

$$p(x_n | Y_{n-1}) \sim \sum_{k=1}^{m_{kn}} \gamma_{kn} \varphi_k(x_n | Y_{n-1})$$

$$p(x_n | Y_n) \sim \sum_{l=1}^{m_{ln}} \delta_{ln} \varphi_l(x_n | Y_n)$$

によって状態ベクトル x_n を推定すればよい。ただし、 $Y_m = \{y_1, \dots, y_m\}$, $m_{kn} = m_{l,n-1} m_{li}$, $m_{ln} = m_{kn} m_{kj}$, $\gamma_{kn} = \alpha_i \delta_{l,n-1}$, $\delta_{ln} = \beta_j \gamma_{kn} \varphi_l(y_n | Y_n)$ であり、 $\varphi_k(x_n | Y_{n-1})$, $\varphi_l(x_n | Y_n)$ の平均および共分散行列はカルマンフィルタを用いて推定することができる。

このガウス和フィルタの実現のためには、各時間ステップにおいて、より少ない項数のガウス分布で近似しなおすための効率的な方法の開発が必要となる。

数値例として、民間企業在庫品増加のデータの解析結果を示した。

数値予測にもとづく降雨確率の推定

石 黒 真木夫

1. 気象予測と統計学

気象予測と統計学がかかわる場面は多いが、ここでとりあげるのは、「短期間予報」におけるかかわりである。短期間予報においても、生データを、数値予報プログラムへの入力に加工する「解析」の段階と、数値予報プログラムの出力を実際の「天気予報」に加工する「天気翻訳作業」の場面における統計学の利用に分けられる。この報告では、後者の「天気翻訳作業」における新しい統計的手法の利用を紹介する。天気翻訳作業のうち確率予測にかかわるものは、どれも同じように扱うことができるが、「降雨確率」の推定を数値例としてとりあげた。

2. 現 状

2.1 降雨確率予報

数値予報で得られる大気の 3 次元的構造の「予測値」を説明変数とし、アメダスの観測にもとづく降雨の「有無」を目的変数とする線型重相関回帰式によって、予測が行われている。説明変数の選択は、50 個ほどの仮予測因子の中から変数増加法によって「5 個」を選んでいる。確率の予測値として、負の数や 100% 以上の値がでる場合があるが、そのような場合には、0% あるいは 100% で打ち切った値が使われる。

表 1. ハイパー・パラメータの値と ABIC

ρ	v	ABIC
1.0	0.50000 D+00	362.74
	0.10000 D+01	307.92
	0.20000 D+01	265.91
	0.40000 D+01	242.28**
	0.80000 D+01	245.23
1.5	0.10000 D+01	346.78
	0.50000 D+00	412.70
	0.20000 D+01	291.03
	0.40000 D+01	265.18*
	0.80000 D+01	305.39

2.2 統計的手法

重回帰モデルにおける変数推定法として従来の変数選択法(変数増加法・増減法, 等を含む)と, リッジ回帰法などにかわるベイズ型重回帰法(石黒(1985))がある. 又, 確率推定のためには, 重回帰モデルより2値回帰モデルの利用が望ましいが, このモデルをデータにあてはめるために必要な計算を高速に実行する環境の整備も大きく進展している.

表2. $p=1, v=4$ の場合のパラメータの推定値
(θ_k の絶対値の降順に並べた θ_k の値と k)

-1.28	21	0.12	48	-0.01	37
-0.61	47	0.12	26	0.01	33
-0.53	38	-0.12	8	-0.01	44
0.52	42	-0.12	34	0.00	11
-0.40	23	-0.10	6	0.00	12
0.36	36	-0.09	24	0.00	13
-0.34	25	-0.09	3	0.00	14
0.28	30	-0.09	29	0.00	15
-0.26	9	0.09	32	0.00	16
0.25	7	0.09	22	0.00	17
-0.25	5	-0.08	31	0.00	18
0.25	41	-0.08	39	0.00	19
-0.22	27	-0.04	40	0.00	20
0.21	49	0.04	46	0.00	50
-0.14	1	-0.03	2	0.00	4
-0.13	10	-0.03	35	0.00	28
0.13	45	-0.01	43		

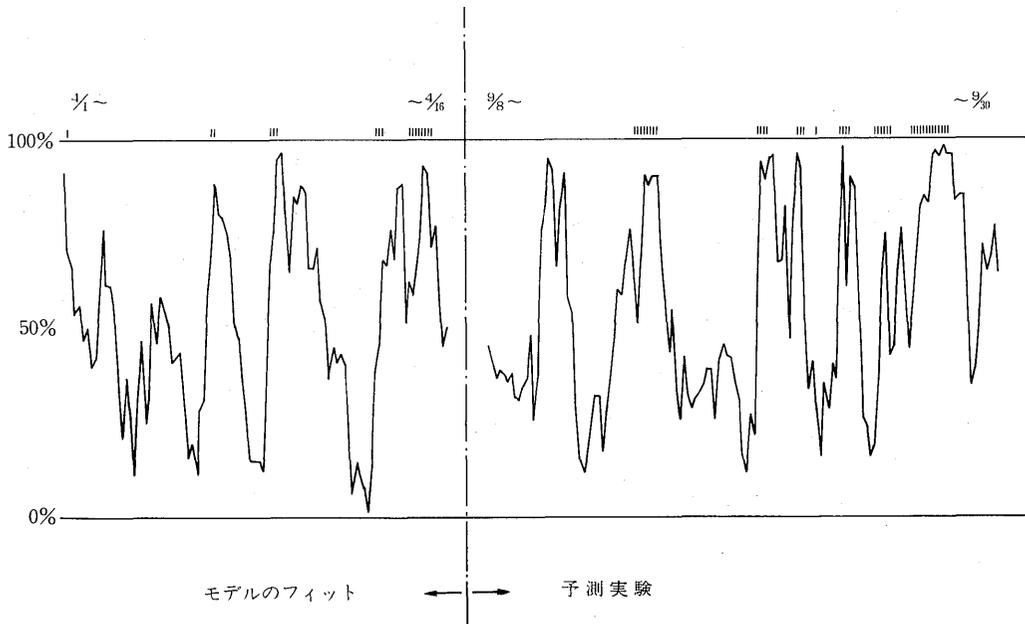


図1. 2値回帰モデルのフィットと予測

3. ベイズ型重 2 値回帰モデル

これは、ベイズ型重回帰の方法を、2 値回帰に応用して、雨の確率 P と説明変数 $x = (x_1, x_2, \dots, x_m)$ の値を結びつける 2 値重回帰の式

$$(3.1) \quad P(\theta) = \frac{\exp\left(\sum_{k=1}^m \theta_k x_k\right)}{1 + \exp\left(\sum_{k=1}^m \theta_k x_k\right)}$$

の係数 $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ を非線形な形、

$$(3.2) \quad \theta_k = \|a_k\|^p \quad (k=1, 2, \dots, m)$$

にパラメトライズした上で、事前分布

$$(3.3) \quad \pi(a) = \prod_{k=1}^m \frac{1}{\sqrt{2\pi} v} e^{-a_k^2/2v^2}$$

のもとで、ベイズ推定する方法である。ただし、(3.2) 式で使った記号 $\|\cdot\|^p$ は非線形変換

$$\|a\|^p = \begin{cases} |a|^p & a \geq 0 \\ -|a|^p & a < 0 \end{cases}$$

を表すものとする。

事前分布のパラメータである p と v は ABIC 最小化によって決める。ただし、ここでの数値例の計算にあたっては、近似的な値の計算で代用した。

4. 数 値 例

富士山の周辺における降雨確率(大雨)の予測式の推定と予測の実験とを行った。使ったデータは 1983 年の暖候期(4-9 月)のものである。データの、前半を係数推定用に、後半を予側実験用に使った。係数推定の結果を表 1, 2 に示した。

表 1 に示すように、 $p=1$ の場合に ABIC が最小となったので、結果的に Ridge 回帰による係数の推定を行った事になる。パラメータの推定値を大きさの順に並べかえた表(表 2)を見ると(説明変数は分散が一定となるように正規化してある)、きわだった説明力を持つ少数の変数があるのではないことがよくわかる。このようなデータの解析に変数選択法はあまり適していないはずである。

図 1 にモデルのフィットの様子と予測例を示す。図中の「折れ線グラフ」がモデルにもとづく降雨確率の推定値を示し、上部の短線が「降雨」が観測されたことを示す。従来の方法との比較が必要だが、かなりうまくいっているように見える。図の 9 月 10 日頃の位置にある「ハズレ」は数値予報の「ハズレ」に起因するものであろうと思われる。

5. 謝 辞

ここで報告した成果は、統計数理研究所共同研究「62-共研-57 数値予報の結果を利用した統計的天気予報手法の改善」のなかで得られたものである。

参 考 文 献

- [1] Akaike, H. (1980). Likelihood and Bayes procedure, in *Bayesian Statistics*, (eds. J. M. Bernardo, M. H. DeGroot, D. U. Lindley and A. F. M. Smith), University Press, Valencia, Spain.
- [2] 坂本・石黒・北川 (1983). 情報量統計学, 共立出版.
- [3] 石黒真木夫 (1985). ベイズ型重回帰モデル, 統計数理, **33**, 8-11.
- [4] 気象庁予報部 (1986). 短期間予報用ガイダンスの解説, 予報技術資料第 35 号.

- [5] Ishiguro, M. and Sakamoto, Y. (1983). A Bayesian approach to binary response curve estimation, *Ann. Inst. Statist. Math.*, **35**, 115-137.
- [6] Sakamoto, Y. and Ishiguro, M. (1985). Bayesian binary regression involving two explanatory variables, *Ann. Inst. Statist. Math.*, **37**, 369-387.

玖珠川水系の流量予測の試み

荒 畑 恵美子

九州の福岡県, 大分県, 熊本県にまたがる玖珠川水系の流量予測を試みた。流量に関連するデータは, ダムおよび河川の水位, 発電量等である。連接水系のすべての河川の水位を観測しているわけではないので, 流量の推定に困難が生じる。

各時点で, 各々のダム貯水池での流量の入出力関係をモデル化すると, パラメータ数がデータ数より多くなる。この不足するデータから流量を推定するために, スムースネス・プライヤを導入して, ベイズモデルを構成した。この推定法を実装化するため, モデルを状態空間表現し, カルマンフィルタを用いて計算するプログラムを作製した。このとき, 観測データ不足を補うため, 2次の階差を観測方程式に加えダミー・データとして, 零を与え, 推定の安定化を計った。

これは, 統計数理研究所, 九州電機製造株式会社, 大阪大学との共同研究の一部である。

山下法と Freund 法について

土 谷 隆

Karmarkar [1]が線形計画問題に対する射影変換を用いた多項式オーダーの内点法を提案して以来, 線形計画問題の内点法に関して精力的な研究が進められている。いわゆる標準形の線形計画問題の双対問題 $\langle D \rangle$ (n 変数 m 制約式) :

$$\begin{aligned} \min c^t x - c_0, & \quad \text{subject to } A^t x \leq b, \\ A \in R^{n \times m}, \quad x, c \in R^n, \quad b \in R^m \end{aligned}$$

($m > n$ で A の rank は n ; 制約領域は空でなく有界; 解は非退化で目的関数の最小値 c_0 を 0 とする。) については, 山下と Freund によって, 独立に, 射影の概念を用いた多項式性を持つ解法が提案されている ([2], [3])。これら2つの方法は, $\langle D \rangle$ を解くことを $\langle D \rangle$ に対する log potential 関数

$$f(x) = m \log(c^t x - c_0) - \log \Pi(A^t x - b)$$

(Πv でベクトル v の各要素の積を表す。) の最小化として捉え, 最小化問題を解くために射影を用いて $\langle D \rangle$ を変換する (log potential 関数の値を不変に保つような射影変換) 点において類似しているが, その変換, 方向の導出法は全く異なる。しかし, 山下法を (その本質は損なわずに) 若干修正したものと Freund の方法とは同一の方向を与える。本発表ではその事情を明らかにし, 2つの方法が, 「 $\langle D \rangle$ を同次化した線形計画問題 $\langle DH \rangle$:

$$\min c^t z - c_0 z_0, \quad \text{subject to } A^t z \leq b z_0, \quad z \in R^n, \quad z_0 \in R$$

に関する log potential 関数

$$f_H(z, z_0) = m \log(c^t z - c_0 z_0) - \log \Pi(A^t z - b z_0)$$

に対し, $\langle DH \rangle$ に関する log barrier 関数 $\log \Pi(A^t z - b z_0)$ の Hesse 行列を計量とする最急降下法を行っている。」あるいは, 「 $\langle DH \rangle$ に対して Adler-Karmarkar 法を適用している。」と考えることによって統一的に捉えられることを示した([4])。さらに, これらの方法と Karmarkar 法自身とのより直接的な対応を見いだし, 山下法, Freund 法が Karmarkar 法自身と, ある意味で等価なものであることも明らか