal of the American Statistical Association, 88, 421, 298-308.
9. Kato, H., Kawahara, H. and Ishiguro, M. (1998), On the Coefficient- Modulated Autoregressive Model, Statistical Research Group Technical Report, No. 98-12, Hiroshima University.
10. Orlikoff, R. F. and Baekn, R.J. (1989), "The effect of the heartmeat on vocal fundamental derequency perturmation", Journal of Speech and Hearing Research, 32, pp. 576-582.

Simplified estimator and simultaneous confidence region in an extended growth curve model

広島工大・環境　神田隆至
広島大・原医研　大瀧　慈
広島大・　理　藤越康祝

Potthoff and Roy (1964) によって導入された通常の成長曲線モデルの拡張として，いくつかの異なる個体内計画行列をもつ次のモデルを考える。

$$E(Y) = A_1\Xi_1 X_{(1)} + \cdots + A_k \Xi_k X_{(k)} \tag{1}$$

ここに, Y は $N \times p$ の 観測値行列で $Y \sim N_{N\times p}(E(Y), I_N \otimes \Sigma)$, A_i は $N \times r_i$ の階数 r_i の既知の個体間計画行列, $X_{(i)}$ は $q_i \times p$ の階数 q_i の既知の個体内計画行列, Ξ_i は $r_i \times q_i$ の未知のパラメター行列である。さらに, $X_{(i)}$ に次の階層的構造を仮定する。

"$q_1 < q_2 < \cdots < q_k$ かつ $X_{(i)}$ は $X_{(k)}$ の始めの $q_i \times p$ の部分行列と同じである。"　(2)

このモデルは, 特別な場合として次数が異なる $k-$標本多項式成長曲線を含んでいる。

条件 (2) をみたすモデル (1) の標準形は次のように与えられる。

$$Z \sim N_{N\times p}(E(Z), I_N \otimes \Omega), \tag{3}$$

ここに,

$$E(Z) = E\left(\begin{bmatrix} Z_{11} & \cdots & Z_{1\ell} \\ \vdots & \ddots & \vdots \\ Z_{k1} & \cdots & Z_{k\ell} \\ Z_{01} & \cdots & Z_{0\ell} \end{bmatrix}\right) = \begin{bmatrix} \Theta_{11} & O & \cdots & O \\ \vdots & \ddots & & \vdots \\ \Theta_{k1} & \cdots & \Theta_{kk} & O \\ O & \cdots & O & O \end{bmatrix}$$

であり, Z_{ij} は $r_i \times b_j (i = 1, \ldots, k; j = 1, \ldots, \ell)$, Z_{0j} は $n \times b_j (j = 1, \ldots, \ell)$, Ω は未知の正定値行列, $n = N - r_1 - \cdots - r_k$, $\ell = k+1$, $b_1 = q_1$, $b_2 = q_2 - q_1, \ldots, b_k = q_k - q_{k-1}$, $b_{k+1} = b_\ell = p - q_k$ である。また,

$$W = [Z_{01} \ \ldots \ Z_{0\ell}]'[Z_{01} \ \ldots \ Z_{0\ell}] = [W_{ij}], \quad W_{ij} = Z'_{0i} Z_{0j}$$

とする。

モデル (3) のもとでの MLE, 検定問題は Fujikoshi and Sato (1996), Fujikoshi, kanda and Ohtaki (1998) 等で考察されている。本報告では, MLE's と密接に関連している簡便推定量を提案し，その推定量の基本的性質に基づいて, 平均パラメータの同時信頼区間の構成を提案する。行列の分割に対して, 次の記号を用いる。

$$Z_i = [Z_{i1} \ Z_{i2} \ \ldots \ Z_{i\ell}], \qquad Z_{i(12\ \ldots i)} = [Z_{i1} \ Z_{i2} \ \ldots \ Z_{ii}],$$

$$\Omega_{(2\ldots\ell)1} = \begin{bmatrix} \Omega_{21} \\ \vdots \\ \Omega_{\ell 1} \end{bmatrix}, \ \Omega_{(2\ldots\ell)(2\ldots\ell)} = \begin{bmatrix} \Omega_{22} & \cdots & \Omega_{2\ell} \\ \vdots & & \vdots \\ \Omega_{\ell 2} & \cdots & \Omega_{\ell\ell} \end{bmatrix},$$

$$\Omega_{11\cdot 23\ldots\ell} = \Omega_{11} - \Omega_{1(23\ldots\ell)} \Omega^{-1}_{(23\ldots\ell)(23\ldots\ell)} \Omega_{(23\ldots\ell)1} \text{ 等々。}$$

ここで用いるモデル (1) の平均パラメータの推定量は, 最尤法の考え方に基づいて

$$\begin{aligned}
\hat{\Theta}_{11} &= Z_{11} - Z_{1(23\ldots\ell)}\hat{\Omega}^{-1}_{(23\ldots\ell)(23\ldots\ell)}\hat{\Omega}_{(23\ldots\ell)1},\\
\hat{\Theta}_{2(12)} &= Z_{2(12)} - Z_{2(3\ldots\ell)}\hat{\Omega}^{-1}_{(3\ldots\ell)(3\ldots\ell)}\hat{\Omega}_{(3\ldots\ell)(12)},\\
&\vdots\\
\hat{\Theta}_{k-1(12\ldots k-1)} &= Z_{k-1(12\ldots k-1)} - Z_{k-1(k\ell)}\hat{\Omega}^{-1}_{(k\ell)(k\ell)}\hat{\Omega}_{(k\ell)(12\ldots k-1)},\\
\hat{\Theta}_{k(12\ldots k)} &= Z_{k(12\ldots k)} - Z_{k\ell}\hat{\Omega}^{-1}_{\ell\ell}\hat{\Omega}_{\ell(12\ldots k)},
\end{aligned}$$

によって推定する。ただし, 共分散行列 Ω を簡便推定量

$$n\hat{\Omega} = W$$

で推定する。Ω の簡便推定量は $Z_{1(2\ldots\ell)}, Z_{2(3\ldots\ell)}, \ldots, Z_{k-2(k-1\ldots\ell)}$ と $Z_{k\ell}$ の情報を無視して, $Z_{0(1,2\ldots\ell)}$ だけを用いたものである。

次に, $C_i : c_i \times r_i, D_i : q_i \times d_i (i = 1, \ldots, k)$ を $\mathrm{rank}(C_i) = c_i \leq r_i, \mathrm{rank}(D_i) = q_i \leq d_i$ の既知行列とし, すべての1次結合 $\mathbf{a}_1'C_1\Theta_{11}D_1\mathbf{b}_1, \mathbf{a}_2'C_2\Theta_{2(12)}D_2\mathbf{b}_2, \ldots, \mathbf{a}_k'C_k\Theta_{k(1\ldots k)}D_k\mathbf{b}_k$ について信頼係数 $1-\alpha$ の同時信頼区間を構成する問題を考える。$\mathbf{a}_i'C_i\Theta_{i(1\ldots i)}D_i\mathbf{b}_i$ に関する場合は通常の成長曲線モデルにおける方法によって構成される。具体的には枢軸統計量

$$\eta_i(\mathbf{a}_i, \mathbf{b}_i) = \frac{\mathbf{a}_i'C_i(\hat{\Theta}_{i(1\ldots i)} - \Theta_{i(1\ldots i)})D_i\mathbf{b}_i}{\{\mathbf{a}_i'C_i[I_{r_i} + Z_{i(i+1\ldots\ell)}W^{-1}_{(i+1\ldots\ell)(i+1\ldots\ell)}Z'_{i(i+1\ldots\ell)}]C_i'\mathbf{a}_i\}^{1/2}(\mathbf{b}_i'D_i'\hat{\Omega}_{(1\ldots i)(1\ldots i)\cdot i+1\ldots\ell}D_i\mathbf{b}_i)^{1/2}}$$

を考え,

$$\begin{aligned}
S_h^{(i)} &= \{C_i(\hat{\Theta}_{i(1\ldots i)} - \Theta_{i(1\ldots i)})D_i\}'\{C_i[I_{r_1} + Z_{i(i+1\ldots\ell)}W^{-1}_{(i+1\ldots\ell)(i+1\ldots\ell)}Z'_{i(i+1\ldots\ell)}]C_i'\}^{-1}\\
&\quad\times\{C_i(\hat{\Theta}_{i(1\ldots i)} - \Theta_{i(1\ldots i)})D_i\},\\
S_e^{(i)} &= D_i'W_{(1\ldots i)(1\ldots i)\cdot i+1\ldots\ell}D_i
\end{aligned}$$

とし, $\ell^{(i)}$ を $S_h^{(i)}(S_e^{(i)})^{-1}$ の最大固有根, $r_{d_i,c_i,\tilde{n}_i}(\alpha), (\tilde{n}_i = n - p_{i+1} - \cdots - p_\ell)$ を $\ell^{(i)}$ の上側 α 点とすると, $\mathbf{a}_i'C_i\Theta_{i(1\ldots i)}D_i\mathbf{b}_i$ の信頼限界は

$$\begin{aligned}
&\mathbf{a}_i'C_i\hat{\Theta}_{i(1\ldots i)}D_i\mathbf{b}_i\\
&\pm[r_{d_i,c_i,\tilde{n}_i}(\alpha)(\mathbf{a}_i'C_i[I_{r_1} + Z_{i(i+1\ldots\ell)}W^{-1}_{(i+1\ldots\ell)(i+1\ldots\ell)}Z'_{i(i+1\ldots\ell)}]C_i'\mathbf{a}_i)\cdot(\mathbf{b}_i'D_i'\hat{\Omega}_{(1\ldots i)(1\ldots i)\cdot i+1\ldots\ell}D_i\mathbf{b}_i)]^{1/2}
\end{aligned}$$

で与えられる。

次にこれらの結果を用いて, $\mathbf{a}_1'C_1\Theta_{11}D_1\mathbf{b}_1, \ldots, \mathbf{a}_k'C_k\Theta_{k(1\ldots k)}D_k\mathbf{b}_k$ の同時信頼区間を求めことができる。個別の信頼区間の信頼係数を適当に大きく設定することにより, 目標の信頼係数 $1-\alpha$ をもつ同時信頼区間を構成することができる。しかし, k が増えるにつれて. 個々の信頼係数を非常に大きくする必要がある。ここで提案するのは

$$\eta_i^2(\mathbf{a}_i, \mathbf{b}_i) \leq \max(\tilde{n}_1\ell^{(1)}, \ldots, \tilde{n}_k\ell^{(k)}) = T^2_{\max}$$

の上側の点を利用する方法である。ここに, $\tilde{n}_i = n - p_{i+1} - \cdots - p_\ell$ である。$T^2_{\max}(\alpha)$ を $T^2_{\max}$ の上側 α 点とする。一般に $T^2 = T^2_{\max}$ の分布を得るのは難しいが, $c_1 = \cdots = c_k = 1$ のとき

$$\lim_{n\to\infty} P(T^2 \leq x) = G_{q_1}(x)G_{q_2}(x)\cdots G_{q_k}(x)$$

および, この結果の精密化が与えられる。ここに, $G_q(x)$ は自由度 q の χ^2 分布の分布関数である。

因子分析における独自分散の最尤推定量の分布の漸近展開

東京外国語大学外国語学部　市川雅教
九州大学大学院数理学研究科　小西貞則

はじめに　因子分析モデルの母数の最尤推定量の漸近分布は，Anderson and Rubin (1956) や Lawley and Maxwell (1971) などで与えられている．しかしながら，観測変数の数に比して標本が大きくない場合には，これらの結果を用いた正規近似の精度が必ずしも十分ではない．ここでは，因子分析における独自分散の最尤推定量の分布の漸近展開と，その精度を数値的に評価した結果について報告した．

因子分析モデルと母数の最尤推定　因子分析モデルのもとで，観測される p 個の変数 $x_1,\ldots,x_p$ の共分散行列 $\mathbf{\Omega}=(\omega_{ij})$ は，$\mathbf{\Omega}=\mathbf{\Lambda}\mathbf{\Lambda}'+\mathbf{\Psi}$ と分解される．ここで，$p\times m$ 行列 $\mathbf{\Lambda}$ は因子負荷量行列とよばれ，対角行列 $\mathbf{\Psi}$ の i 番目の対角要素 $\psi_i(>0)$ は独自分散とよばれる．x_i の分散 ω_{ii} のうち，共通因子により説明されない部分の割合 $\psi_i^*=\psi_i/\omega_{ii}$ は，独自性とよばれる．多変量正規母集団 $\mathrm{N}_p(\boldsymbol{\mu},\mathbf{\Omega})$ からの大きさ $N=n+1$ の無作為標本にもとづく $\mathbf{\Omega}=\mathbf{\Lambda}\mathbf{\Lambda}'+\mathbf{\Psi}$ の不偏推定量を $\mathbf{S}$ とし，その尤度にもとづく $\mathbf{\Lambda},\mathbf{\Psi}$ の最尤推定量を $\hat{\mathbf{\Lambda}},\hat{\mathbf{\Psi}}$ とする．

標本共分散行列の関数の分布の漸近展開　いま，$\mathbf{U}=n^{1/2}(\mathbf{S}-\mathbf{\Omega})$ とし，$\mathbf{S}$ の関数 $h(\mathbf{S})$ に対して，$w_n=n^{1/2}\{h(\mathbf{S})-h(\mathbf{\Omega})\}$ が $w_n\approx\mathrm{tr}\,\mathbf{AU}+\frac{1}{2}n^{-1/2}\sum_{a,b,c,d}t(ab,cd)u_{ab}u_{cd}$ とテイラー展開できるものとする．このとき，w_n の漸近バイアス，漸近分散，漸近歪度を b_1,σ^2,b_3 とすると，b_1, σ^2, b_3 は次式で与えられる (Siotani, Hayakawa and Fujikoshi, 1985).

$$b_1=\frac{1}{2}\sum_{a,b,c,d}(\sigma_{ac}\sigma_{bd}+\sigma_{ad}\sigma_{bc})t(ab,cd),\ \ \sigma^2=2\mathrm{tr}\,(\mathbf{A\Omega})^2,\ \ b_3=12T+8\mathrm{tr}\,(\mathbf{A\Omega})^3$$

ただし，$T=\sum_{a,b,c,d}t(ab,cd)[\mathbf{\Omega A\Omega}]_{ab}[\mathbf{\Omega A\Omega}]_{cd}$ とする．また，w_n の分布の漸近展開は，

$$\mathrm{P}(\frac{w_n}{\sigma}\le x)=\Phi(x)-n^{-1/2}\left\{\frac{b_1}{\sigma}+\frac{b_3}{6\sigma^3}(x^2-1)\right\}\phi(x)+O(n^{-1})$$

で与えられる．ここで，$\Phi(\cdot)$ と $\phi(\cdot)$ は，それぞれ標準正規分布の分布関数と密度関数である．

次に，$\hat{\sigma}^2=\sigma^2(\mathbf{S})$ を σ^2 の一致推定量とし，ステューデント化した統計量を $u_n=w_n/\hat{\sigma}$ とする．$\sigma^2(\mathbf{S})$ が $\mathbf{S}=\mathbf{\Omega}$ の近傍で微分可能で，$\hat{\sigma}^2-\sigma^2\approx n^{-1/2}\mathrm{tr}\,\mathbf{HU}$ とすると，u_n の漸近バイアス b_1' と漸近歪度 b_3' は

$$b_1'=\sigma^{-1}b_1+\sigma^{-3}\mathrm{tr}\,\mathbf{A\Omega H\Omega},\ \ b_3'=\sigma^{-3}b_3-6\sigma^{-3}\mathrm{tr}\,\mathbf{A\Omega H\Omega}$$

で与えられ，u_n の分布の漸近展開は次式で与えられる．

$$\mathrm{P}(u_n\le x)=\Phi(x)-n^{-1/2}\left\{b_1'+\frac{b_3'}{6}(x^2-1)\right\}\phi(x)+O(n^{-1})$$

独自分散と独自性の最尤推定量の漸近バイアス，漸近分散，漸近歪度　いま，$\hat{\psi}_i$ が $\hat{\psi}_i=\psi_i+n^{-1/2}p_i^{(1)}+n^{-1}p_i^{(2)}+\cdots$ と展開できることを仮定し，摂動法により $\hat{\psi}_i$ の近似を求める．行列 $\mathbf{\Phi}$ を $\mathbf{\Phi}=\mathbf{\Psi}^{-1}-\mathbf{\Psi}^{-1}\mathbf{\Lambda}(\mathbf{\Lambda}'\mathbf{\Psi}^{-1}\mathbf{\Lambda})^{-1}\mathbf{\Lambda}'\mathbf{\Psi}^{-1}$ とし，$\Xi\equiv(\phi_{ij}^2)>0$ を仮定する．$\Xi^{-1}=(\xi^{ij})$ とし，Ξ^{-1}

の第 i 行 (列) を対角要素とする対角行列を $\mathbf{\Delta}_i$ とすると，$p_i^{(1)} = \mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi U}, p_i^{(2)} = \mathrm{tr}\,\mathbf{\Delta}_i\mathbf{Q}$ を得る．ここで，

$$\begin{aligned}\mathbf{Q} &= -2\mathbf{\Phi}(\mathbf{U}-\mathbf{P}^{(1)})\mathbf{\Phi}(\mathbf{U}-\mathbf{P}^{(1)})\mathbf{B} + 2\mathbf{\Phi}(\mathbf{U}-\mathbf{P}^{(1)})\mathbf{\Phi P}^{(1)}(\mathbf{\Psi}^{-1}-\mathbf{\Phi}) \\ &\quad -\mathbf{\Phi}(\mathbf{U}-\mathbf{P}^{(1)})\mathbf{B}(\mathbf{U}-\mathbf{P}^{(1)})\mathbf{\Phi}\end{aligned}$$

であり，$\mathbf{B} = \mathbf{\Psi}^{-1}\mathbf{\Lambda\Gamma}^{-2}\mathbf{\Lambda}'\mathbf{\Psi}^{-1}$ とする．以上のことから，$h(\mathbf{S}) = \hat{\psi}_i$ の場合，

$$\begin{aligned}b_1 &= -\psi_i \mathrm{tr}\,(\mathbf{I}+\mathbf{\Gamma}^{-1}) - 2\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi}(\mathbf{B}\odot\mathbf{\Xi}^{-1}) \\ \sigma^2 &= 2\xi^{ii} \\ \mathrm{tr}\,(\mathbf{A\Omega})^3 &= \mathrm{tr}\,(\mathbf{\Phi\Delta}_i)^3 \\ T &= -2\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{B\Delta}_i \\ \mathrm{tr}\,\mathbf{A\Omega H\Omega} &= 4\mathrm{tr}\,(\mathbf{\Phi\Delta}_i)^3 - 8\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{B\Delta}_i\end{aligned}$$

一方，$h(\mathbf{S}) = \hat{\psi}_i^*$ の場合，

$$\begin{aligned}b_1 &= -\psi_i^*\mathrm{tr}\,(\mathbf{I}+\mathbf{\Gamma}^{-1}) - 2\omega_{ii}^{-1}\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi}(\mathbf{B}\odot\mathbf{\Xi}^{-1}) + 2\psi_i^*(1-\psi_i^*) \\ \sigma^2 &= 2\omega_{ii}^{-2}\xi^{ii} + 2\psi_i^{*2}(1-2\psi_i^*) \\ \mathrm{tr}\,(\mathbf{A\Omega})^3 &= \omega_{ii}^{-3}\mathrm{tr}\,(\mathbf{\Phi\Delta}_i)^3 - 3\omega_{ii}^{-1}\psi_i^{*3}[\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{\Phi}]_{ii} - \psi_i^{*3}(1-3\psi_i^*) \\ T &= -2\omega_{ii}^{-3}\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{B\Delta}_i + 4\omega_{ii}^{-1}\psi_i^{*3}[\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{\Phi}]_{ii} \\ &\quad +2\psi_i^*(1-3\psi_i^*)\{\omega_{ii}^{-2}\xi^{ii} + \psi_i^{*2}(1-\psi_i^*)\} \\ \mathrm{tr}\,\mathbf{A\Omega H\Omega} &= 4\omega_{ii}^{-3}\mathrm{tr}\,(\mathbf{\Phi\Delta}_i)^3 - 8\omega_{ii}^{-3}\mathrm{tr}\,\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{B\Delta}_i \\ &\quad +4\omega_{ii}^{-1}\psi_i^{*3}[\mathbf{\Phi\Delta}_i\mathbf{\Phi\Delta}_i\mathbf{\Phi}]_{ii} + 4\psi_i^*(1-3\psi_i^*)\{2\omega_{ii}^{-2}\xi^{ii} + \psi_i^{*2}(1-2\psi_i^*)\}\end{aligned}$$

数値例 Emmett (1949) による相関係数行列 ($p = 9$) を標本共分散行列とみなして最尤法により 2 因子モデルをあてはめ，得られた推定値を母数として用いた．標本の大きさが 100 と 200 の場合のそれぞれについて，シミュレーションによりステューデント化した統計量の実現値を 10^6 回求め，確率分布 $\mathrm{P}(u_n \le x)$ やバイアス，分散，歪度，尖度の正確な値を数値的に求めた．その結果，真の値が 1 に近い場合の独自性を除いては，漸近展開式の近似精度が良好であることが分かった．

文献

Anderson, T. W. and Rubin, H. (1956) Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, **5**, 111–150.

Emmett, W. G. (1949) Factor analysis by Lawley's method of maximum likelihood. *British Journal of Psychology, Statistical Section*, **2**, 90–97.

Lawley, D. N. and Maxwell, A. E. (1971) *Factor Analysis as a Statistical Method.* (2nd ed.) London: Butterworths.

Siotani, M., Hayakawa, T., and Fujikoshi, Y. (1985) *Modern Multivariate Statistical Analysis: A Graduate Course and Handbook.* Columbus: American Sciences Press.

反復測定分散分析理論と、教育・心理の分野における適用の現状について

愛知学院大学文学部 千野　直仁

1　はじめに

縦断的データもしくは経時測定データの解析法としては、これまでに、ANOVA、MANOVA、GMANOVA 等の 1 変量・多変量解析、各種定常・非定常時系列解析、マルコフ連鎖、分岐過程、点過程等の確率過程分析法、3 相因子分析、STATIS、PVA、DYNASCAL (Chino and Nakagawa, 1990) 等の計量心理学的データ解析、力学系、複素力学系等の力学系モデルによる解析などが、提案されている。

一般に、教育や心理の分野では、同一被験者にいろいろな条件を課して反復測定を行う場合が多々ある。それらの代表的なケースは、学習実験での練習効果や忘却経過の測定、各種精神作業や一般の作業における作業量や疲労の測定、脳波や呼吸数などの測定、教育や心理の各分野、とりわけ実験心理学の分野でのサンプル数節約や少ないサンプル数による実験条件の均質化のための反復測定などである。

本報告では GMANOVA も含めたいわゆる反復測定（測度）分散分析の理論及び教育や心理の分野におけるそれらの適用の現状と問題点について、以下の 3 項目、とりわけ第 1 項目について詳しく報告した：

- 教育・心理の分野での反復測定分散分析の現状
- 反復測定分散分析の理論の現状
- 現状での球形仮説 (sphericity hypothesis) への対応と今後の課題

2　反復測定分散分析の現状と今後の課題

まず、教育・心理の分野での反復測定分散分析の現状については、これらの分野での反復測定分散分析デザインは非常にポピュラーであるが、反復測定データに対する同デザイン分散分析の適切な適用例は全般にきわめて少ないことを指摘した。

つぎの反復測定分散分析の理論の現状、としては以下の項目に分けて報告した：

- 反復測定データに対する同デザイン分散分析の利用の現状
- F-比が歪まないための必要十分条件についての証明の歴史

- F-比の歪みに対する内外の楽観的見解とその危険性
- 反復測定分散分析の入門書の不足
- 国際的統計ソフトである SAS、SPSS 等の対応の現状

とりわけ、F-比の歪みに対する内外の楽観的見解とその危険性については、筆者のSAS による簡単なシミュレーション結果を引用し、強調した。また、国際的な統計解析パッケージである、SAS や SPSS でさえ、幾つかの点で不十分であり、注意して利用しないと間違った結論を下す危険性があることを指摘した。

最後の、現状での球形仮説への対応と今後の課題の項では、反復測定デザインデータに対する以下のような球形仮説への対応方法

1. まず球形検定を行い、球形仮説が満たされていると判断出来る時は、通常の（修正なしの）F-検定を、満たされていない時は、Box/G-G のイプシロン修正をした近似 F-検定を、それぞれ行う、
2. 3段階 G-G 法を行う、
3. 最初からイプシロン修正をした近似 F-検定を行う、
4. 球形検定として LBI 検定 (Locally Best Invariant test) を用い、あとはうえと同様の手続きによる、
5. Box/G-G イプシロンによる球形検定を行う、
6. ノンパラメトリック検定を行う、

の長短の一層綿密な検討の必要性に言及した。

参考文献

千野 直仁 (1993). 反復測度デザイン概説ーその1　愛知学院大学文学部紀要 **23**, 223-235.

千野 直仁 (1994). 反復測度デザイン概説ーその2　愛知学院大学文学部紀要 **24**, 103-119.

千野 直仁 (1995). 教育や心理の分野における ANOVA、MANOVA、GMANOVA 適用上の問題点　愛知学院大学文学部紀要 **25**, 71-96.

多次元正規定常過程における標本自己共分散行列の漸近有効性について

北海道大学経済　柿沢　佳秀

1. はじめに

定常時系列で最も基本的な量である自己共分散関数の推定問題は古くから論じられている。特に標本自己共分散の一致性や漸近正規性は線形過程に対して white noise のきわめて緩い仮定の下で研究されている。しかしながら、標本自己共分散の漸近有効性を論じたものは少ない。

今、時系列データを生成する過程が、スペクトル密度 $f_\theta(\lambda)$ をもつスカラー定常過程であるとしよう。k 次の自己共分散は $\int_{-\pi}^{\pi} f_\theta(\lambda)\mathrm{e}^{ik\lambda}\,d\lambda$ のように積分表示されるので、それを有効に推定する問題はスペクトル母数 $\boldsymbol{\theta}$ の推定問題になる。推定量 $\widehat{\boldsymbol{\theta}}_n$ の漸近有効性は様々な意味で研究されており、ここでの興味は「自然な推定量である（不偏な）標本自己共分散がどのようなモデルについて漸近有効になっているのか」という問題である。正規 ARMA(p,q) 過程に対し、Porat (1987) が次のような結果を示した。k 次標本自己共分散は (I) $p \geq q$ ならば、$k = 0, 1, \ldots, p-q$ に対して漸近有効である（その漸近分散がその下限に達する）が、(II) $q > p$ ならば任意の k に対し漸近有効ではない。Porat の証明は ARMA 過程固有のものであって、Kakizawa and Taniguchi (1994) はスペクトル密度を用いて漸近分散及びその下限を表現し、両者が一致するための必要十分条件を与えた。

本講演の前半では、Kakizawa and Taniguchi (1994) の結果を多次元正規定常過程へ拡張する。報告者は最近正規定常過程の 2 次形式の大偏差理論を数学会等で議論しており、後半では標本自己共分散行列の Bahadur 漸近有効性を調べる。

2. 準備

$\{\boldsymbol{X}(t) = (\ X_1(t), \ldots, X_m(t)\ )' : t \in \mathbf{Z}\}$ を平均 $\mathbf{0}$ の m 次元正規定常過程とし、その自己共分散行列関数 $\boldsymbol{R}_\theta(k) = E\boldsymbol{X}(t)\boldsymbol{X}(t+k)'$ $(k \in \mathbf{Z})$ は $\sum_{\ell=-\infty}^{\infty} ||\boldsymbol{R}_\theta(\ell)||_E < \infty$ を満たしているとする。ここで、$\boldsymbol{\theta} = (\ \theta_1, \ldots, \theta_r\ )'$. このとき、$\{\boldsymbol{X}(t)\}$ のスペクトル行列 $\boldsymbol{f}_\theta(\lambda) = \{f_{\theta,ab}(\lambda) : a, b = 1, \ldots, m\}$ が存在し、$\boldsymbol{f}_\theta(\lambda) = (2\pi)^{-1}\sum_{\ell=-\infty}^{\infty} \boldsymbol{R}_\theta(\ell)\mathrm{e}^{-i\ell\lambda}$. n 観測 $\boldsymbol{X}(1), \ldots, \boldsymbol{X}(n)$ に基づいた $r \times r$ Fisher 情報行列 $\boldsymbol{\mathcal{F}}_n(\boldsymbol{\theta})$ に関して

$$\lim_{n\to\infty} \frac{1}{n}\boldsymbol{\mathcal{F}}_n(\boldsymbol{\theta}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \boldsymbol{\Delta}(\lambda)^*[\{\boldsymbol{f}_\theta(\lambda)^{-1}\}' \otimes \boldsymbol{f}_\theta(\lambda)^{-1}]\boldsymbol{\Delta}(\lambda)\,d\lambda \equiv \boldsymbol{\mathcal{F}}(\boldsymbol{\theta})$$

が $\boldsymbol{f}_\theta(\lambda)$ の適当な条件の下で示される。ただし、$\boldsymbol{\Delta}(\lambda) = (\ \mathrm{vec}[\partial_1 \boldsymbol{f}_\theta(\lambda)], \ldots, \mathrm{vec}[\partial_r \boldsymbol{f}_\theta(\lambda)]\ )$ $(m^2 \times r$ 行列). 以下では、$r \times r$ 行列 $\boldsymbol{\mathcal{F}}(\boldsymbol{\theta})$ は任意の $\boldsymbol{\theta}$ について正定値であると仮定する。

3. Kakizawa and Taniguchi (1994) の拡張

$\boldsymbol{R}_\theta(k) = \boldsymbol{R}_\theta(-k)'$ に注意して一般性を失うことなく $k \geq 0$ を仮定する。k 次の（不偏な）標本自己共分散行列 $\widehat{\boldsymbol{R}}(k) = (n-k)^{-1}\sum_{t=1}^{n-k} \boldsymbol{X}(t)\boldsymbol{X}(t+k)'$ $(0 \leq k < n)$ の漸近分散は

$$\lim_{n\to\infty} nVar(\ \mathrm{vec}[\widehat{\boldsymbol{R}}(k)]\ ) = 4\pi\int_{-\pi}^{\pi} \widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)^*[\{\boldsymbol{f}_\theta(\lambda)^{-1}\}' \otimes \boldsymbol{f}_\theta(\lambda)^{-1}]\widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)\,d\lambda. \qquad (1)$$

ここで、$\boldsymbol{\phi}_{ab}^{[k]}(\lambda) = (\mathrm{e}^{ik\lambda}\boldsymbol{e}_b\boldsymbol{e}_a' + \mathrm{e}^{-ik\lambda}\boldsymbol{e}_a\boldsymbol{e}_b')/2$（ただし、$\boldsymbol{e}_a = (\ 0, \ldots, 1, 0, \ldots, 0\ )'$ は $m \times 1$ 単位ベクトル）から、$\boldsymbol{\Phi}^{[0]}(\lambda) = (\ \mathrm{vec}[\boldsymbol{\phi}_{ab}^{[0]}(\lambda)] : a \geq b\ )$ $(m^2 \times m(m+1)/2$ 行列) 及び、$1 \leq k < n$ に対し $\boldsymbol{\Phi}^{[k]}(\lambda) = (\ \mathrm{vec}[\boldsymbol{\phi}_{ab}^{[k]}(\lambda)] : a, b = 1, \ldots, m\ )$ $(m^2 \times m^2$ 行列) を定義し、$\widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda) = [\boldsymbol{f}_\theta(\lambda)' \otimes \boldsymbol{f}_\theta(\lambda)]\boldsymbol{\Phi}^{[k]}(\lambda)$. なお、(1) は $\mathrm{vec}[\widehat{\boldsymbol{R}}(0)]$ を $\mathrm{vech}[\widehat{\boldsymbol{R}}(0)]$ として正しい。一方、$\mathrm{vec}[\boldsymbol{R}_\theta(k)]$ の不偏推定量のクラメル・ラオの下限（$\mathrm{vec}[\boldsymbol{R}_\theta(0)]$ を $\mathrm{vech}[\boldsymbol{R}_\theta(0)]$ と理解）は

$$CRB_n(k) = \left\{\frac{\partial}{\partial\boldsymbol{\theta}'}\mathrm{vec}[\boldsymbol{R}_\theta(k)]\right\}\boldsymbol{\mathcal{F}}_n^{-1}(\boldsymbol{\theta})\left\{\frac{\partial}{\partial\boldsymbol{\theta}'}\mathrm{vec}[\boldsymbol{R}_\theta(k)]\right\}'$$

で与えられ

$$\lim_{n\to\infty} nCRB_n(k) = \left\{\int_{-\pi}^{\pi} \widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)^*[\{\boldsymbol{f}_\theta(\lambda)^{-1}\}' \otimes \boldsymbol{f}_\theta(\lambda)^{-1}]\boldsymbol{\Delta}(\lambda)\,d\lambda\right\}\mathcal{F}(\boldsymbol{\theta})^{-1}$$
$$\left\{\int_{-\pi}^{\pi} \widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)^*[\{\boldsymbol{f}_\theta(\lambda)^{-1}\}' \otimes \boldsymbol{f}_\theta(\lambda)^{-1}]\boldsymbol{\Delta}(\lambda)\,d\lambda\right\}^*. \tag{2}$$

この節では、もし (1)=(2) のとき、$\widehat{\boldsymbol{R}}(k)$ は漸近有効であるという。

Theorem 1. $\widehat{\boldsymbol{R}}(k)$ が漸近有効であるための必要十分条件は、

$$[\boldsymbol{f}_\theta(\lambda)' \otimes \boldsymbol{f}_\theta(\lambda)]\boldsymbol{\Phi}^{[k]}(\lambda) = (\ \mathrm{vec}[\partial_1 \boldsymbol{f}_\theta(\lambda)], \ldots, \mathrm{vec}[\partial_r \boldsymbol{f}_\theta(\lambda)]\ )\boldsymbol{C} \tag{3}$$

(同値な条件として) あるいは

$$\boldsymbol{\Phi}^{[k]}(\lambda) = -(\ \mathrm{vec}[\partial_1 \boldsymbol{f}_\theta(\lambda)^{-1}], \ldots, \mathrm{vec}[\partial_r \boldsymbol{f}_\theta(\lambda)^{-1}]\ )\boldsymbol{C} \tag{4}$$

となるような λ に無関係な行列 $\boldsymbol{C}$ が存在することである。

VMA(q) 過程では (3) が成立することはあり得ず、VAR(p) 過程では $0 \le k \le p$ のときに限り (4) が成立する。これらはスカラーの Kakizawa and Taniguchi に対応するが、行列は一般に可換でなく残念ながら VARMA(p,q) 過程では $q > p$ のとき (4) が成立しないこと以外は線形従属性をうまく扱えない。

4. 標本自己共分散行列の Bahadur 漸近有効性

任意の零でない $m \times m$ 定数行列 $\boldsymbol{\Gamma} = (\gamma_{ab})$ (なお、$k=0$ のときは下三角行列とする) をとり、

$$g(\boldsymbol{\theta}) = \mathrm{tr}[\boldsymbol{\Gamma}'\boldsymbol{R}_\theta(k)] = \int_{-\pi}^{\pi} \mathrm{tr}\{\boldsymbol{\phi}^{[k]}(\lambda)\boldsymbol{f}_\theta(\lambda)\}\,d\lambda \tag{5}$$

とおく。ただし、$\boldsymbol{\phi}^{[k]}(\lambda) = \sum_{a,b}^{m} \gamma_{ab}\boldsymbol{\phi}_{ab}^{[k]}(\lambda)$. $g(\boldsymbol{\theta})$ の一致推定量の exponential rate に関して次の Theorem 2 が IID 設定を仮定した Bahadur (1960) の拡張として証明される。

Theorem 2. S_n が $g(\boldsymbol{\theta})$ の一致推定量ならば、任意の $\boldsymbol{\theta} \in \Theta$ について

$$\limsup_{\varepsilon\to 0}\limsup_{n\to\infty}\left[-\frac{1}{n\varepsilon^2}\log P_\theta(|S_n - g(\boldsymbol{\theta})| \ge \varepsilon)\right] \le \frac{1}{2v(\boldsymbol{\theta})}.$$

ここで、$v(\boldsymbol{\theta})$ は (2) の $\widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)$ を $[\boldsymbol{f}_\theta(\lambda)' \otimes \boldsymbol{f}_\theta(\lambda)]\mathrm{vec}[\boldsymbol{\phi}^{[k]}(\lambda)]$ として得られる。

(5) の一致推定量として $S_n = \mathrm{tr}[\boldsymbol{\Gamma}'\widehat{\boldsymbol{R}}(k)]$ を考え、この exponential rate を評価すると

Theorem 3.

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\left[-\frac{1}{n\varepsilon^2}\log P_\theta(|S_n - g(\boldsymbol{\theta})| \ge \varepsilon)\right] = \frac{1}{2w(\boldsymbol{\theta})}.$$

ここで、$w(\boldsymbol{\theta})$ は (1) の $\widetilde{\boldsymbol{\Phi}}^{[k]}(\lambda)$ を $[\boldsymbol{f}_\theta(\lambda)' \otimes \boldsymbol{f}_\theta(\lambda)]\mathrm{vec}[\boldsymbol{\phi}^{[k]}(\lambda)]$ として得られる。

Theorem 2 より、$w(\boldsymbol{\theta}) \ge v(\boldsymbol{\theta})$ が成立しており、任意の $m \times m$ 行列 $\boldsymbol{\Gamma}$ について $w(\boldsymbol{\theta}) = v(\boldsymbol{\theta})$ であるとき、$\widehat{\boldsymbol{R}}(k)$ は Bahadur の意味で漸近有効であるということにする。しかし、$w(\boldsymbol{\theta}), v(\boldsymbol{\theta})$ の定義から、Theorem 1 と同じ必要十分条件が Bahadur の意味でも得られることが示された。なお、不偏推定量 $\widehat{\boldsymbol{R}}(k)$ の代わりに $\widehat{\widehat{\boldsymbol{R}}}(k) = n^{-1}\sum_{t=1}^{n-k}\boldsymbol{X}(t)\boldsymbol{X}(t+k)'$ としても Theorem 3 が正しい。

A model selection criterion
for accurate forecasting of nonlinear time series

神奈川大・工　松田 安昌

1 Introduction

Chen and Tsay (1993) proposed functional-coefficient autoregressive (FAR) models

$$X_t = f_1(X_{t-d})X_{t-1} + \cdots + f_p(X_{t-d})X_{t-p} + Z_t, \tag{1}$$

which contain many useful nonlinear time series models as special cases. In fact, by identifying $\{fi(\cdot)\}$ using adequate parameters $\theta \in \Theta$, FAR models are reduced to the following models.

- If $f_i(x) = \theta_i$, linear autoregressive models.
- If $f_i(x) = \begin{cases} \theta_i & \text{if } x < r \\ \phi_i & \text{if } x \geq r \end{cases}$, threshold autoregressive (TAR) models.
- If $f_i(x) = \theta_i + \phi_i \exp(-rx^2)$, exponential autoregressive models.

In this report, we propose a method useful for identifying a delay parameter d and a functional form of $f(\cdot)$.

2 A selection Criterion

Suppose we have observed $X_1, \ldots, X_n$ from FAR model (1) when d and p are knwon. When we try to fit a parametric model $g_i(\cdot, \theta)$, $\theta \in \Theta$ to unknown function $f_i(\cdot)$, the goodness-of-fit of the models is evaluated as fowllows:

The following two kinds of estimators for $f_i(\cdot)$.

1. parametric estimator $g(\cdot, \hat{\theta}) = (g_1(\cdot, \hat{\theta}), \ldots, g_p(\cdot, \hat{\theta}))'$, where $\hat{\theta}$ is given by minimizing

$$Q(\theta) = \sum_{t=p+1}^{n} (X_t - g_1(X_{t-d}, \theta)X_{t-1} - \cdots - g_p(X_{t-d}, \theta)X_{t-p})^2.$$

2. nonparametric estimator $\hat{f}_{h_n}(\xi) = (\hat{b}_1, \ldots, \hat{b}_p)'$, which is given by minimizing

$$Q(b) = \sum_{t=p+1}^{p} (X_t - b_1 X_{t-1} - \ldots - b_p X_{t-p})^2 K_{h_n}(\xi - X_{t-d})$$

and its solution is

$$\hat{f}_{h_n}(\xi) = \left(\frac{1}{n} \sum_{t=p+1}^{n} Y_t Y_t' K_{h_n}(\xi - X_{t-d}) \right)^{-1} \frac{1}{n} \sum_{t=p+1}^{n} Y_t X_t K_{h_n}(\xi - X_{t-d}),$$

where $K_{h_n}(\cdot) = h_n^{-1} K(\cdot / h_n)$ and $Y_t = (X_{t-1}, \ldots, X_{t-p})'$ and h is bandwidth.

The difference of the two estimators is sample fluctuation for a correct model, while it is large for incorrect model. Here, we define a metric between the two estimators and use it for testing a goodness-of-fit of the model $g(\cdot, \theta)$.

$$T_n = \sum_{j=1}^{m} nh(\hat{f}_h(\xi_j) - g(\xi_j, \hat{\theta}))' \hat{\Omega}^{-1}(\xi_j)(\hat{f}_h(\xi_j) - g(\xi_j, \hat{\theta})), \tag{2}$$

where

$$\begin{aligned}
\hat{\Omega}(\xi) &= \hat{\sigma}^2 \hat{A}_1^{-1}(\xi) \int K(z)^2 dz, \\
\hat{\sigma}^2 &= \frac{1}{n} \sum_{t=p+1}^{n} \left(X_t - \sum_{i=1}^{p} g_i(X_{t-d}, \hat{\theta}) X_{t-i} \right)^2, \\
\hat{A}_1(\xi) &= \frac{1}{n} \sum_{t=p+1}^{n} Y_t Y_t' K_h(\xi - X_{t-d}).
\end{aligned}$$

3 Asymptotic Properties

When the correct model is fitted, ie.

$$f(\cdot) = g(\cdot, \theta)$$

for some $\theta \in \Theta$, asymptotic properties of our criterion are shown.

We have the following result.

Theorem .1 *Under some conditions, for m distinct points $x_1, \ldots, x_m$, $\sqrt{nh}(\hat{f}_h(x_j) - f(x_j, \hat{\theta}))$ for $j = 1, \ldots, m$ converges in distribution to a mutually independent p-dimensional normal variable with mean 0 and variance*

$$\Omega(x_j) = \sigma^2 \Big(E(Y_t Y_t' | X_{t-d} = x_j) w_1(x_j) \Big)^{-1} \int K(z)^2 dz,$$

Hence, T_n converges in distribution to χ^2_{mp}, when the model $g(\cdot, \theta)$ is correctly specified.

References

[1] Chen, R. and Tsay, R. (1993), Functional-coefficient autoregressive models. *Journal of the American Statistical Association*, **88**, 298-308.

[2] Härdle, W. and Mammen, E. (1993), Comparing nonparametric versus parametric regression fits. *Annals of Statistics*,**21**(4), 1926-47.

[3] Matsuda, Y. (1998), A diagnostic statistic for functional-coefficient autoregressive models. *Communications in Statistic-Theory and Methods*,**27**(9), 2257-2273.

[4] Tong, H. (1990), *Non-linear Time Series: A Dynamical System Approach.* Oxford University Press.

Calculations of level probabilities for normal random variables with unequal variances with applications to Bartholomew's test in unbalanced one-way models

Tetsuhisa Miwa
Laboratory of Statistics
National Institute of Agro-Environmental Sciences
3-1-1 Kannondai, Tsukuba 305-8604, Japan

A. J. Hayter
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205, U.S.A.

Wei Liu
Department of Mathematics
University of Southampton
Southampton, S09 5NH, England

An effective procedure is presented to compute the level probabilitites under simple order of independent normal random variables with unequal variances. Bartholomew proposed the likelihood ratio test for testing the homogeneity of the treatment effects against the simply ordered alternative hypothesis. Although there is some literature showing that Bartholomew's test has good properties, its null distribution for the unbalanced model has been difficult to calculate except for small k. The calculation of the level probabilities allows the computation of the p-values of Bartholomew's test for unbalanced models.

Consider the unbalanced one-way analysis of variance model

$$X_{ij} = \mu_i + \epsilon_{ij}, \quad 1 \le i \le k, \quad 1 \le j \le n_i,$$

where the ϵ_{ij} are independent $N(0, \sigma^2)$ random variables. Let $\bar{X}_i$, $1 \le i \le k$, be the ith sample mean based on n_i observations, and let S^2 be an unbiased estimate of σ^2 distributed independently of the $\bar{X}_i$ as $S^2 \sim \sigma^2 \chi^2_\nu / \nu$ for some degrees of freedom ν. Usually the mean squared error in the analysis of variance will be used as the estimate S^2 with $\nu = \sum_{i=1}^k n_i - k$.

Suppose that we are interested in testing the null hypothesis

$$H_0 \colon \mu_1 = \mu_2 = \cdots = \mu_k$$

against the simply ordered alternative hypothesis

$$H_A \colon \mu_1 \le \mu_2 \le \cdots \le \mu_k$$

with at least one strict inequality. Bartholomew derived the likelihood ratio test for this problem. With known σ^2, Bartholomew's test statistic is

$$\bar{\chi}^2 = \sum_{i=1}^k n_i (\hat{\mu}_i - \bar{X})^2 / \sigma^2,$$

where $\bar{X} = \sum_{i=1}^k n_i \bar{X}_i / \sum_{i=1}^k n_i$ is the overall mean of all the data observations, and $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k$ are the maximum likelihood estimators of $\mu_1, \mu_2, \ldots, \mu_k$ under the order restriction $\mu_1 \le \mu_2 \le$

$\ldots \leq \mu_k$. These isotonic estimators can be obtained from an algorithm whereby any adjacent means which violate the ordering are pooled.

However, one important obstacle to the implementation of Bartholomew's test has been that tables of critical points are available only for equal sample sizes. In particular, no explicit expression has been available to calculate the critical points for unequal sample sizes with $k > 5$. Bartholomew showed that the null distribution of the statistic $\bar{\chi}^2$ can be expressed as

$$\Pr\{\bar{\chi}^2 \leq a\} = P(1,k;n) + \sum_{l=2}^{k} P(l,k;n) \Pr\{\chi^2_{l-1} \leq a\},$$

for $a > 0$ where χ^2_i is a χ^2 random variable with i degrees of freedom, and $P(l,k;n)$ is the probability that $\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_k$ consist of l different values. These probabilities $P(l,k;n)$ are called "level probabilities" and depend upon the weight vector $n = (n_1, n_2, \ldots, n_k)$. The problem in the evaluation of this null distribution for unequal sample sizes has been the difficulty in calculating these level probabilities for unequal weights.

The purpose of this talk is to present a method which enables the easy and quick evaluation of the level probabilities with unequal weights. This then allows the easy computation of the p-values of Bartholomew's test for the unbalanced one-way model with an arbitrary number of treatments k. A crucial step in calculating the level probabilities is the calculation of orthant probabilities of the form $P(k,k;n) = \Pr\{\bar{X}_1 < \bar{X}_2 < \cdots < \bar{X}_k\}$, and a recursive method proposed in Hayter and Liu is employed to calculate these probabilities.

The arrangement of the talk is as follows. The computation of the orthant probabilities is discussed first and then the computation of all of the level probabilities is described. The applications to Bartholomew's test are then given together with an example.

In this talk it is shown how the level probabilities under simple ordering can be calculated for independent normal random variables with unequal variances. This enables us to implement Bartholomew's tests for unbalanced one-way layouts. The calculation of the orthant probabilities $\Pr\{\bar{X}_1 < \bar{X}_2 < \cdots < \bar{X}_k\}$ plays the main role in the computation of the level probabilities $P(l,k;n)$, and we have demonstrated a simple technique to calculate these recursively implementing cubic polynomial approximation and an optimum choice of grid points. Our method provides a fast and accurate computation of the null distribution of Bartholomew's statistic which, for example, takes about fourteen seconds to compute for the comparison of $k = 10$ treatments.

When the variance σ^2 is unknown, the likelihood ratio test statistic $\bar{E}^2$ has been suggested to test the homogeneity of treatment effects against the alternative hypothesis of simple order. However, in section 4 we examined another statistic which has also been suggested in the literature, namely $\bar{B}^2$ which is a simple extension of the $\bar{\chi}^2$ statistic with the unknown σ^2 replaced by its estimate S^2. As we demonstrate, the statistic $\bar{B}^2$ has the advantage of providing simultaneous confidence intervals for all monotone contrasts of the treatment effects.

Finally, it should be remembered that although Bartholomew's test was constructed as the likelihood ratio test for the one-way layout, it can also be applied to many other experimental designs. Notice that the only requirement is that the statistic $\bar{B}^2$ is calculated from a set of independent normal sample means and an independent estimate S^2.

Misspecified Prediction in Time Series Analysis

In bong CHOI and Masanobu TANIGUCHI

はじめに

時系列解析の予測問題において spectral density に misspecification のある予測問題は Grenander-Rosenblatt(1957) によって紹介された。その後、AR model の h-step prediction という形で Yamamoto(1976), Baillie(1975) 等で研究された。実は h-step prediction は spectrum に misspecification がある場合の予測ととらえることができる。本報告ではまず spectral density に misspecification のある場合の予測をするとき、その予測誤差に対する漸近期待値を導く。そして、時系列回帰モデルにおける残差項の spectral density に misspecification のある場合の予測の平均二乗誤差を研究し、その予測誤差に対する漸近期待値を導く。これらの結果は今までの種々の h-step 予測等の結果を特別な場合として含む。最後に、misspecification による予測の悪さがどのような least favourable な spectrum の contamination でおきるか研究する。そして、いくつかの例として具体的な値を与えて misspecification による予測の悪さを実際に確かめてみる。

参考文献

[1] Baillie, R. T.(1979) The asymptotic mean square error of multistep prediction from the regression model with autoregressive errors. *J. Amer. Statist. Assoc. 74. 175-184*

[2] Brillinger, D. R.(1975) *Time Series ;Data Analysis and Theory.* Holt, Rinehart & Winston.

[3] Grenander-Rosenblatt(1957) *Statistical Analysis of Stationary Time Series* . Wiley, New York.

[4] Hosoya, T. and Taniguchi, M.(1982) A central limit theorem for stationary processes and the parameter estimation of linear process. *Ann. Statist. 10. 132-153.* Correction: (1993) 21,1115-1117.

[5] Rozanov, Yu. A.(1969) On a new class of statistical estimates. *Soviet-Japanese Sympos. Theory of Probability, Novosibirsk. 239-252.*

[6] Taniguchi, M.(1980) On selection of the order of the spectral density model for a stationary process. *Ann. Inst. Statist. Math. 32. 401-419.*

[7] Taniguchi, M.(1980) Regression and interpolation for time series. *Recent Development of Statistical Inference and Data Analysis. K.Matusita, Ed. 311-321. North-Holland.*

[8] Yamamoto, T.(1976) Asymptotic mean square prediction error for an autoregressive model with estimated coefficents. *Appl. Statist. 25. 123-127.*

縮小ランク成長曲線モデルにおける次元の推定

広島大・理 藤越康祝

要約

正準変数分析法, あるいは正準判別分析法は, いくつかのグループからの多変量観測値が与えられているとき, これらのグループ間の差異をできるだけ小数次元の空間で記述するための方法である. とくに, 各グループの共分散行列が同一で, かつ, 平均ベクトルに構造が仮定されていない場合に対して, その方法および関連した推測問題が古くから研究されており, その起源は Fisher (1936), Bartlett (1938) に遡る. これらの研究およびその後の発展等については, 例えば, Reinsel and Velu (1998) による研究書等を参照されたい. 一般に, 各グループの平均ベクトルが平面上にあると, グループ間の差異はその平面の次元に等しい数の判別変数で記述される. 正準変数解析の次元とは, このような平面の最小次元数のことである. Fujikoshi and Veitch (1979) は正準変数分析の次元を推定するための AIC 型基準および C_p 型基準を提案している.

多変量観測値が経時データ, あるいは成長曲線データである場合について, 上で述べた正準変数分析をそのまま適用することは適切でなく, 適当な修正が必要である. とくに, Potthoff and Roy (1964) によって導入された成長曲線モデルの場合を扱うが, これについては Albert and Kshirsagar (1993) によっても扱われている. 彼等が扱っている推定・検定問題は, より一般な定式化のもとでFujikoshi (1974) においても考察されていることを注意したい. 一般に, 正準変数分析のモデルは, 成長曲線モデルにおいてある種のパラメータ行列のランクが縮小しているものとして表され, この意味で縮小ランク成長曲線モデルと呼ばれる. 縮小ランク成長曲線モデルは, Reinsel and Velu (1998) の第6章においても解説されている.

本報告では, 縮小ランク成長曲線モデルにおける推測問題を概観し, 次元の推定法, グループ間の差異や類似性を視学的に捉えるための方法を提案した。この問題は次のように要約される。変数 y が各個体について, p 個の時点または処理条件のもとで観測されるものとする. また n 個の個体はk 個のグループに分類され, 第 i グループの第 j 観測値を

$$\boldsymbol{y}_{ij} = (y_{ij1}, \cdots, y_{ijp})', \quad j = 1, \ldots, n_i, \quad i = 1, \ldots, k$$

とする. Potthoff and Roy (1964) によって導入された成長曲線モデルは,

$$\boldsymbol{y}_{i1}, \cdots, \boldsymbol{y}_{in_i} \quad \sim \quad i.i.d. \quad \mathrm{N}_p(X\boldsymbol{\xi}_i, \Sigma)$$

と表される. ただし, X は $p\times q$ の個体内計画行列, $\Xi=(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_k)'$ は $k\times p$ の未知パラメーター行列, Σ は未知正定値行列. グループ間の比較を目的とする縮小ランク成長曲線モデルは, パラメーター行列 Ξ が

$$\mathrm{rank}(C\Xi)=s\,(\le m=\min\{p,k-1\})$$

をみたすモデルである. ここに,C は $(k-1)\times k$のコンストラスト行列である。この制約条件は, k 個の成長曲線係数ベクトル$\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_k$ が s 次元平面上にあることを意味し,

$$C\Xi=\Theta\Gamma',\quad \Theta=[\boldsymbol{\theta}_1,\cdots,\boldsymbol{\theta}_s],\quad \Gamma=[\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_s]$$

と表される. このとき, Θ,Γ への分解は一意的でないが, Γ の列ベクトルは縮小次元空間表現において重要な役割を演じている. 本報告では、次元 s を推定するためのモデル選択基準, および、各個体を2または3次元空間に表現するための正準変数分析法を提案した。

参考文献

[1] Bartlett, M.S. (1938). Further aspects of the theory of multiple regression. *Proc. Camb. Phil. Soc.* **34**, 33-40.

[2] Fujikoshi, Y. (1974). The likelihood ratio tests for the dimensionality of regression coefficients. *J. Multivariate Anal.* **4**, 327-340.

[3] Fujikoshi, Y. (1977). Asymptotic expansions for the distribution of some multivariate tests. *Multivariate Analysis IV* (P.R. Krishnaiah, Ed.), 55-70, North-Holland, Amsterdam.

[4] Fujikoshi, Y. and Veitch, L.G. (1979). Estimation of dimensionality in canonical correlation analysis. *Biometrika* **66**, 345-351.

[5] Fisher, R.A. (1936). The statistical utilization of multiple measurements. *Ann. Eugenics* **8**, 376-386.

[6] Potthoff, R.F. and Roy, S.N. (1964) A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* **51**, 313-326.

[7] Reinsel, G.C. and Velu, R.P. (1998). Reduced-Rank Models for Regression. *Lecture Notes in Statistics* No.**137**, Springer-Verlag.

３相主成分分析による時系列データの解析 ― 問題点と数値例 ―

村上 隆（名古屋大学・教育学部）

1. 序

同一の質問紙（questionnaire）を同一の個体群に複数回反復実施することによって得られた縦断的データ（longitudinal data）について考える。この種のデータは心理学を始めとする社会科学的研究においてしばしば得られる。

各時点のデータ自体がデータ行列の形をとるから，データ全体は3つの添え字をもつ3相データとなる。すなわち，x_{ijk} $(i=1,\ldots,N, j=1,\ldots,p, k=1,\ldots,m)$ であり，ここで，N は個体の数，p は質問項目 (items) の数，m は反復の数である。反復の回数は少ないが，各時点（occasion）における個体 (individuals) 数は多くとれる。

こうしたデータの分析にあたっては，次のことが前提条件である。

個々の変量，すなわち，質問項目への反応の意味は，必ずしも明確でない。むしろ，それらの合成変量としての尺度，あるいはそれらの基底にある因子（潜在変量）の方が，しばしば明確な意味をもつ。これらは，個人の特性（traits），あるいは理論的な構成概念（constructs）と呼ばれる。そして，分析の主要な目的は平均値ではなく，特性の時系列的安定性の評価，つまり機会間相関の正確な評価にある。

特性をあらわす尺度は相互に相関する複数の項目によって作られる必要があるが，そのことは，意味の一様性（内部一貫性 internal consistency）だけでなく，意味の多様性（帯域幅 bandwidth）にも配慮する必要がある。つまり，尺度を定義する際には，時点間相関だけでなく，時点内の項目内相関にも配慮する必要がある。その結果，合成変量の定義には，単なる回転以上の任意性が存在する。その際，個別のPCAも拡張されたCAも理想の方法とは言えない。

2. Tucker2 とその問題点

3相主成分分析のサブモデルである Tucker2 model は，$Z_k(k=1,\ldots,m)$ を $n\times p$ のデータ行列，G を $n\times r$ の主成分得点行列，A^* を $p\times q$ の1次負荷行列，$C_k(k=1,\ldots,m)$ を $q_k\times r$ の2次負荷行列とするとき，

$$f(F,A^*,C_1,\ldots,C_m)=\sum_{k=1}^{m}\parallel Z_k-GC_k'A^{*\prime}\parallel^2 \tag{1}$$

を，$n^{-1}G'G=I$，$m^{-1}\sum_{k=1}^{m}C_kC_k'=I$ なる制約条件の下で最小化することをめざすものである (Kroonenberg & De Leeuw, 1980; 村上，1990)。ただし，$q\le p, r\le mp, mr\le q, r\le mq$ とする。なお，データ行列は，各機会ごとに中心化され，全機会を通じて標準化されているものとする。すなわち，$Z_k'\mathbf{1}=\mathbf{0}$ $(k=1,\ldots,m)$，$(nm)^{-1}\mathrm{diag}\sum_{k=1}^{m}Z_k'Z_k=I$ である。

これを、心理学的縦断データに適用する際には，次のような問題が生ずる。

1. モデル(1) には，各occasionに直接対応する主成分のmatrixが与えられていない。それに最も近いのは，$\ddot{F}_k=GC_k'$ であろうが，その時点間共分散行列は，$n^{-1}\ddot{F}_k'\ddot{F}_l=C_kC_l'$によって与えられる。しかしながら，$[\ddot{F}_1\ldots\ddot{F}_m]$ のrankは，r を超えられないから，多くの場合，C_kC_l' によって得られる機会間相関は過大評価となる。もっと直接的な時点間相関の指標が欲しい。

2. もう１つ，各主成分の機会間相関行列は，次のような形のシンプレックス行列，

$$R=\begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{m-1}\\ \rho & 1 & \rho & \ldots & \rho^{m-2}\\ \vdots & \vdots & \vdots & \ldots & \vdots\\ \vdots & \vdots & \vdots & \ldots & \vdots\\ \rho^{m-1} & \rho^{m-2} & \rho^{m-3} & \ldots & 1\end{bmatrix} \tag{2}$$

に準じたquasi simplex行列の形をとる可能性が高い。しかし，simplex行列は m より小さい次元に分解できない。すなわち，(1) において，$r << mq$ として2次主成分分析を行うことが，機会間相関を歪める可能性もあるように思われる。

3. Stationary weightsをもつ1次主成分

第1の問題を解決するために，機会を通じて共通の（stationary）重みを用いる合成変量を定義する方法を考える。すなわち，

$$F_k = Z_k V \quad k = 1, \dots, m \tag{3}$$

における重み行列 V を求める。それには，

$$g(V; W_1, \dots, W_m) = \mathrm{tr} \sum_{k=1}^{m} \sum_{l=1}^{m} W_k'(V' S_{kl} V) W_l \tag{4}$$

を，$V'V = I$，$\sum_{k=1}^{m} W_k' W_k = I$ の下で最大化すればよい。

この基準は，Compositesの分散と，その機会間相関の両方を大きくすることを目指す，compromisingなものであり，Psychometricな尺度構成に適するように思われる。またこれは，アルゴリズム的にはTuckals2と等価であることが示される（Murakami, 1998)。V，W_kには，単純構造を目指す直交回転を行う。

また，第2の問題については，r を変えて計算を試みることによって，直接検討することができる。

4. 数値例と結果

雇用促進事業団職業研究所による，39項目からなる職業レディネステスト（vocational interests inventory）を，367名の中学生に，1年次，2年次，3年次においてそれぞれ実施したものを適用例とした。主な結果は以下のようである。

1. 縦断的データのTucker2 的な扱いは，Psychometricな尺度構成の目的を反映しており，意味のある結果を導き出しているように思われる。

2. しかしながら，Tuckals2は，questionnaireの反復実施による縦断的データの分析において，単なる同時的PCAと著しく異なる1次合成変量を産み出すわけではなかった。ただし，p に比して m が大きい場合については，実質的に異なる解が得られる可能性がある。実際，(研究会以後の検討で) そのような例も見出された。

3. 各尺度の時点間相関は，おおよそsimplexの 形態を示す。Tuckals2のアルゴリズムが，特に時点間構造に歪みをもたらす兆候は認められない。

4. 重みの直交回転は，解釈可能な解を産み出すのに寄与する。

5. もとのモデルの1次負荷行列 $\tilde{A}^*$ は，斜交構造行列 $\tilde{L}$とほとんど同じ解釈を示すが，2次負荷行列 C_k は，2次重み行列 W_k とは，かなり異なった解釈につながった。これは，このデータの場合，回転前の第II合成変量の「位置づけ」にかかわるものと思われる。

文　献

Kroonenberg, P.M. and De Leeuw, J. (1980): Principal component analysis of three-mode data by means of alternating least squares algorithm. *Psychometrika*, **45**, 69-97.

村上　隆　(1990)　3相データの階層的主成分分析 柳井晴夫・岩坪秀一・石塚智一（編）『人間行動の計量分析－多変量データ解析の理論と応用』　東京大学出版会 71-94.

Murakami, T. 1998 Tucker2 as a second-order principal component analysis. C. Hayashi et al. (Eds.) *Classification, Data Analysis and Related Methods.* Springer-Verlag. Pp.575-586.

Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, **31**, 279-311.

An Overview of Three-way Component Analysis Models and Some Recent Developments for Analysis of Longitudinal Data

Henk A.L. Kiers[1]
University of Groningen

Keywords: three-way data, three-way models, multi-way data, multivariate longitudinal data.

Three-way data are data associated with three sets of units. As an examples of three-way data one may, in behavioral research, have scores of a set of individuals on a set of variables at different situations or time points. In this example the three sets of units defining the three "ways" of the data are different. However, one also commonly encounters three-way data where two sets of units are equal, for instance in the case of data consisting of the similarities between a number of stimuli as judged by a number of judges, or, data consisting of several covariance matrices for a set of variables. Such data are denoted as *two-mode three-way data*. Finally, one may have data where all ways pertain to the same mode. Such data are denoted as *one-mode three-way* data. An example is data on transitions from party to party over three consecutive elections. The difference in data type leads to a difference in techniques for analyzing three-way data.

In this presentation, an overview is given of various three-way techniques for different data types. Most attention is given to the two most popular three-way techniques for the analysis of three-mode three-way data: PARAFAC and three-mode principal component analysis (3MPCA). It is demonstrated that both techniques give components for all three modes, while in 3MPCA there is an additional outcome matrix, the core, which relates the components of all modes to each other. A further difference between 3MPCA and PARAFAC is that the latter gives unique solutions, whereas in the former, as in ordinary PCA, the solution is determined only up to rotation of the component matrices and the core. Next some attention is paid to techniques for the analysis of two-mode three-way data:

INDSCAL, PARAFAC2, three-mode scaling, IDIOSCAL, and a very general multi-mode structural equations model.

The second part of the presentation focuses on application of three-way methods to multivariate longitudinal data. When longitudinal data for a reasonably large set of subjects on a number of variables are available, three-mode three-way methods can be used upon considering the subjects, the variables and the time points as the three-modes of the data. A technique for displaying the results from this special application is presented. However, naive application of three-way methods to such three-way methods implies that the natural ordering of time points is not taken into account. In particular in cases of relatively long time series, it is attractive to explicitly take this ordering into account, for instance by using smoothness constraints on the component matrix, as will be illustrated.

In many cases one has two-mode three-way data. One may for instance have longitudinal covariance matrices, where the raw data have been aggregated over the subjects. Such data can nicely be analyzed by models like INDSCAL, IDIOSCAL, three-mode scaling and the like. Such data sets usually pertain to few time points. Therefore, one may fruitfully employ models that do not take the ordering of the time points into account. Instead, by analyzing such data, and keeping certain parts of the solution invariant over time, one can inspect how the other outcomes (e.g., correlations between components) change over time. Alternatively, one may have long time series for few individuals, where, furthermore, the time series for different subjects may not be comparable. In such cases, one may fruitfully analyze the data after first aggregating over time points. In some models one can explicitly take the time ordering into account by modeling the relation of the variables at time t with components at time t *and time t-1*.

Finally, an example will be given of the analysis of one-mode three-way transition data by means of a multidimensional scaling model. In this model, three-way distances (distances between *three* entities are used, as well as a *slide* vector that indicates the main trend for each transition.

[1] Author Address: Heymans Institute (PA), University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands. email: h.a.l.kiers@ppsw.rug.nl

経時測定データに基づく処理比較
—血圧および心拍数２４時間値データへの応用—

東大工学系研究科
広津千尋・安達絵里

1. はじめに

各被験者につき血圧および心拍数を３０分おきに２４時間測定したデータがある。本データの特徴は２４時間後にほぼ元の水準に戻り、線形なトレンドには興味のないことである。この種のデータでは臨床的に、昼間平均、睡眠時平均、昼間上昇および夜間下降時の水準、両者の差で定義される振幅等に興味が持たれ解析されている。とくに夜間平坦および上昇は病態であり、ごく近年は夜間過度の下降も問題視され始めている。さらに双生児解析により、これらプロファイルパラメータのうち遺伝素因に依るもの、依らないものの分類が興味を集めている。しかしながら、これら個々のパラメータは相互に関連があり、別個の検定にはよく知られた多重性の問題も伴う。そこでこれら興味あるパラメータを総括したものとして推移曲線の凹凸性に注目し、被験者の分類および特徴付けを行う方法を提案する。

2. 統計モデル

$$y_{ij} \;=\; \mu_{ij}+\varepsilon_{ij},\;\; i=1,\cdots,n;\; j=1,\cdots,t$$

$$\mu_{ij}: i\text{th subject の}j\text{期における反応};\;\; \varepsilon_{ij}: \text{測定誤差 n.i.i.d.}(0,\sigma^2)$$

ここでμ_{ij}について時刻に関する特別なパラメトリックモデルは仮定しないが、ゆっくりとした系統的な変化を想定する。一方、ε_{ij}はランダムな偶然誤差であり、特に相関構造は仮定しない。この仮定は例えば体内コレステロール量を一ヶ月おきに６期測定した経時データでは問題なく受け容れられようが、分単位の血圧データでは難しいかもしれない。今回の時間単位データは微妙であるが、接近したデータ間の相関はもし存在したとしてもμ_{ij}のゆっくりした変化の方に吸収させるという考え方である。以下ではμ_{ij}に母数模型を仮定するが、変量模型を仮定してもほぼ同じように議論を展開することができる。さて、このように問題を整理すると、結局、繰返しのない２元配置データ$(n\times t)$において、ある系統的な交互作用を検出する問題として定式化することができる。そこで、興味ある傾向性仮説として

$$\text{帰無仮説}\,H_0 \;:\; \mu_{ij}-\mu_{ij+1}-(\mu_{i'j}-\mu_{i'j+1})=\mu_{ij+1}-\mu_{ij+2}-(\mu_{i'j+1}-\mu_{i'j+2}),$$

$$\text{対立仮説}\,H_1 \;:\; \mu_{ij}-\mu_{ij+1}-(\mu_{i'j}-\mu_{i'j+1})\geq\mu_{ij+1}-\mu_{ij+2}-(\mu_{i'j+1}-\mu_{i'j+2})$$

を想定する。H_0はスロープ変化が被験者iとi'で一致すること、H_1は凸（下に）性の程度がiに対しi'の方が強いことを意味する。ただしH_1は、iとi'には順序関係が無く交換可能なため、本質的に両側仮説である。以下でH_1に対して適切な検定統計量を導き、それに基づいて似通った推移パターンを示す被験者を分類する方法を提案する。これは母数模型の場合、凸性の程度が一定の被験者をグルーピングすることに相当し、変量模型の場合には$\boldsymbol{\mu}_i=(\mu_{i1},\cdots,\mu_{it})'$ の分散行列Ωが$\Omega\propto 1_t1_t'$ という特別な構造となるサブグループを構成することに相当する。

3. 系統的成分の推定

3.1 被験者間の二乗距離

対立仮説H_1は２階差分行列L_t'を用いて次のように表すことができる、

$$H_1: L_t'(\boldsymbol{\mu}_{i'} - \boldsymbol{\mu}_i \geq \mathbf{0}, \quad L_t' = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \multicolumn{8}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix}_{t-2\times t} .$$

このとき検定の完全類は$(L_t'L_t)^{-1}L_t'(\boldsymbol{y}_{i'} - \boldsymbol{y}_i)$, $\boldsymbol{y}_k = (y_{k1}, \cdots, y_{kt})'$の各要素につき単調増大、かつ凸な受容域を持つ検定全体で与えられる。そこで規準化定数η_k, δ_k（略）により、$Q_t' = \mathrm{diag}(\eta_k^{1/2}\delta_k)(L_t'L_t)^{-1}L_t$ を導入し、検定統計量$\chi^{\dagger 2}(i;i') = \| Q_t'(\boldsymbol{y}_{i'} - \boldsymbol{y}_i \|^2 /2$ を構成する。これを被験者i、i'の間の二乗距離と呼ぶ。$\chi^{\dagger 2}$はH_0の下で

$$\chi^{\dagger 2}(i;i') = \sigma^2\{\omega_2\chi^2_{(2)} + \omega_3\chi^2_{(3)} + \cdots + \omega_{t-1}\chi^2_{(t-1)}\} \tag{1}$$

と展開される。ただし、$\chi^2_{(k)}$はk次多項式に沿ったトレンドを検出するための互いに独立な自由度１のχ^2統計量である。(1) 式は$\chi^{\dagger 2}$の分布論と同時に明解な特徴付けを与える。すなわち$\chi^2_{(k)}$の荷重ω_kが急速に減衰することから、$\chi^{\dagger 2}$は２次式的変化に 重みがあり、それに高次成分を加味した統計量である。なお、１次成分$\chi^2_{(1)}$が消去されていることに注意する。

3.2 群間の一般化二乗距離

群$G_1, \cdots, G_m$の間の一般化二乗距離は次のように定義する。まず、群$G_1, \cdots, G_k$の間の規準化直交対比を$\bar{\boldsymbol{y}}(G_1; \cdots; G_m) = \sum_{i=1}^n \gamma_i \boldsymbol{y}_i$, $\gamma_i = \lambda_l$ $(i \in G_l)$; $\sum q_l\lambda_l = 0$; $\sum q_l\lambda_l^2 = 1$によって定義する。このとき、$\chi^{\dagger 2}(G_1; \cdots; G_m) = \max_{\lambda_1, \cdots, \lambda_m} \| Q_t'\bar{\boldsymbol{y}}(G_1; \cdots; G_m) \|^2$ を群間の一般化二乗距離と呼ぶ。これは$\chi^{\dagger 2}(i;i')$の一般化である。$\chi^{\dagger 2}(G_1; \cdots; G_m)$はさらに

$$\chi^{\dagger 2}_{\max} = \max_{\boldsymbol{j}'\boldsymbol{\gamma}=0, \|\boldsymbol{\gamma}\|^2=1} \| Q_t'(\gamma_1\boldsymbol{y}_1 + \cdots + \gamma_n\boldsymbol{y}_n) \|^2$$

により上から押さえられる。$\chi^{\dagger 2}_{\max}$はH_0の下で漸近的に$\sigma^2\omega_2\chi^2_{(2)}(n-1)$の分布に従うことが示される。この漸近分布は行列$Q_t'Q_t$において固有値$\omega_2$が$\mathrm{tr}(Q_t'Q_t) = \sum\omega_k$の３／４を超えることから、$n$がある程度大きいとき、非常に良い近似を与える。

4. 誤差成分の推定

誤差成分の推定には系統的成分推定と逆特性の統計量を構成すればよい。そこで２次形式$\chi^{\dagger 2}(i;i')$の一般化逆行列により、$S_i^- = \| \mathrm{diag}\{(\eta_k\delta_k)^{-1/2}\}L_t'(\boldsymbol{y}_i - \bar{\boldsymbol{y}}) \|^2$ を定義し、$S^- = \sum S_i^-$ を構成する。S^-は (1) 式の$\omega_k\chi^2_{(k)}$を$\omega_k^{-1}\chi^2_{(k)}(n-1)$と置き換えた形式に展開される。

5. 検定統計量とその分布

§§3.4 の結果を用い、統計量$S(i;i')/S^-$ および$S(G_1; \cdots; G_m)/S^-$に基づいて群分けを行う。群分けの有意性は

$$\chi^{\dagger 2}_{\max}/S^- = \omega_2\chi^2_{(2)}(n-1)/\sum\omega_k^{-1}\chi^2_{(k)}(n-1)$$

の分布に従って評価することができる。

6. 応用例

慶応大学医学部で取られた２０３人のデータに適用した結果、５群への有意な群分けが得られた。５群は夜間上昇、平坦、やや下降、下降、極度の下降と明解に特徴付けられ、従来の医師による３分類（上昇、平坦、下降）の合理的な精密化を与えることが確認された。

Non-Null Distributions of the Wald's Criteria in Random-Effects Growth Curve Models

熊本大・工　横山　隆久

ABSTRACT. This paper deals with profile analysis in two extended growth curve models. The first is a growth curve model with parallel mean profiles, which has a random-effects covariance structure based on a single response variable; the second is a multivariate growth curve model with parallel mean profiles, which has a multivariate random-effects covariance structure based on several response variables. For testing "no condition variation" and "level" hypotheses concerning parallel mean profiles of several groups, we obtain the Wald's criteria (Wald (1943)) and their asymptotic non-null distributions. A numerical example is also given.

1. Growth curve model with parallel mean profiles

Suppose that a response variable x has been measured at p different occasions on each of N individuals, and each individual belongs to one of k groups. Let $\boldsymbol{x}_j^{(g)} = (x_{1j}^{(g)}, \ldots, x_{pj}^{(g)})'$ be a p-vector of measurements on the j-th individual in the g-th group, and assume that $\boldsymbol{x}_j^{(g)}$'s are independently distributed as $N_p(\boldsymbol{\mu}^{(g)}, \Sigma)$, and $\boldsymbol{\mu}^{(g)}$'s have parallel profiles, i.e., $\boldsymbol{\mu}^{(g)} = \delta^{(g)}\mathbf{1}_p + \boldsymbol{\mu}$, where $\mathbf{1}_p = (1, \ldots, 1)'$, $\boldsymbol{\delta} = (\delta^{(1)}, \ldots, \delta^{(k-1)})'$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$ are vectors of unknown parameters, Σ is an unknown $p \times p$ positive definite matrix, $j = 1, \cdots, N_g$, $g = 1, \ldots, k$. Without loss of generality we may assume that $\delta^{(k)} = 0$. Then the model of $X = [\boldsymbol{x}_1^{(1)}, \ldots, \boldsymbol{x}_{N_1}^{(1)}, \ldots\ldots, \boldsymbol{x}_1^{(k)}, \ldots, \boldsymbol{x}_{N_k}^{(k)}]'$ can be written as

$$X \sim N_{N\times p}(A_1\boldsymbol{\delta}\mathbf{1}_p' + \mathbf{1}_N\boldsymbol{\mu}',\ \Sigma\otimes I_N), \tag{1}$$

where A_1 is an $N \times (k - 1)$ between-individual design matrix of rank $k - 1$ $(\leq N - p - 1)$, $N = N_1 + \cdots + N_k$. Further, we assume that Σ in (1) has a random-effects covariance structure (see, Rao (1965))

$$\Sigma = \lambda^2\mathbf{1}_p\mathbf{1}_p' + \sigma^2 I_p, \tag{2}$$

where $\lambda^2 \geq 0$ and $\sigma^2 > 0$. Srivastava (1987) obtained the LR tests for "no condition variation" hypothesis

$$H_{01}: \boldsymbol{\mu} = \nu\mathbf{1}_p \quad \text{vs.} \quad H_{11}: \boldsymbol{\mu} \neq \nu\mathbf{1}_p \tag{3}$$

and "level" hypothesis

$$H_{02}: \boldsymbol{\delta} = \mathbf{0} \quad \text{vs.} \quad H_{12}: \boldsymbol{\delta} \neq \mathbf{0} \tag{4}$$

when Σ is unknown positive definite, where $-\infty < \nu < \infty$. Under the assumption that mean response curves of k groups are parallel, the null hypothesis H_{01} means that mean response curves have a flat parallelism, i.e., are horizontal over time. On the other hand, the null hypothesis H_{02} means that mean response curves are coincident. In this section we obtain asymptotic non-null distributions of the Wald's criteria for the hypotheses (3) and (4) under the random-effects covariance structure (2).

2. Multivariate growth curve model with parallel mean profiles

Next we consider an extension of the model (1) to the multiple-response case when m response variables have been measured. Let $\boldsymbol{x}_j^{(g)} = (x_{11j}^{(g)}, \ldots, x_{1mj}^{(g)}, \ldots\ldots, x_{p1j}^{(g)}, \ldots, x_{pmj}^{(g)})'$ be an mp-vector of measurements, and assume that $\boldsymbol{\mu}^{(g)}$'s satisfy $\boldsymbol{\mu}^{(g)} = (\mathbf{1}_p \otimes I_m)\,\boldsymbol{\delta}^{(g)} + \boldsymbol{\mu}$, $g = 1, \ldots, k$. Then the model of X can be written as

$$X \sim N_{N\times mp}(A_1 \Delta (\mathbf{1}_p' \otimes I_m) + \mathbf{1}_N \boldsymbol{\mu}', \ \ \Omega \otimes I_N), \tag{5}$$

where A_1 is the same as described in (1), $\Delta = [\,\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(k-1)}\,]'$ is an unknown $(k-1) \times m$ parameter matrix, $\boldsymbol{\mu}$ is an mp-vector of unknown parameters, Ω is an unknown $mp \times mp$ positive definite matrix. Further, we assume that Ω in (5) has a multivariate random-effects covariance structure (see, Reinsel (1982))

$$\Omega = (\mathbf{1}_p \otimes I_m)\Sigma_\lambda(\mathbf{1}_p' \otimes I_m) + I_p \otimes \Sigma_e, \tag{6}$$

where Σ_λ and Σ_e are arbitrary $m \times m$ positive semi-definite and positive definite matrices, respectively. In this section we obtain asymptotic expansions of the non-null distributions of modified Wald-type statistics for the hypotheses

$$H_{01}: \boldsymbol{\mu} = \mathbf{1}_p \otimes \boldsymbol{\nu} \quad \text{vs.} \quad H_{11}: \boldsymbol{\mu} \neq \mathbf{1}_p \otimes \boldsymbol{\nu} \tag{7}$$

and

$$H_{02}: \Delta = 0 \quad \text{vs.} \quad H_{12}: \Delta \neq 0 \tag{8}$$

under the multivariate random-effects covariance structure (6), where $\boldsymbol{\nu}$ is an m-vector of free parameters. The hypotheses (7) and (8) are extensions of "no condition variation" and "level" hypotheses in the single-response case due to Srivastava (1987) to ones in the multiple-response case.

3. Numerical example

In this section we give a numerical example to illustrate our tests when the sample size is large. We apply the asymptotic results of Section 1 to the repeated measures data (see, e.g., Srivastava (1987)) of the time required for each of 51 students to solve each of four different mathematical problems. Each student belongs to one of four groups (a), (b), (c) and (d). For the observation matrix X, we assume the model (1) in the case $p = 4$, $k = 4$ and $N = 51$. Now we consider the problems of testing the hypotheses (3) and (4) under the random-effects covariance structure (2).

References

[1] C. R. Rao, The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves, Biometrika, **52** (1965), 447 – 458.

[2] G. Reinsel, Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure, J. Amer. Statist. Assoc., **77** (1982), 190 – 195.

[3] M. S. Srivastava, Profile analysis of several groups, Commun. Statist.–Theor. Meth., **16** (3), (1987), 909 – 926.

[4] A. Wald, Tests of statistical hypotheses concerning several parameters when the number of observations is large, Trans. Am. Math. Soc., **54** (1943), 426 – 482.

主成分分析の新しい方法

統計数理研究所 江口 真透

1. はじめに

ニューラルネットワークスにおいて提案されている多くの算法は統計的方法論と密接な関連がある．特に多変量解析の分野とニューラル計算の分野は多くの共通な問題を異なるアプローチから異なる研究者によって独立に提案されている．データの共有化，普遍化が進むにつれて関連する分野が再統合される方向が重要視されるだろう．

この発表では主成分分析についてニューラル計算で提案された自己組織化の方法を統計理論の枠組みで考察した．

この方法の特徴は次の点である．主成分分析の従来の方法は各データに対して共通の重み（データ数の逆数）を与えた標本共分散行列の主固有ベクトルによって定められる．一方でこの方法は各データの重みを固有ベクトルへの残差２乗和に反比例して与えた共分散行列の固有ベクトルによって定められる．重み係数の導出は固有ベクトルから成され，固有ベクトルの導出は重み係数から求める．このように再帰的に重み係数を更新する自然なアルゴリズムが提案された．この自己組織化による方法から定義された主成分ベクトルの影響関数を陽に求めた．影響関数の大域的な振る舞いが安定していることが示された（Higuchi and Eguchi, 1998）．従来の主成分ベクトルの影響関数は非有界な振る舞いをすることが知られている（Critchley, 1985）．一方で正規性を仮定した下での漸近効率の意味では古典的方法が有効であり，その相対効率の公式を与えた．

このような理論的考察を基にして幾つかのモンテカルロ実験を行った．更に Kendall の土壌データ，1998 年の参議院比例区の選挙における都道府県別データの解析を古典的方法と比較した．

2. 理論

$x_1,\cdots,x_n$ を p 変量のデータ

$$\mathcal{P} = \{\rho:(0,\infty)\to(0,\infty)\,|\,\rho'(z)>0,\ \rho''(z)\le 0,\ \rho(0)=0,\ \rho'(0)=1\}$$

とする．この時，

$$L_\rho(\gamma;G)=E_G\{\rho(z_\gamma(\boldsymbol{x}-\bar{\boldsymbol{x}}))\}$$

と定める．ここで G は分布，$z_\gamma(\boldsymbol{x}):=(\|\boldsymbol{x}\|^2-(\gamma^{\mathrm{T}}\boldsymbol{x})^2)/2$ を表す．これより $\rho\in\mathcal{P}$ に対して

$$T_\rho(G):=\mathrm{argmin}\{\ L_\rho(\gamma;G):\gamma\ \text{ s.t. }\ \|\gamma\|=1\}$$

を考察しよう．G の代わりにデータ $x_1,\cdots,x_n$ から作られた経験分布 $\bar{G}_n$ を代入すると $T_\rho(\bar{G}_n)$ は主成分ベクトルの ρ-推定である．ρ の典型は

(i) 古典的方法（Hotelling, 1993）$\rho(z)=1$

(ii) 自己組織化の方法（Oja, 1982, Xu and Yuille, 1995)

$$\rho(z)=\rho_0(z)=-\frac{1+\mathrm{e}^{-\beta\eta}}{\beta}\log\frac{1+\mathrm{e}^{-\beta(z-\eta)}}{1+\mathrm{e}^{\beta\eta}}\ \propto$$

ここで β,η はスケール因子．$T_\rho(G)$ を求めるアルゴリズムは

(i) Given γ, calculate

$$S_{\boldsymbol{\gamma}}:=\sum_{i=1}^{n}\psi(\gamma,\boldsymbol{x}_i)(\boldsymbol{x}_i-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\bar{\boldsymbol{x}})^{\mathrm{T}}\models(\text{1st-eigenvector})\Longrightarrow\ \gamma\ .$$

(ii) Given $S(\gamma)$, calculate γ, by

$$\gamma_1 = \mathrm{argmax}\{\gamma_*^{\mathrm{T}} S(\gamma)\gamma_* : \|\gamma_:\| = 1\}$$

の交代によって作られる. $\gamma_1 = \phi(\gamma)$ とおくと $\gamma_k = \phi^k(\gamma)$ は任意の $k \geq 1$ に対して $\gamma_{k+1} \neq \gamma_k$ の時,

$$L_\rho(\gamma_{k+1}, G) < L_\rho(\gamma_k, G)$$

が示された.

次に, T_ρ の一致性及び影響関数を求める. 影響関数については (Huber, 1985) を参照.

定理. (Kamiya and Eguchi, 1998) $N(\mu, \Sigma)$ を平均 μ, 分散 Σ の p 変量正規分布とし,

$$\Sigma = (\gamma_1 \cdots \gamma_p)\mathrm{diag}(\lambda_1 \cdots \lambda_p)(\gamma_1 \cdots \gamma_p)^{\mathrm{T}}$$

と分解する. ただし $\lambda_1 > \cdots > \lambda_p$. この時

(1) $T_\rho(N(\mu, \Sigma)) = \gamma_1 (\forall \rho \in \mathcal{P})$

(2) $\mathrm{IF}(x, T_\rho, N(\mu, \Sigma)) := \lim_{\epsilon \to 0} \frac{T_\rho(G_\epsilon = (1-\epsilon)N(\mu,\Sigma) + \epsilon\delta_x) - \gamma_1}{\epsilon}$

$$= \psi\left(\frac{\sum_{j=2}^p a_j^2(x)}{2}\right) a_1(x) \sum_{j=2}^p \frac{\lambda_j a_j(x)}{\lambda_j^*(\lambda_1 - \lambda_j)} \gamma_j.$$

ここで $a_j(x) = \gamma_j^{\mathrm{T}} x$,

$$\lambda_j^* = E\left\{\psi\left(\frac{\sum_{k=2}^p a_k^2(x)}{2}\right) a_j^2(x)\right\}.$$

この定理から

(1) $\sup_{z>0} \sqrt{z}\psi(z) = c_0 \Rightarrow \sup_{x : |a_1(x)| \leq c_1} \|\mathrm{IF}(x, T_\rho, N(\mu, \Sigma))\|^2 < (c_0 c_1)^2$

(2) $\mathrm{Eff}(T_\rho) := \frac{\det^+(\mathrm{AV}(T_\rho))}{\det^+(\mathrm{AV}(T_{\mathrm{classic}}))} = \prod_{j=2}^p \frac{\lambda_j^{*2}}{\lambda_j^{**}\lambda_j}$,

が示せる. ここで

$$\lambda_j^{**} = E\left[\psi^2\left(\frac{\sum_{k=2}^p a_k^2(x)}{2}\right) a_j^2(x)\right].$$

参考文献

Critchley, F. (1985). Influence in principal components analysis. *Biometrika*, **72**, 627–636.

Higuchi, I. and Eguchi, S. (1998). The influence function of principal component analysis by self-organizing rule. *Neural Computation*, **10**, 1435–1444.

Hotteling, H. (1933). Analysis of complex of statistical variables into principal components. *J. Educational Psychology*, bf 24, 417-441, 498-520.

Huber, P. J. (1981). *Robust Statistics.* New York: Wiley.

Kamiya, H. and Eguchi, S. (1998). A class of robust principal component vectors. ISM research memo 699.

Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Math. Biol.*, **15**, 267–273.

Xu, L. and Yuille, A. L. (1995). Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE Trans. on Neural Networks*, **6**, 131–143.

平滑化母数の推定方程式とその適用

統計数理研究所　柳本　武美
食品総合研究所　柳本　正勝

1. 序

回帰モデル

$$y = \mu + \epsilon\,, \qquad \epsilon \sim N\,(\,0\,,\ \sigma^2\ I\,)$$

を考える。ここで、$y\epsilon \in R^p$ で $0 \in R^p$ はゼロベクトル、Iは$p \times p$単位行列である。母数モデルとは異なり、平滑化法では $\mu = (\mu_1, ..., \mu_p)'$ において μ_i は i と共に緩やかに変動する場合に良い推定量を求める問題になる。現在多くの方法が提案されている。その中では経験ベイズ法が最も良く用いられている。また GCV 法も良く研究されている。

経験ベイズ法では $\mu \sim N\,(\alpha_1 e_1 + \alpha_2 e_2,\ \tau^2 D^-)$ を仮定する。ここで $e_1 = (1, ..., 1)'$, $e_2 = (1, 2, ..., p)'$, $\quad y' D y = \sum (y_{i-2} y_{i+1} + y_{i+2})^2$ である。この仮定の下で推定量を求める（たとえば Yanagimoto and Yanagimoto 1987）と、$\hat{\mu}\,(\gamma\,;\,y) = (I + \gamma D)^{-1}\,y, \quad \gamma = \tau^2/\sigma^2$ $\hat{\sigma}^2 = y'(I - (I + \gamma D)^{-1})\,y\,/\,(n-2)$。$\gamma$ の推定は$-\,\|y - \hat{\mu}\,\|^2\,/\,tr(I - (I + \gamma D)^{-1}) + \hat{\sigma}^2 = 0$ で求める。

一方 GCV は $\hat{\mu} = (I + \gamma\,D)^{-1}\,y$ が用いられ、$\hat{\sigma}^2 = \|\,(I - (I + \gamma\,D)^{-1})\,y\,\|^2\,/\,tr(I - (I + \gamma\,D)^{-1})$ として、γ は$GCV(\gamma) = \|\,y - \hat{\mu}(\gamma\,;\,y)\,\|^2\,/\,p\{\,tr(I - (I + \gamma\,D)^{-1})\}^2$ を最小にするよう選ばれる。GCV での γ の推定は $\{\mu(\gamma; y); \gamma \in \Gamma\}$ として自由に選んで適用できる。

その他の推定量としては

- $Lowess(x, y, f, iter, delta)$; f が上の γ に相当する
- 移動平均
- Penalized L.S. では $\|\,y - \mu\,\|^2 + \gamma y' D\, y$ を最小にする μ として上記の$\hat{\mu}(\gamma; y)$ が誘導される

がある。

2. 動機

ここでの研究は経験ベイズ法の拡張であって、$\{\mu(\gamma; y); \gamma \in \Gamma\}$ を自由に選べるようにしたい。そうすると

a) 経験ベイズ法の手法を prior 分布を仮定しないで説明する。経験ベイズ法で μ が超母集団からの実現値という仮定は望ましくない場合が多い。

b) Lowess での f を推定する規準を作る。

c) 例えば移動平均

$$\hat{\mu}(\gamma\,;\,y) = (1 - \gamma)\,E\,(\,E'E)^{-1}\,E'y + \gamma\,M\,y$$

ただし M は次の行列

$$
M = \frac{1}{3}\begin{pmatrix} 3 & & & & & 0 \\ 1 & 1 & 1 & & & \\ & 1 & 1 & 1 & & \\ & & & \ddots & & \\ & & & 1 & 1 & 1 \\ 0 & & & & & 3 \end{pmatrix}
$$

この推定量は、M が非負定値行列でないから、Bayes 推定量ではないけれども適用できるようにしたい。その為には経験ベイズ法を prior を用いないで説明して、これを形式的に適用する。

3. 提案する推定方程式

まず $L(\mu_1, \mu_2) = \| \mu_1, -\mu_2 \|^2$ として、$PD(\mu_1, \mu_2, \mu_3) = L(\mu_1, \mu_3) - L(\mu_1, \mu_2) - L(\mu_2, \mu_3)$ とおく。ここで、$\mu_{\mathrm{w}} = E(E'E)^{-1}E'y, \quad E = (e_1, e_2)$ として

$$
E\{PD(y, \hat{\mu}(\gamma\,;\,y), \mu_0\} = E\{g_1(\gamma, \sigma^2\,;\,y)\}
$$

$$
E\{PD(\mu_{\mathrm{w}}, \hat{\mu}(\gamma\,;\,y), \mu_0\} = E\{g_2(\gamma, \sigma^2\,;\,y)\}
$$

となる２つの関数を求める。上の等式で Stein identity を適用すると、

$$
\begin{aligned}
g_1(\gamma, \sigma^2\,;\,y) &= -\|y - \mu(\gamma\,;\,y)\|^2 + \sigma^2(p - \nabla_y \mu(\gamma\,;\,y)) \\
g_2(\gamma, \sigma^2\,;\,y) &= -(y - \mu(\gamma\,;\,y))'(y - \mu_{\mathrm{w}}(y)) + \sigma^2(p - \nabla_y \mu_{\mathrm{w}}(y))
\end{aligned}
$$

ここで $\nabla_y \mu(\gamma; y) = \Sigma \partial \mu_i(\gamma; y)/\partial y_i$ となる。これらが 0 になるよう γ と σ^2 を選ぶと、経験ベイズ法に一致する。この導出では prior 分布を仮定していないことに注意する。上記では $g_1(\gamma; \sigma^2; y)$ についての条件は明示的にではないが文献上見られる。しかし $g_2(\gamma; \sigma^2; y)$ については見当たらない。

拡張の可能性としては Random effect model、GLM があると期待される。

4. 適用

演者らは穀物の生産動向について研究を行ってきた（柳本・柳本 1987, 1997a, 1997b）。現時点での演者らの知見をまとめると次のようになる。

a) 穀物生産は好調であったが、 1980 年代後半から様変わりしている。

b) 人口は増加、収穫面積は頭打ち、単収は増加している。

c) 単収／人口が実際的指標である。これは減少し始めている。

d) Mitchell ら（1997）の「世界食料の展望」は 1997 年に出版されたにも拘わらず、データは 1990 年までしか用いていない。従って、その結論は今日的でない。

e) 経験ベイズ法及び Lowess, 移動平均法で平滑化母数を推定した方法はほぼ同じ傾向を示す。これらの手法は有用である。

時系列解析における離散データの解析について

鹿児島大学理学部　　近藤正男
大学入試センター　　藤井光昭

1　時系列解析における離散データ

時系列解析においては一般に連続値も離散値も現われるが、離散データとして特別な解析法を考えるものとして、ここでは、A）観測時点が離散、B）観測値が離散、の2点に焦点を合わせて論じたい。

時系列が定常性を持つ場合と非定常な場合とで解析法は異なるがここでは定常性を持つ場合に限って話を進めることにする。特に断らない限り時系列 $\{X(t)\}$ の平均は 0 とし、定常性とは $E\,X(t)X(t+h)=R(h)$ を意味する。

2　観測時点が離散

2.1　確率表現

一定時間間隔で観測値を得る場合で、この場合に対応する t として整数値をとることにする。もし実際には連続的な時間パラメータのもとで値が変動しているのに観測は離散時点でしか出来ない場合には、離散時点の観測値から連続的変動のどのような情報が得られるかを明確にする問題が生じる。$X(t)$ 、 $R(h)$ は

$$X(t)=\int \cos 2\pi t\lambda\, dZ_1(\lambda)+\int \sin 2\pi t\lambda\, dZ_2(\lambda) \qquad (Z(\lambda)\text{ は直交増分性を持つ確率過程})$$

$$R(h)=\int \cos 2\pi h\lambda f(\lambda)\, d\lambda \qquad (f(\lambda)\text{ はスペクトル密度})$$

と表現出来、t が連続の場合は積分範囲は $(-\infty,\infty)$ となり、t が離散の場合は積分範囲は $[-\frac{1}{2},\frac{1}{2}]$ となる。t が連続の場合の $f(\lambda)$ を $g(\lambda)$ で表し、離散の場合はそのまま $f(\lambda)$ を用いると

$$f(\lambda)=\sum_{k=-\infty}^{\infty} g(\lambda+k). \qquad (*)$$

2.2　時間パラメータが連続の場合と離散の場合のモデル

時系列モデルには多くの場合、時点ごとに無相関な系列 $\{\epsilon(t)\}$ が介在する。t が連続の場合は形式的に $d\epsilon(t)/dt$ を考えなければならず、通常の意味で $\epsilon(t)$ が微分可能でないため取り扱いが複雑になる。一般に t が連続の場合にはブラウン運動を用いて表現されるなど $X(t)$ に多くの数学的仮定が必要ある。

離散時点モデルとしてよく用いられる ARMA 過程を、時間パラメータが連続の対応する確率過程のなかへ埋めこむ問題などが論じられているが、簡単ではない。t が離散の時

$$X(t)+\phi_1 X(t-1)+\cdots+\phi_p X(t-p)=\theta_0\epsilon(t)+\theta_1\epsilon(t-1)+\cdots+\theta_q\epsilon(t-q)$$

であらわされるモデルを DARMA(p,q) 、t が連続の時

$$X(t)+a_1DX(t)+\cdots+a_PD^PX(t)=b_0B(t)+b_1DB(t)+\cdots+b_QD^QB(t)$$

であらわされるモデルを CARMA(P,Q) とする（$D=\partial/\partial t$ 、$B(t)$ はブラウン運動）。

$X(t)$ が CARMA(P,Q) なら、離散時点 t では DARMA(p,q) である。しかし、DARMA(p,q) が与えられたとき、離散時点でそのようになる CARMA(P,Q) が見出せるかというと必ずしもそうではない。$R(h)$ が h が整数の時一致することと、(*) の条件が成り立つ必要があるからである。

2.3　時間パラメータが連続な定常過程の特性量を離散時点観測で求められるか？

スペクトル密度の推定等でしばしば問題となる。高い周波数の密度が小さいときには、観測時点間隔を短くして可能になる。

3　観測値が離散

観測値が離散の場合の分析法を連続値の場合と区別した形ではあまり論じられていない。

3.1　連続的な観測値が得られるが離散化して解析

連続値離散時点時系列 $\{X(t)\}$ が正規過程のとき、自己相関係数 $\rho(h)$ の通常用いられている推定量は

$$\hat{\rho}_1(h) = \frac{\sum_{t=1}^{n} X(t)X(t+h)}{\sum_{t=1}^{n} X(t)^2}$$

であるが、応用面では他の推定量、例えば $X(t+h)$ の値を正、0、負に応じて 1、0、-1 に変換した推定量

$$\hat{\rho}_2(h) = \frac{\sum_{t=1}^{n} X(t) sgn(X(t+h))}{\sum_{t=1}^{n} |X(t)|}$$

等が用いられている。また、2 値化 (level crossing) した離散値離散時点時系列

$$Z(t) = \begin{cases} 1 & X(t) > u \\ 0 & X(t) \le u \end{cases}$$

の自己相関係数を $\rho_z(h)$ とする。$u=0$ のとき、$\rho_z(h) = \frac{2}{\pi}\sin^{-1}\rho(h)$ が成り立ち、zero-crossings 数 $D = \sum_{t=2}^{n}(Z(t)-Z(t-1))^2$ とおくと、

$$\rho(1) = \cos\Big(\frac{\pi E(D)}{N-1}\Big)$$

が一般の ellipsoidal(elliptically symmetric) 過程でも成り立つことが知られている。

符号化、2 値化の良さとして、計算が簡単、計算が早い、データ打ち切り (truncated)、混合 (contaminated) の影響を受けにくい、アウトライヤーに対してロバストである、相関がより弱くなる、情報がそれほど減少しない等が考えられる。

参考文献

1.Robinson, P. M. (1978),Continuous models fitting from discrete data. *Direction in Time Series.*(ed. D.R.Brillinger and G.C.Tiao) 263-278.

2.Brockwell, P. J. (1995),A note on the embedding of discrete-time ARMA processes.*J. Time Ser. Anal.*, **16**, 451-460.

3.Brockwell, A. E and Brockwell, P. J. (1998),A class of non-embeddable ARMA processes. *J. Time Ser. Anal.*, to appear.

4.Huzii, M.(1962),On a simplified method of the estimation of the correlogram for a stationary Gaussian process. *Ann. Inst. Stat. Math.*, **14**, 451-460.

5.Inagaki, N. and Kondo, M. (1980),Several estimators of the autocorrelation based on limiter estimating functions for a stationary Gaussian process.

6.Kedem, B. (1994),*Time Series Anarysis by Higher Order Crossings.* IEEE PRESS