参 考 文 献

S. Amari. (1982). Differential geometry of curved exponential families-curvature and information loss, *Ann. Statist.*, **10**, 357-387.

S. Amari and M. Kumon. (1983). Differential geometry of Edgeworth expansions in curved exponential family, *Ann. Inst. Statist. Math.*, **35**A, 345-365.

S. Amari. (1984). Differential geometry of statistics—towards new developments, METR 83-1 (London Workshop on Differential Geometry in Statistical Inference), Univ. Tokyo.

S. Amari. (1985). Differential-Geometrical Methods in Statistics, Springer Lecture Notes in Statistics, vol. **28**.

ベイズ型重回帰モデル

統計数理研究所 石 黒 真木夫

1. 問 題 意 識

最近，Akaike (1980) の考えにもとづいた ABIC 最小化法によるベイズモデルの利用が活発に行われるようになってきた．

ベイズモデルは性格を異にする2つのモデル，データ分布のモデルとこのモデルのパラメータの分布（事前分布）のモデル，で構成されるが，事前分布のモデルとして現実に利用されているのはいわゆる smoothness prior と呼ばれる形の，パラメータの間に自然な順序がつけられて，その順序にパラメータの値が滑らかに変化する場合のものだけである．

赤池の方法は，すくなくとも形式的には，いかなる事前分布を持って来ても使える一般的な形をしている．従って，様々なデータ分布と事前分布を組合わせることによって多種多様の統計的方法を構成することができる．しかし，そうして導かれる統計的方法のすべてが実用的なものであるとは限らない．

ABIC 最小化法はこれまで成功をおさめてきているが，これは単にベイズモデルを利用したからそうなったということではなく，smoothness prior という極めて特殊な形をした事前分布が何らかの意味で自然なものであったからだ，と考えられる．では，smoothness prior ではない事前分布であって，同じように自然で実用的なものは在りうるか．もし在るものならそのような事前分布を構成してみたい．

2. ベイズモデルと AIC 最小化法

smoothness prior をもちいたベイズモデルによる解析が有効な典型的な例として回帰分析がある（石黒•荒畑，1982）．ベイズモデルを利用することによってデータの大局的な動きを捕えた適度に滑らかな回帰曲線が得られる．

従来の AIC 最小化法によって滑らかな回帰曲線を得ることは，たとえば，多項式回帰における次数選択によって実現されていた(坂元他，1983)．一般に，何等かの意味であまり激しい変動をしない滑らかな推定を得ようとしている場合に「次数選択」という手段が使われるといってよい．逆にある場面で AIC が「次数選択」に使われているなら，そこではベイズモデル，

smoothness prior を用いたベイズモデル，の利用が考えられるのである．

AIC 最小化法とベイズモデルのあいだの，このような関係を手掛りにして次のように考えを進める．

AIC が使われるのは「次数選択」の場面だけではない．「変数選択」にも使われる．「次数選択」に対応して smoothness prior が存在するとすると，「変数選択」に対応する事前分布があって然るべきである。「変数選択」が必要になる典型的な場面は重回帰である．重回帰モデルのパラメータの推定に適した事前分布があってよい．

3. ベイズ型重回帰モデル

重回帰モデル

$$\boldsymbol{y}=X\boldsymbol{a}+\boldsymbol{r} \tag{1}$$

のパラメータ $\boldsymbol{a}$ の推定に使われる 2 つの対照的な方法として，

(a) subset regression（例えば，杉山，1983）

(b) ridge regression（例えば，Marquardt, 1970）

がある．(a) が「変数選択」による方法である．(b) はパラメータの事前分布として球対称の正規分布を仮定した時のベイズ型の推定法とみなすことができる．これら 2 つの方法をくらべてみると，(a) は「意味のある変数」と「無意味な変数」がはっきりと分かれている場合に有効であり，(b) はその境があいまいな場合に効果的である．

見方によっては，(a) の方法も，ベイズ型の推定法の事前分布がパラメータの部分空間に退化した極限と考えることができる．(a) と (b) をつなぐ形として，$\boldsymbol{a}=(a_1, a_2, \cdots, a_m)$ の事前分布に，$p>1$ として，

$$\pi(\boldsymbol{a})=\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi}\ vp\,|\,a_j\,|^{1-1/p}}e^{-\frac{a_j^{2/p}}{2v^2}} \tag{2}$$

を考えてみることにする．ここで a_j^{α} は $\alpha>0$ に対して，

$$a_j^{\alpha}=\begin{cases} \ \ |\,a_j\,|^{\alpha} & a_j\geqq 0 \\ -|a_j|^{\alpha} & a_j<0 \end{cases}$$

で定義する．$p\to\infty$ の極限が (a) の場合，$p\to 1$ の極限が (b) に対応する．v と p が超パラメータということになる．(1) 式の $\boldsymbol{r}$ に正規分布を仮定して尤度関数を構成すれば，これと (2) でベイズ型モデルということになるが，(2) は $a_j\to 0$ の極限で無限大になるので扱いが難しい．(1) と (2) における a_j に $\theta_j=a_j^{1/p}$ なる変換をほどこして得られる

$$\boldsymbol{y}=X(\theta_1^p, \theta_2^p, \cdots, \theta_m^p)^T+\boldsymbol{r} \tag{1$'$}$$

$$\pi(\boldsymbol{\theta}\,|\,v^2)=\prod_{j=1}^{m}\frac{1}{\sqrt{2\pi}\ v}e^{-\frac{\theta_j^2}{2v^2}} \tag{2$'$}$$

の組合せをベイズ型重回帰モデルと呼ぶことにする．

4. 数　値　例

年齢に対する歯の咬耗度の重回帰係数を AIC 最小化による変数選択法によって求めた杉山(1983)の結果と我々のベイズ型重回帰モデルによる推定結果を表 1 に示す．杉山の推定値で空

欄になっているのは変数選択で落とされた変数である．ベイズ型重回帰モデルによる推定値としては(1′)(2′)で決まる事後分布のmodeを採用した．p, v 決定のためのABICの厳密な計算はむずかしいので，(1′)の線形近似による近似値で代用した．

このデータの場合，重回帰係数が全て正の値をとるのが自然であるが変数選択法による推定値の一つが負になっており，説明変数である歯の咬耗度の間の相関が推定に大きく影響していることをうかがわせる．

表1． 年齢に対する歯の咬耗度の重回帰係数の推定値

変数番号	1	2	3	4	5	6	7	8	9	10	11	12	13	14
杉山の結果	1.5		2.4		2.9		1.2							3.3
ベイズ推定	1.1	0.1	2.1	−0.1	1.3	0.0	0.8	0.4	0.0	0.7	0.2	0.2	0.5	2.6

変数番号	15	16	17	18	19	20	21	22	23	24	25	26	27	28
杉山の結果			1.2	2.8	−2.1		1.1							1.1
ベイズ推定	0.1	−0.1	0.1	2.2	−0.1	−0.2	0.6	−0.0	−0.1	−0.1	−0.0	−0.2	0.6	0.7

5． ベイズ型重回帰と変数選択法の比較

(1)式の X として杉山のデータで与えられるもの，$\boldsymbol{a}$ として表1の推定値を用いて人工データを作って，ベイズ型重回帰と変数選択法の各々でパラメータを推定しなおしてみた結果を表2a, bに与える．表中の数値は，発生したデータ数を N，真の分布を f^*，推定されたモデルを f として

$$-2\sum_{i=1}^{N}\int f^*(y_i)\log f(y_i)dy_i$$

表2a． 推定されたモデルの平均対数尤度（表1のベイズ推定を真値とした場合）

実験番号	1	2	3	4	5	6	7	8	9
変数選択法	76.2	48.1 (x)	69.4 (x)	48.1	98.8 (x)	59.2	84.4 (x)	52.0	100.7 (x)
ベイズ法	56.3	38.0 (x)	54.4 (x)	40.5	51.2	47.5	47.7	45.8	58.6

表2b． 推定されたモデルの平均対数尤度（表1の杉山の結果を真値とした場合）

実験番号	10	11	12	13	14	15	16	17
変数選択法	64.5 (x)	91.4 (x)	52.8	48.8 (x)	62.4 (x)	97.1 (x)	74.9 (x)	47.43 (x)
ベイズ法	48.7	60.3	43.5	41.9	48.0	55.6	47.5	42.6

で定義される「平均対数尤度」(とデータ数の積）である．この値が小さいほどモデルのあてはまりが良い．表中のx印は推定されたパラメータにいちじるしい負の値が含まれていることを示す．

参考文献

Akaike, H. (1980), Likelihood and Bayes procedure, in Bayesian Statistics (eds. J.M. Bernardo, M.H. De Groot, D.U. Lindley and A.F.M. Smith), University Press, Valencia, Spain.

Donald W. Marquardt (1970) Generalized Inverses, Ridge Regression, Baiased Linear Estimation, and

Nonlinear Estimation, *Technometrics*, vol. **12**, No. 3.
杉山高一（1983）多変量データ解析入門，朝倉書店.
石黒・荒畑（1982）ベイズ型スプライン回帰，統計数理研究所彙報，第30巻，1号，pp. 29-35.
坂元・石黒・北川（1983），情報量統計学，共立出版.

討　論

質　　問：説明変数のスケーリングはどうするのか.
　答　　：平均0，分散1になるように正規化している.
コメント：subset model の尤度を AIC で定義して平均化する方法と比較してみるとよい.
　答　　：やってみる.

変形する地球の運動と統計モデル

緯度観測所　大　江　昌　嗣

1. はじめに

地球の運動には，色々のモード及び周期があり，それらの観測から地球の構造及び弾性的又は非弾性的性質を知る事が出来る．即ち，地球の内部診断が出来る訳である．弾性又は非弾性的効果は地球の固有周期を左右し，特に非弾性的効果はエネルギーの損失を生じ，運動における位相ずれを起こす．その量は，一般に Q^{-1} によって表わされる．ある周期の変形エネルギーを E とすれば Q^{-1} は一周期で失われるエネルギーの割合であり，

(1)
$$Q^{-1}=\frac{1}{2\pi E}\oint\frac{dE}{dt}\,dt$$

で定義される．

こうした Q の値は一般に周波数に依存する．従って色々の周波数帯域のデータについて Q を求める必要がある．ところで，色々の現象の観測方程式は運動の原因となっている力（励起関数と呼ぶことにする）とそれに対するシステムの応答とが分っていて初めて解く事が出来る．しかし実際の地球においては，この励起関数が100% 分っている訳ではなく，更に観測の擾乱や誤差の影響も見逃がせない．ここに統計的モデルの重要な役割が生じてくる．

2. 極　運　動

極運動はこうした現象の典型的な例であろう．極運動は地球上の質量の移動や偶力の作用によって，自転軸が地球の形状軸（地球の平均的形から決まる軸）に対して動く現象である．地球には，自転軸に対して動き易い固有周期が2つある．1つは約1.2年のチャンドラー周期であり，もう1つは約1日の自由コア章動と呼ばれる周期である．実際の地球では，これらの周期のところに明確な周期をもつ励起関数はなさそうであるが，大気分布等には季節変動があり，それによって極の強制振動が生じる．極運動の季節項は，大気圧分布及び風が山岳部に当たるストレスによって90% 以上説明がつく．一方，チャンドラー運動は不規則な変動によって励起されるランダムプロセスと考えられる．しかし，大気分布の不規則成分によって説明できそうな