Generalised information criteria in model selection for locally stationary processes

新潟大学　理学部　蛭川　潤一
Institute for Cancer Research, Rikshospitalet-Radiumhospitalet; University of Oslo
Solvang Kato, Hiroko
早稲田大学　基幹理工学部　玉置　健一郎
早稲田大学　基幹理工学部　谷口　正信

統計的モデルの良さを評価する問題は、情報理論の枠組みを用いて発展してきた。モデルを評価するために、通常、確率過程の構造がある関数によって記述されると仮定する。そのような関数の例としては、i.i.d. の場合の確率密度関数 $p(x)$ 、回帰モデルのトレンド関数 $\mu(u)$、定常過程のスペクトル密度関数 $f(\lambda)$、非線形モデルのダイナミックシステム関数 $\mathbf{F}(X_{t-1},\ldots,X_{t-p})$、等が挙げられる。時系列解析の理論は、定常性の仮定の下で発展してきた。しかしながら、定常時系列モデルは、実際の時系列データを記述するのに十分でない。近年、重要な非定常過程のクラスの一つである局所定常過程が Dahlhaus (1996) によって提案された。局所定常過程の構造はある滑らかな関数で規定される。即ち、時変スペクトル密度関数 $g(u,\lambda)$ によって記述される。ここでは、時変スペクトルモデル $\mathcal{P}=\{f_\theta(u,\lambda):\theta\in\Theta\subset\mathbf{R}^q\}$ を真のスペクトル密度関数 $g(u,\lambda)$ を持つ局所定常過程にあてはめることを考える。Dahlhaus (1996)、 Van Bellegem and Dahlhaus (2006) は、正規 Kullback-Leibler 情報量に基づく局所定常過程についてのモデル選択基準を提案した。ここでは、より一般的に、時変スペクトル密度関数の非線形汎関数に基づく局所定常過程についてのモデル選択基準を提案する。また、真の時変スペクトル密度関数 $g(u,\lambda)$ がモデル $\mathcal{P}$ に属するという仮定を必要としない。

局所定常過程における様々な重要な量は、多くの場合、時変スペクトル密度関数の汎関数によって記述される。線形な汎関数については、自然な推定量が未知の時変スペクトル密度関数を局所ペリオドグラムで置き換えることで得ることが出来る。しかしながら、興味のある汎関数は、いつでも線形であるとは限らない。そのような場合には、非一致性を避けるために、局所ペリオドグラムの代わりに、ノンパラメトリックなカーネル型時変スペクトル密度関数推定量を用いる必要がある。

簡単のため $\{X_{t,T}\}$ $(t=1,\ldots,T;T\geq 1)$ は平均零の 1 変数局所定常過程で、時変スペクトル密度関数 $g(u,\lambda)$ を持つとする。今、観測系列 $\mathbf{X}_T=\left\{X_{2-N/2,T},\ldots,X_{1,T},\ldots,X_{T,T},\ldots,X_{T+N/2,T}\right\}$ が利用可能であるとする。本発表では、時刻 u における局所距離関数

$$D(\boldsymbol{\theta},g,u)=\int_{-\pi}^{\pi}K\{\boldsymbol{\theta},g(u,\lambda),u,\lambda\}d\lambda \tag{1}$$

を用いる。定常過程についての、このような距離関数の具体例は、Dahlhaus and Wefelmeyer (1996) や Taniguchi and Kakizawa (2000) に見つけることができ、それらは、局所定常過程に自然に拡張される。

汎関数 S を

$$D\{S_g(u),g,u\}=\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}D(\boldsymbol{\theta},g,u) \tag{2}$$

を満たすものとして定義する。

$\hat{g}_T$ をノンパラメトリック時変スペクトル密度関数推定量

$$\hat{g}_T(u,\lambda) = \int_{-\pi}^{\pi} W_M(\lambda-\mu) I_N(u,\mu) d\mu, \tag{3}$$

とする。ここに、

$$I_N(u,\lambda) = \frac{1}{2\pi H_N} \left| \sum_{s=0}^{N-1} h\left(\frac{s}{N}\right) X_{[uT]-N/2+s+1,T} \exp(i\lambda s) \right|^2 \tag{4}$$

は時刻 u における局所ペリオドグラムとする。$S_g(u)$ の推定量は $\hat{g}_T$ を用いて、$S_{\hat{g}_T}(u)$ によって自然に定義される。

定理 1. $S_{\hat{g}_T}(u)$ は

$$\begin{aligned} &\sqrt{N}\{S_{\hat{g}_T}(u) - S_g(u)\} \\ &\xrightarrow{d} N(\mathbf{0}, v(h) D_g(u)^{-1} \Gamma_g(u) D_g(u)^{-1}) \end{aligned} \tag{5}$$

を満たす。

特に、コントラスト型の $K(\cdot,\cdot,\cdot,\cdot)$ について、ガウス過程の場合、または、 $\boldsymbol{\theta}$ が *innovation free* のとき、 $\sqrt{N}\{S^c_{\hat{g}_T}(u) - \boldsymbol{\theta}\}$ の漸近分散は $4\pi v(h) \tilde{D}_g(u)^{-1}$ になる。

さて、$D(\boldsymbol{\theta}, \hat{g}_T, u)$ を最小にする値 $S_{\hat{g}_T}(u)$ によって $\boldsymbol{\theta}$ を推定して、パラメトリックモデル $\mathcal{P} = \{f_{\boldsymbol{\theta}}(u,\lambda) : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbf{R}^q\}$ を g にあてはめたことを思い出そう。$f_{S_{\hat{g}_T}(u)}$ と g の近さは $D(S_{\hat{g}_T}(u), g, u)$ によって測ることが出来る。$D(S_{\hat{g}_T}(u), g, u)$ の自然な推定量は g に $\hat{g}_T$ を代入して $D(S_{\hat{g}_T}(u), \hat{g}_T, u)$ によって得ることが出来る。これは、通常 $D(S_{\hat{g}_T}(u), g, u)$ を過小評価する。そのバイアスを

$$b_g(u) = E_{X_T}\left\{D(S_{\hat{g}_T}(u), \hat{g}_T, u) - D(S_{\hat{g}_T}(u), g, u)\right\} \tag{6}$$

として、$D(S_{\hat{g}_T}(u), \hat{g}_T, u) - b_{\hat{g}_T}(u)$ によって、時刻 u における一般化情報量基準を定義する。

$D_g(u)$ と $\Gamma_g(u)$ は g を含むので、g を $\hat{g}_T$ で置き換えると、$E_{X_T}\left\{D(S_{\hat{g}_T}(u), g, u)\right\}$ の推定量として、$G_N(q) = D(S_{\hat{g}_T}(u), \hat{g}_T, u) + N^{-1} v(h) \mathrm{tr}\left\{D_{\hat{g}_T}(u)^{-1} \Gamma_{\hat{g}_T}(u)\right\}$ を得る。$G_N(q)$ を N 倍して

$$\mathrm{GIC}(q) = N D(S_{\hat{g}_T}(u), \hat{g}_T, u) + v(h) \mathrm{tr}\left\{D_{\hat{g}_T}(u)^{-1} \Gamma_{\hat{g}_T}(u)\right\} \tag{7}$$

を一般化情報量基準と呼ぶ。

特に、コンストラクト型の $K(\cdot,\cdot,\cdot,\cdot)$ について、ガウス過程の場合、または、 $\boldsymbol{\theta}$ が innovation free のとき、情報量基準 $\mathrm{GIC}(q)$ は赤池情報量基準

$$\mathrm{AIC}(q) = N D(S_{\hat{g}_T}(u), \hat{g}_T, u) + 4\pi H^{(2)}(1) v(h) q \tag{8}$$

になる。

[1] Dahlhaus, R. (1996). Maximum likelihood estimation and model selection for locally stationary processes. *J. Nonparametr. Statist.*, **6**, 171–191.

[2] Dahlhaus, R. and Wefelmeyer, W. (1996). Asymptotically optimal estimation in misspecified time series models. *Ann. Statist.*, **24**, 952–974.

[3] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic theory of statistical inference for time series* (New York: Springer-Verlag).

[4] Van Bellegem, S. and Dahlhaus, R. (2006). Semiparametric estimation by model selection for locally stationary processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **68**, 721–746.