ベイズアプローチに基づくモデル評価規準

九大・数理　藤岡　由美
九大・数理　井元　清哉
九大・数理　小西　貞則

1　はじめに

説明変数 $\boldsymbol{X}$ と目的変数 Y に関して n 組のデータ $\{(\boldsymbol{x}_\alpha, y_\alpha); \alpha = 1, ..., n\}$ が観測されたとする．非線形回帰においては，平均構造 $\mu_\alpha = E[Y_\alpha|\boldsymbol{x}_\alpha]$ はある滑らかな関数によって記述されるとし，実用的にはスプライン，B-スプライン，局所多項式などが主に用いられる (de Boor, 1978; Green and Silverman, 1994; Simonoff, 1996)．本稿では，B-スプラインに基づく非線形回帰モデルを用いてデータから有効に情報を抽出する方法について検討する．

Eilers and Marx (1996) は，B-スプライン非線形回帰モデルを罰則付き対数尤度関数の最大化に基づいて推定し，この方法を P-スプラインと呼んだ．P-スプラインの特徴は，B-スプラインの節点位置の選択を平滑化パラメータの選択問題に置き換えたことにあり，平滑化パラメータの選択がモデル構築において本質的となる．そこで，Eilers and Marx (1996) は，情報量規準 AIC (Akaike, 1973; 1974) のバイアス補正項を，ハット行列の対角和で置き換えたモデル評価規準に基づいて平滑化パラメータを選択する方法を提案した．しかし，情報量規準 AIC は最尤法に基づいて推定されたモデルの評価規準であり，罰則付き最尤法に基づいて推定された B-スプライン非線形回帰モデルの評価に直接用いることには理論的問題が残る．また，ハット行列による推定曲線の自由度の近似に関してもさらに研究の必要がある．そこで，井元・小西 (1999b) は，これらの問題を理論的に解決し Kullback-Leibler 情報量の 1 つの推定量として情報量規準 SPIC を導出し，B-スプライン基底関数の個数と平滑化パラメータを選択する方法を提案した．

本稿では，これらの Kullback-Leibler 情報量に基づく方法に対して，ベイズ理論，つまりモデルの事後確率に基づくモデル評価について検討する．罰則付き対数尤度関数に含まれる罰則項は，モデルに含まれる未知パラメータの事前分布，平滑化パラメータはその事前分布を規定

するハイパーパラメータと見なすことができる．ベイズ理論に基づくモデル評価規準としては，最尤法に基づいて推定されたモデルを評価するためのベイズ型情報量規準 BIC (Schwarz, 1978) が挙げられる．Akaike (1980) は，事前分布のハイパーパラメータ選択規準として，データの周辺尤度に基づく赤池ベイズ型情報量規準 ABIC を提案した．ABIC に基づく非線形回帰モデルの推定に関しては，石黒・荒畑 (1982)，田辺・田中 (1983)，井元・小西 (1999a) を参照されたい．

本稿では，罰則付き最尤法に基づいて推定されたモデルの評価をベイズの観点から考察し，一つの評価規準を一般的な枠組みで導出する．提案するモデル評価規準は，その特別な場合として BIC を含み，また，その精密化となっていることで BIC の拡張といえる．適用例として罰則付き最尤法に基づいて推定された B-スプラインセミパラメトリック回帰モデル，B-スプライン非線形回帰モデルを評価するためのモデル評価規準を導出し，Kullback-Leibler 情報量に基づく規準と理論的・数値的に比較検討する．

2　モデル設定

説明変数 $\boldsymbol{X} = (X_1, \cdots, X_p) \in R^p$ と目的変数 Y に関して n 個のデータ $\{(\boldsymbol{x}_\alpha, y_\alpha); \alpha = 1, 2, \cdots, n\}$ が観測されたとする．一般に回帰モデルは，各実験点 $\boldsymbol{x}_\alpha$ における目的変数 Y_α の確率的変動に関する成分とその期待値 $\mu_\alpha = E[Y_\alpha|\boldsymbol{x}_\alpha]$ $(\alpha = 1, \cdots, n)$ に対して仮定される系統的成分からなる．各実験点 $\boldsymbol{x}_\alpha$ における目的変数 Y_α の確率的変動は次の指数型分布族に属する密度関数，あるいは確率関数に従うとする．

$$f(y_\alpha|\boldsymbol{x}_\alpha; \xi_\alpha, \phi) = \exp\left\{\frac{y_\alpha \xi_\alpha - u(\xi_\alpha)}{\phi} + v(y_\alpha, \phi)\right\}. \tag{1}$$

ただし，$u(\cdot)$, $v(\cdot, \cdot)$ は既知の関数である．ここで，各条件付き期待値 $E[Y_\alpha|\boldsymbol{x}_\alpha] = \mu_\alpha = u'(\xi_\alpha)$ は，連結関数 h によって $h(\mu_\alpha) = h(u'(\xi_\alpha)) = \eta_\alpha$ と予測子 η_α に関係付けられる．系統的成分では，この予測子 η_α に対してデータの構造と分析目的に応じて次のようなモデルが想定される．

1. 加法モデル: 各説明変数に対して滑らかな関数 w_j $(j = 1, \cdots, p)$ を用いて

$$\eta_\alpha = \sum_{j=1}^{p} w_j(x_{\alpha j}), \quad \alpha = 1, \cdots, n \tag{2}$$

と仮定する．予測子の構造をそれぞれの変数の和で表現した柔軟性の高いモデルである．

2. セミパラメトリック回帰モデル (Green and Silverman; 1994): 説明変数のうちの一つ (z_α) に滑らかな関数 $w(z_\alpha)$ を仮定する．

$$\eta_\alpha = \boldsymbol{x}'_\alpha \boldsymbol{\beta} + w(z_\alpha), \quad \alpha = 1, \cdots, n. \tag{3}$$

ただし，$\boldsymbol{x}_\alpha$ は元のデータから 1 つの説明変数に対するデータを取り除いた $p-1$ 次元ベクトル，$\boldsymbol{\beta}$ は $p-1$ 次元パラメータベクトルである．

ここでは，セミパラメトリック回帰モデルに着目し，(3) 式の滑らかな関数 $w(\cdot)$ として B-スプラインに基づく非線形関数

$$w(z_\alpha) = \sum_{j=1}^{m} a_j B_j(z_\alpha) = \boldsymbol{b}'_\alpha \boldsymbol{a} \tag{4}$$

を用いる．ただし，$B_j(\cdot)$ $(j = 1, \cdots, m)$ は B-スプライン基底関数，$\boldsymbol{b}_\alpha = (B_1(z_\alpha), \cdots, B_m(z_\alpha))'$ は既知の基底関数ベクトル，$\boldsymbol{a} = (a_1, \cdots, a_m)'$ は m 次元パラメータベクトルとする (de Boor, 1978; 井元・小西, 1999a, b).

従って，(1), (3), (4) 式より B-スプラインセミパラメトリック回帰モデルは次式で与えられる．

$$f(y_\alpha | \boldsymbol{x}_\alpha, z_\alpha; \boldsymbol{\beta}, \boldsymbol{a}, \phi) = \exp\left\{ \frac{y_\alpha r(\boldsymbol{x}'_\alpha \boldsymbol{\beta} + \boldsymbol{b}'_\alpha \boldsymbol{a}) - s(\boldsymbol{x}'_\alpha \boldsymbol{\beta} + \boldsymbol{b}'_\alpha \boldsymbol{a})}{\phi} + v(y_\alpha, \phi) \right\}. \tag{5}$$

井元・藤岡・小西 (1998) は，B-スプラインセミパラメトリック回帰モデル (5) 式に含まれる未知パラメータ $\boldsymbol{\beta}$, $\boldsymbol{a}$, ϕ を，次式の罰則付き対数尤度関数の最大化に基づく推定値 $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\phi}$ に置き換えることによって得られる統計モデルを評価するための情報量規準を提案している．

$$l_\lambda(\boldsymbol{\beta}, \boldsymbol{a}, \phi) = \sum_{\alpha=1}^{n} \log f(y_\alpha | \boldsymbol{x}_\alpha, z_\alpha; \boldsymbol{\beta}, \boldsymbol{a}, \phi) - \frac{n\lambda}{2} \boldsymbol{a}' D'_k D_k \boldsymbol{a}. \tag{6}$$

ただし，$\lambda(>0)$ は平滑化パラメータ，D_k は k 階差分を与える $m \times (m-k)$ 行列である．

3　モデル評価規準の構築

説明変数 $\boldsymbol{X} = (X_1, \cdots, X_p) \in R^p$ と目的変数 Y に関して n 個のデータ $\{(\boldsymbol{x}_\alpha, y_\alpha); \alpha = 1, 2, \cdots, n\}$ が観測されたとする．非線形回帰モデルを $f(y_\alpha | \boldsymbol{x}_\alpha; \boldsymbol{\theta})$ で表し，q 次元パラメータベクトル $\boldsymbol{\theta}$ に対する事前分布を $\pi(\boldsymbol{\theta}|\lambda)$ とする．例えば，(5) 式で与えられる B-スプラインセミパラメトリック回帰モデルに対しては $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{a}', \phi)'$ とする．ここで, λ は事前分布を規定

するハイパーパラメータである．モデルの事前確率を π_f とし，$\int \prod_{\alpha=1}^{n} f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\lambda)d\boldsymbol{\theta}$ をモデルの周辺分布に基づく n 個のデータの尤度と考える．このときモデルの事後確率は，分子の規格化定数を無視すると次式で与えられる．

$$\pi_f \int \prod_{\alpha=1}^{n} f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\lambda)d\boldsymbol{\theta}. \tag{7}$$

想定したモデル族の中で (7) 式で与えられる事後確率を最大とするモデルを最適なモデルとして採用する．従って，(7) 式の高次積分をいかに求めるかが問題となり，積分のラプラス近似 (Tierney and Kadene, 1986) を用いることにより

$$\begin{aligned}
&\pi_f \int \prod_{\alpha=1}^{n} f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\lambda)d\boldsymbol{\theta} \\
&\qquad = \pi_f \int \exp\left\{nL(\boldsymbol{\theta}|\lambda)\right\} d\boldsymbol{\theta} \\
&\qquad = \frac{(2\pi)^{q/2}}{n^{q/2}|J(\hat{\boldsymbol{\theta}}|\lambda)|^{1/2}} \exp\left\{nL(\hat{\boldsymbol{\theta}}|\lambda)\right\} \\
&\qquad\quad \times \left[1 + \frac{1}{n}\left\{\frac{1}{24}\sigma^{(4)}L_{(4)}(\hat{\boldsymbol{\theta}}|\lambda) + \frac{1}{72}\sigma^{(6)}L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda)L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda)\right\} + O(n^{-2})\right]
\end{aligned}$$

を得る．ただし，

$$\begin{aligned}
L(\boldsymbol{\theta}|\lambda) &= \frac{1}{n}\sum_{\alpha=1}^{n} \log f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta}) + \frac{1}{n}\log\pi(\boldsymbol{\theta}|\lambda), \\
J(\hat{\boldsymbol{\theta}}|\lambda) &= -\frac{\partial^2 L(\hat{\boldsymbol{\theta}}|\lambda)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}, \\
\sigma_{ij} &= nE[(\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j)], \\
\sigma^{(4)}L_{(4)}(\hat{\boldsymbol{\theta}}|\lambda) &= \sum_{i,j,k,l}^{q} (\sigma_{ij}\sigma_{kl} + \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk})\frac{\partial^4 L(\hat{\boldsymbol{\theta}}|\lambda)}{\partial\theta_i\partial\theta_j\partial\theta_k\partial\theta_l}, \\
\sigma^{(6)}L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda)L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda) &= \sum_{i,j,k,l,m,h}^{q} (\sigma_{ij}\sigma_{kl}\sigma_{mh} + \cdots + \sigma_{ih}\sigma_{jk}\sigma_{lm})\frac{\partial^3 L(\hat{\boldsymbol{\theta}}|\lambda)}{\partial\theta_i\partial\theta_j\partial\theta_k}\frac{\partial^3 L(\hat{\boldsymbol{\theta}}|\lambda)}{\partial\theta_l\partial\theta_m\partial\theta_h},
\end{aligned}$$

とし，$\hat{\boldsymbol{\theta}}$ は $L(\boldsymbol{\theta}|\lambda)$ のモードである．

いま，$\log\pi(\boldsymbol{\theta}|\lambda) = O(1)$ ならば，$\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{ML} = O_p(n^{-1})$ が成り立つ．ただし，$\hat{\boldsymbol{\theta}}_{ML}$ は最尤推定量である．また，$\log\pi(\boldsymbol{\theta}|\lambda) = O(n)$ ならば，罰則付き最尤法に基づく推定量と一致する．以後 $\log\pi(\boldsymbol{\theta}|\lambda) = O(n)$ として議論する．

従って，積分のラプラス近似より，モデルのパラメータ $\boldsymbol{\theta}$ を $L(\boldsymbol{\theta}|\lambda)$ のモード $\hat{\boldsymbol{\theta}}$ で置き換えた統計モデル $f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})$ の評価規準

$$\mathrm{GBIC} = -2\log\left\{\pi_f \int \prod_{\alpha=1}^{n} f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\lambda)d\boldsymbol{\theta}\right\}$$

$$
\begin{aligned}
&= -2\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) + q\log(n/2\pi) + \log|J(\hat{\boldsymbol{\theta}}|\lambda)| - 2\log\pi(\hat{\boldsymbol{\theta}}|\lambda) - 2\log\pi_f \\
&\quad -\frac{2}{n}\left\{\frac{1}{24}\sigma^{(4)}L_{(4)}(\hat{\boldsymbol{\theta}}|\lambda) + \frac{1}{72}\sigma^{(6)}L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda)L_{(3)}(\hat{\boldsymbol{\theta}}|\lambda)\right\} \qquad (8)
\end{aligned}
$$

が得られる．導出した GBIC の値が最小となるモデルを事後確率最大という意味で最適なモデルとして選択する．

(例) 式 (5) で表される B-スプラインセミパラメトリック回帰モデルに含まれるパラメータ $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{a}', \phi)'$ の事前分布に対して $\pi(\boldsymbol{\beta}, \boldsymbol{a}, \phi|\lambda) = \pi_1(\boldsymbol{a}|\lambda)\pi_2(\boldsymbol{\beta}, \phi)$ と仮定する．m 次元パラメータベクトル $\boldsymbol{a}$ の事前分布 $\pi_1(\boldsymbol{a}|\lambda)$ として退化した正規分布

$$
\pi_1(\boldsymbol{a}|\lambda) = \left(\frac{n\lambda}{2\pi}\right)^{(m-k)/2} |D_k'D_k|_+^{1/2} \exp\left(-\frac{n\lambda}{2}\boldsymbol{a}'D_k'D_k\boldsymbol{a}\right) \qquad (9)
$$

を仮定する．ただし，$|D_k'D_k|_+$ は $D_k'D_k$ の 0 でない $m-k$ 個の固有値の積を表す．また，$p-1$ 次元パラメータベクトル $\boldsymbol{\beta}$ とスケールパラメータ ϕ の事前分布は $\pi_2(\boldsymbol{\beta}, \phi) = \text{constant}$ とおく．このとき，

$$
\begin{aligned}
L(\boldsymbol{\beta}, \boldsymbol{a}, \phi|\lambda) &= \frac{1}{n}\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha, z_\alpha; \boldsymbol{\beta}, \boldsymbol{a}, \phi) + \frac{1}{n}\log\pi(\boldsymbol{\beta}, \boldsymbol{a}, \phi|\lambda) \\
&\sim \frac{1}{n}\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha, z_\alpha; \boldsymbol{\beta}, \boldsymbol{a}, \phi) + \frac{1}{n}\log\pi_1(\boldsymbol{a}|\lambda)
\end{aligned}
$$

のモードは罰則付き最尤法 (6) に基づく推定量と一致する．従って，罰則付き最尤法によって推定された B-スプラインセミパラメトリック回帰モデル (5) 式を評価するモデル評価規準 GBIC は定数項を除いて次式で与えられる．

$$
\begin{aligned}
\text{GBIC} &= -2\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha, z_\alpha; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\phi}) - 2\log\pi_1(\hat{\boldsymbol{a}}|\lambda) \\
&\quad + (p+m)\log(n/2\pi) + \log|J(\hat{\boldsymbol{\theta}})| - 2\log\pi_f.
\end{aligned}
$$

4　数値実験

前節で導出したベイズ理論に基づくモデル評価規準 GBIC と従来提案されていたモデル評価規準を，B-スプライン非線形回帰モデルの平滑化パラメータの選択を通して比較検討する．用いるデータは真のモデル

$$
y_\alpha = g(x_\alpha) + \varepsilon_\alpha, \qquad \alpha = 1, \cdots, n
$$

に従って発生させ，真の曲線 $g(x)$ には次の曲線を仮定した．

1. $g(x) = e^{-3x}\sin(3\pi x) \quad (0 \le x \le 1)$
2. $g(x) = 1 - 48x + 218x^2 - 315x^3 + 145x^4 \quad (0 \le x \le 1)$

また，データ点 x_α は一様乱数により発生し，誤差項 ε_α に対しては，混合正規分布を仮定した．

$$\varepsilon_\alpha \sim rN(0,(0.05R_y)^2) + (1-r)N(0,(0.2R_y)^2), \qquad r = 0, 0.1, 0.2, \cdots, 0.9, 1.0$$

ここで，R_y は定義域における真の曲線の値域の幅を表す．図 1, 2 はそれぞれ真の曲線と $r = 0.7$ のとき発生されたデータ $(n = 100)$ の例である．

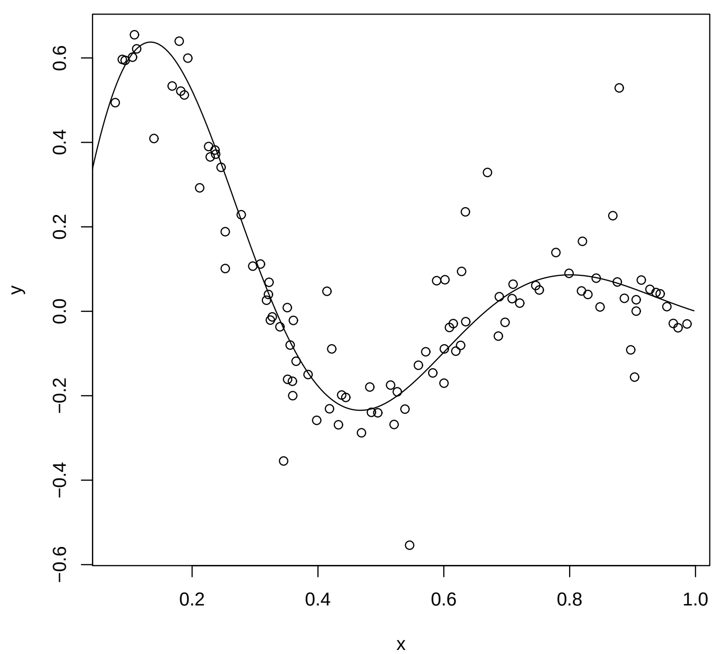

図 1: $g(x) = e^{-3x}\sin(3\pi x)$

このようなデータに対して，次式で定義される B-スプライン非線形回帰モデルを当てはめる．

$$f(y_\alpha|x_\alpha;\boldsymbol{a},\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(y_\alpha - \boldsymbol{b}_\alpha'\boldsymbol{a})^2}{2\sigma^2}\right\}.$$

ただし，$\mu_\alpha = E[Y_\alpha|x_\alpha] = \boldsymbol{b}_\alpha'\boldsymbol{a}$ である．パラメータベクトル $\boldsymbol{a}$ の事前分布には，(9) 式で定義される退化した正規分布，σ^2 に対しては一様分布を仮定した．このとき，$L(\boldsymbol{a},\sigma^2|\lambda) =$

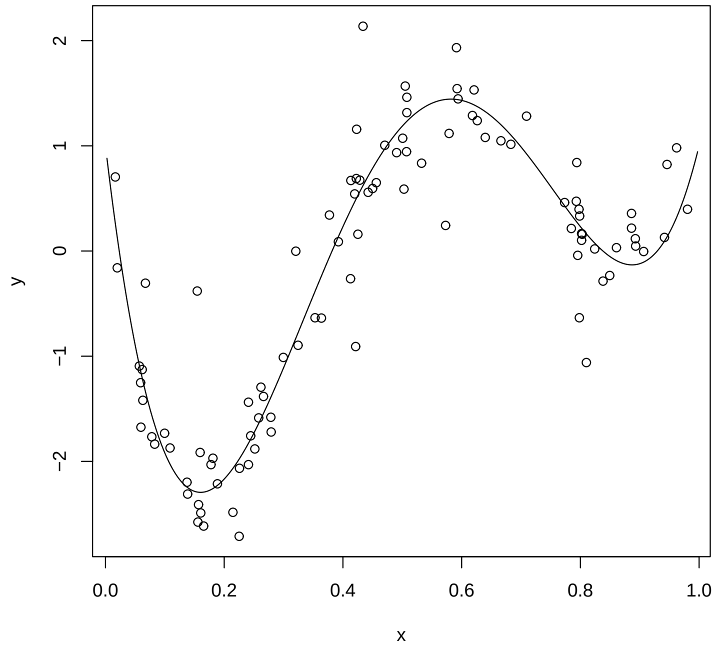

図 2: $g(x) = 1 - 48x + 218x^2 - 315x^3 + 145x^4$

$\sum_{\alpha=1}^{n} \log f(y_\alpha|x_\alpha; \boldsymbol{a}, \sigma^2)/n + \log \pi_1(\boldsymbol{a}|\lambda)/n$ のモードとして与えられる $\hat{\boldsymbol{a}}$, $\hat{\sigma}^2$ に基づく統計モデル $f(y_\alpha|x_\alpha; \hat{\boldsymbol{a}}, \hat{\sigma}^2)$ を評価するための GBIC は次式で与えられる．

$$\begin{aligned}\mathrm{GBIC} &= ||\boldsymbol{y} - B\hat{\boldsymbol{a}}||^2/\hat{\sigma}^2 + n\log(2\pi\hat{\sigma}^2) \\ &\quad + n\lambda\hat{\boldsymbol{a}}' D_k' D_k \hat{\boldsymbol{a}} + k\log n - (m-k)\log\lambda \\ &\quad + \log\left|\frac{1}{n\hat{\sigma}^2}(B'B + n\lambda\hat{\sigma}^2 D_k' D_k)\right| - \log|D_k' D_k|_+ - k\log(2\pi).\end{aligned}$$

提案するモデル評価規準 GBIC と従来の方法を平滑化パラメータの選択を通して比較検討する．ここでは，選択した平滑化パラメータの標準偏差と，推定曲線と真の曲線との平均二乗誤差に基づく比較を行う．

各点 x_α における条件付き期待値 $\mu_\alpha = E[Y_\alpha|x_\alpha]$ の予測値 $\hat{y}_\alpha$ は,

$$S = B(B'B + n\hat{\sigma}^2\lambda D_k' D_k)^{-1}B'$$

と定義される B-スプライン非線形回帰モデルのハット行列 S に対して $\hat{\boldsymbol{y}} = S\boldsymbol{y}$ と表せる．このとき，平滑化パラメータ選択で用いられる従来の規準として，次のようなものが挙げられている．

1. 交差検証法 (CV):

$$\mathrm{CV} = \frac{1}{n}\sum_{\alpha=1}^{n}\left(\frac{y_\alpha - \hat{y}_\alpha}{1 - S_{\alpha\alpha}}\right)^2.$$

ただし，$S_{\alpha\alpha}$ はハット行列 S の第 α 対角成分である．

2. 一般化交差検証法 (GCV, Craven and Wahba, 1979): 交差検証法において，$S_{\alpha\alpha}$ を S の対角成分の平均である $\mathrm{tr}S/n$ で置き換えたものに相当し次式で与えられる．

$$\mathrm{GCV} = \frac{1}{n}\sum_{\alpha=1}^{n}\left(\frac{y_\alpha - \hat{y}_\alpha}{1 - \mathrm{tr}S/n}\right)^2.$$

3. 赤池のベイズ型情報量規準 (ABIC, Akaike, 1980):

$$\begin{aligned}\mathrm{ABIC} &= (n-k)\log(2\pi\hat{\sigma}^2) - (m-k)\log(n\beta) - \log|D_k'D_k|_+ \\ &+ \log|B'B + n\beta D_k'D_k| + n - k\end{aligned}$$

で与えられる．ただし，$\beta = \hat{\sigma}^2\lambda$, $\hat{\sigma}^2 = (||\boldsymbol{y} - B\hat{\boldsymbol{a}}||^2 + n\beta\hat{\boldsymbol{a}}'D_k'D_k\hat{\boldsymbol{a}})/(n-k)$, $\hat{\boldsymbol{a}} = (B'B + n\beta D_k'D_k)^{-1}B'\boldsymbol{y}$ である．

4. 修正 AIC (mAIC, Eilers and Marx, 1996):

$$\mathrm{mAIC} = -2\sum_{\alpha=1}^{n}\log f(y_\alpha|x_\alpha;\hat{\boldsymbol{a}},\hat{\sigma}^2) + 2(\mathrm{tr}S + 1).$$

ただし，$\hat{\boldsymbol{a}}$, $\hat{\sigma}^2$ は罰則付き最尤推定値である．

5. スプライン情報量規準 (SPIC, 井元・小西, 1999b):

$$\mathrm{SPIC} = -2\sum_{\alpha=1}^{n}\log f(y_\alpha|x_\alpha;\hat{\boldsymbol{a}},\hat{\sigma}^2) + 2\mathrm{tr}\left(\hat{I}_{\lambda,m}\hat{J}_{\lambda,m}^{-1}\right).$$

ただし，$\hat{I}_{\lambda,m}$ と $\hat{J}_{\lambda,m}$ は $(m+1)\times(m+1)$ 次行列で次式で与えられる．

$$\hat{I}_{\lambda,m} = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial\psi_\alpha(\boldsymbol{a},\sigma^2)}{\partial\boldsymbol{\theta}}\frac{\partial\log f(y_\alpha|x_\alpha;\boldsymbol{a},\sigma^2)}{\partial\boldsymbol{\theta}'}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \quad \hat{J}_{\lambda,m} = -\frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial^2\psi_\alpha(\boldsymbol{a},\sigma^2)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$$

$$\psi_\alpha(\boldsymbol{a},\sigma^2) = \log f(y_\alpha|x_\alpha;\boldsymbol{a},\sigma^2) - \lambda\boldsymbol{a}'D_k'D_k\boldsymbol{a}/2, \qquad \boldsymbol{\theta} = (\boldsymbol{a}',\sigma^2)'$$

また，BIC の第 2 項 (パラメータ数) を $\mathrm{tr}S$ で置き換えた次式で与えられる修正 BIC (mBIC) も比較対象として用いた．

$$\mathrm{mBIC} = -2\sum_{\alpha=1}^{n}\log f(y_\alpha|x_\alpha;\hat{\boldsymbol{a}},\hat{\sigma}^2) + (\mathrm{tr}S + 1)\log n.$$

表 1: $g(x) = e^{-3x}\sin(3\pi x), r = 0.7, n = 50$

	GBIC	CV	GCV	mAIC	SPIC	ABIC	mBIC
平均 ($\times 10^{-4}$)	4.15	3.07	4.05	3.15	2.12	4.64	7.88
分散 ($\times 10^{-3}$)	1.72	1.84	1.85	1.87	1.95	1.73	1.89
平均二乗誤差	2.05	2.06	2.06	2.06	2.06	2.05	2.05

表 2: $g(x) = e^{-3x}\sin(3\pi x), r = 1.0, n = 50$

	GBIC	CV	GCV	mAIC	SPIC	ABIC	mBIC
平均 ($\times 10^{-5}$)	6.25	6.22	8.44	6.70	3.89	7.42	15.6
分散 ($\times 10^{-4}$)	3.49	3.61	3.61	3.62	3.69	3.50	3.70
平均二乗誤差	2.04	2.04	2.04	2.04	2.04	2.04	2.03

表 3: $g(x) = e^{-3x}\sin(3\pi x), r = 0.7, n = 100$

	GBIC	CV	GCV	mAIC	SPIC	ABIC	mBIC
平均 ($\times 10^{-4}$)	1.91	1.60	2.10	1.85	1.35	2.10	5.32
分散 ($\times 10^{-4}$)	8.37	8.93	8.91	8.91	9.04	8.41	9.70
平均二乗誤差	0.0111	0.0112	0.0112	0.0112	0.0112	0.0111	0.0112

表 4: $g(x) = 1 - 48x + 218x^2 - 315x^3 + 145x^4, r = 0.5, n = 50$

	GBIC	CV	GCV	mAIC	SPIC	ABIC	mBIC
平均 ($\times 10^{-4}$)	7.97	5.48	7.71	6.11	4.07	8.88	16.4
分散	0.0468	0.0514	0.0510	0.0516	0.0541	0.0469	0.0516
平均二乗誤差	0.344	0.348	0.348	0.348	0.351	0.344	0.349

表 5: $g(x) = 1 - 48x + 218x^2 - 315x^3 + 145x^4, r = 0.2, n = 100$

	GBIC	CV	GCV	mAIC	SPIC	ABIC	mBIC
平均 ($\times 10^{-4}$)	5.90	4.28	4.95	4.43	3.21	6.15	15.9
分散	0.0366	0.0404	0.0398	0.0401	0.0415	0.0366	0.0416
平均二乗誤差	0.498	0.502	0.501	0.501	0.503	0.498	0.502

5　考察

1. 情報量規準に基づいた規準 SPIC では選択される平滑化パラメータの値が他の規準によるものより小さくなる傾向がある．
2. mBIC では，選択される平滑化パラメータの値が大きい．このことから，mBIC で選ばれる曲線は平滑化がよく効いていることがわかる．
3. データ数 $n = 100$ では，真の曲線との平均二乗誤差は選択規準による差がほとんどない．
4. 提案した規準 GBIC は，選択される平滑化パラメータの変動を表す分散と真の曲線との平均二乗誤差が他の規準と比べて小さくなっている．ゆえに，データ発生による誤差の影響をあまり受けず，安定した曲線の推定を行うことと，得られた曲線は真の曲線との誤差も小さいということがわかった．

参考文献

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N. and Csaki, F. eds. *2nd Inter. Symp. on Information Theory* Budapest: Akademiai Kiado, pp. 267-281. (Reproduced in S. Kotz and N. L. Johnson, eds. *Breakthroughs in Statistics*, Volume 1, New York: Springer, 1992, pp. 610-624.)

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Contr.*, **AC-19**, 716-723.

Akaike, H. (1980). Likelihood and the Bayes procedure. *Bayesian Statistic* (eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), University Press, Valencia, Spain.

Craven, P. and Wahba, G. (1979) Smoothing noisy data with spline functions. *Numer. Math.*, **31**, 377-403.

de Boor, C. (1978) *A practical Guide to Sprines.* Springer, Berlin.

Eilers, P. and Marx, B. (1996) Flexible smoothing with *B*-splines and penalties (with discussion). *Statistical Science*, **11**, 89-121.

Green, P. J. and Silverman, B. W. (1994) *Nonparametric regression and generalized linear models.* Chapman and Hall, London.

井元清哉・小西貞則 (1999a). B-スプラインによる非線形回帰モデルと情報量規準, 統計数理, 第 47 巻第 2 号, 359-373.

井元清哉，小西貞則 (1999b). 情報量規準に基づく B-スプライン非線形回帰モデルの推定, 応用統計学, **28** (3), 137-150.

井元清哉，藤岡由美，小西貞則 (1998). B-スプラインセミパラメトリックモデルとその推定, 日本数学会秋期総合分科会, 大阪 (大阪大学).

石黒真木夫，荒畑恵美子 (1982). ベイズ型スプライン回帰，統数研彙報，**30** (1), 29-36.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.*, **6**, 461-464.

Simonoff, J. S. (1996) *Smoothing method in statistics.* Springer, New York.

田辺國士，田中輝夫 (1983). ベイズモデルによる曲線・曲面のあてはめ，地球，**5** (3), 179-186.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *J. Amer. Statist. Assoc.* **81**, 82 - 86.