状態空間モデル

吉田　亮

状態空間モデルは，状態変数（state variables）と呼ばれる k 個の潜在確率変数（latent random variables）と p 個の出力変数の時間発展を表す確率モデルの総称である．システムの制御対象に加わる入力と制御対象が生成する出力に対する表現形式の呼称であり，状態空間表現（state space representation）と呼ばれることも多い．適用分野は制御工学，信号処理，時系列解析など多岐にわたる．

1．線形動的システム

時刻 $t=n\Delta t\ (n=0,1,\cdots)$ のシステムの内部状態を k 次元状態ベクトル $x_n=(x_{1n},\cdots,x_{kn})$ で表す．x_n は出力変数に内在する未知の潜在変数として取り扱われる．このような取り扱いは解析上の便宜的要請であることも多い．いま同時刻の p 次元観測値ベクトルを $y_n=(y_{1n},\cdots,y_{pn})$ と記す．線形状態空間モデル（linear state space model）は状態変数と観測値の関係を次の二つの方程式で表現したものである．

$$x_n=F_nx_{n-1}+G_nv_n \tag{1}$$

$$y_n=H_nx_n+w_n. \tag{2}$$

ここで v_n と w_n は互いに独立な白色雑音（white noise）で，それぞれ観測ノイズ（measurement noise），システムノイズ（system noise）と呼ばれる．初期状態ベクトル x_0 とこれらのノイズ列には統計的独立性が仮定される．H_n，F_n，G_n はそれぞれ適当な次元の行列である．式(1)は，システム方程式（system model）あるいは状態遷移方程式（state transition equation）と呼ばれ，状態ベクトルの時間推移を1次マルコフ過程によって表現したものである．式(2)は，観測方程式と呼ばれる．F_n と H_n はそれぞれ状態遷移行列，観測行列と呼ばれる．二つのノイズ項 (v_n, w_n) に多変量正規分布を仮定したモデルに対して，線形ガウス状態空間モデル（linear Gaussian state space model）という特別な呼称が使われる．

一般に，ARMA モデル（autoregressive moving average model），トレンド成分モデル，季節調整モデルなど，時系列解析で用いられる多くのモデルを線形ガウス状態空間モデルの形式で表現できる[2][5]．その場合，統計的推定に付随する諸計算（尤度計算や状態の条件付分布の計算）にカルマンフィルタ（Kalman filter）を援用することができる．例えば，式(2)を出力 y_n に対する回帰式と見なせば，H_n は入力変数を要素に持つ行列，状態 x_n はシステムモデルに従う時変回帰係数と解釈できる．また，状態ベクトルを時系列データに含まれるトレンド成分あるいは季節変動成分など，推定すべき未知の信号と考えることもできる．信号変化の動的特性と出力 y_n の観測メカニズムを(1)と(2)の形式で表現した後，一般的な統計理論を適用することで，系の内部状態に対する統一的な推定方式を設計することができる．

時系列解析において H_n，F_n，G_n が時間変化することは比較的稀であり，そのような場合は時間を表す添え字 n が省略される (H,F,G)．また，H，F，G の値が未知の場合，状態とともにこれらの値に対しても推定を行う．例えば，自己回帰モデル（autoregressive model）の状態空間表現では，H，F，G の各要素は自己回帰係数によって構成される[2]．この場合は，状態の推定よりもむしろ，H，F，G の推定が主たる興味となる．システムノイズや観測ノイズの分布に関するパラメータも同様に未知パラメータとして取り扱われることも多い．状態 x_n 以外のすべての未知パラメータ θ は超パラメータ（hyper-parameter）と呼ばれる．線形ガウス状態空間モデルの超パラメータと状態の最尤推定あるいはベイズ推定は，EM アルゴリズム（expectation-maximization algorithm）とカルマンフィルタの繰り返し計算によって比較的容易に実現できる[3][5]．

よしだ　りょう
統計数理研究所
〒190-8562 立川市緑町 10-3

2．一般化状態空間表現

一般化状態空間表現（generalized state space representation）は，系の内部状態の時間推移と観測メカニズムに対して，一般の条件付き確率分布を使ってモデル化したものである．

$$x_n|x_{n-1} \sim p(x_n|x_{n-1}, \theta) \tag{3}$$

$$y_n|x_n \sim p(y_n|x_n, \theta). \tag{4}$$

条件付き確率分布 $p(a|b)$ は，二つの確率変数 a と b が構成する系において，b の値を固定したときの a の確率分布を表す．θ は超パラメータである．システムモデル(3)と観測モデル(4)には，任意の分布型が与えられる．一例として，(1)と(2)のシステムノイズと観測ノイズに対して正規白色ノイズを仮定した線形ガウス状態空間モデルの表現形式を示す（$v_n \sim N(0, V)$, $w_n \sim N(0, R)$）．ここで $X \sim N(\mu, \Sigma)$ は，確率変数 X が平均 μ，分散共分散行列 Σ を持つ正規分布にしたがうことを意味する．この場合，一般化状態空間表現の二つの条件付き分布は共に多変量正規分布になり，それぞれ $x_n|x_{n-1} \sim N(Fx_{n-1}, GVG^{\mathrm{T}})$, $y_n|x_n \sim N(Hx_n, R)$ という形で表現される．

内部状態や超パラメータのベイズ推定量は，観測値が与えられたもとでの θ や x_n の事後分布 $p(\theta, x_0, x_1, \cdots, x_n|y_1, \cdots, y_n)$ に関する何らかの統計量である．最も自然な推定量は事後分布の最大値や平均である．事後分布の計算には条件付き分布の逐次計算則である平滑化公式を利用することができるが，非ガウスモデルの事後分布計算を実現するには近似アルゴリズムを適用しなくてはならない．

3．推定アルゴリズム

超パラメータ θ について任意の値を設定したもとで，j 時点までの全観測値 $Y_j=\{y_1, \cdots, y_j\}$ から時刻 n の状態 x_n を推定する問題について考える．ベイズ統計学では，x_n の推定量は事後分布 $p(x_n|Y_j)$ に関する何らかの統計量によってあたえらえる．例えば，事後分布の最大値 $\arg\max_{x_n} p(x_n|Y_j)$ や平均 $\mathbb{E}[x_n|Y_j]$ が自然な推定量である．観測区間より先の将来 $n>j$ を推定する問題は予測と呼ばれる．現在の全観測値を用いて最終時刻 $n=j$ の状態を推定する問題をフィルタと呼ぶ．また $n<j$ の場合は平滑化と呼ばれる．いずれの問題も，条件付き分布 $p(x_n|Y_j)$ を求めることが推定量の計算において必須となる．

線形ガウス状態空間モデルの場合，予測，フィルタ，平滑化の条件付き分布はすべて多変量正規分布になる．したがって，事後分布の評価に関わる操作は，予測，フィルタ，平滑化分布の一次と二次のモーメント，すなわち平均ベクトルと分散共分散行列を計算することに帰着される．カルマンフィルタおよびカルマンの平滑化と呼ばれるアルゴリズムは，状態変数の条件付き平均と分散共分散行列を効率的に計算するための漸化式である[2]．

非ガウスあるいは非線形状態空間モデルでは，状態の条件付き分布 $p(x_n|Y_j)$ も非ガウス分布になり，分布型を陽に求めることはできない．したがって，確率分布自体を数値的に近似評価する必要がある．正規分布の平均と分散共分散行列を調整しながら $p(x_n|Y_j)$ の近似を実現する方法として，拡張カルマンフィルタ（extended Kalman filter）と呼ばれる手法がある[1]．非線形モデルを線形近似したもとで，近似用正規分布の平均ベクトルと分散共分散行列をカルマンフィルタの更新則で逐次的に計算するというのが基本思想である．しかしながら当然，線形近似が妥当でない系を取り扱う際，推定結果に大きなバイアスが生じる．

近年の計算機の高速化によって，粒子フィルタ（particle filter）やマルコフ連鎖モンテカルロ法（MCMC：Markov chain Monte Carlo）など，モンテカルロ法の原理を状態分布の近似計算に援用する方法が広く使われるようになった[2][4]．いずれの方法も予測，フィルタ，平滑化分布を有限個のモンテカルロサンプルで近似する．したがって，状態の次元が小さい問題では比較的低コストで良好な近似性能を実現することができるが，問題が高次元化するにつれ必要なモンテカルロサンプル数が指数的に増大するという欠点もある．モンテカルロ法に基づくアプローチとは別に，変分ベイズ（variational Bayes method）あるいはベイズ平均場近似（mean-field approximation）と呼ばれる汎関数最適化に基づく手法も提案されている[6]．状態の同時条件付き分布 $p(X_n|Y_j)$ と分布型を指定しない任意の近似用分布 $q(X_n)$ の間にカルバック＝ライブラー情報量を導入して，交互最適化によって $q(X_n)$ の関数形を求めるという手法である．状態と同時に超パラメータも推定する場合には，拡大した事後分布 $p(\theta, X_n|Y_j)$ に対して粒子フィルタやマルコフ連鎖モンテカルロ，変分ベイズ法を適用することで，状態空間モデルの推定問題を解くことができる．

参考文献

[1] 片山徹,「応用カルマンフィルタ」, 朝倉書店, 2000.

[2] 北川源四郎,「時系列解析入門」, 岩波書店, 2005.

[3] P. J. ブロックウェル, R. A. デービス,「入門 時系列解析と予測」(改訂第2版)(訳：逸見功, 田中稔, 宇佐美嘉弘, 渡辺則生), シーエーピー出版, 2004.

[4] A. Doucet, N. de Freitas and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer, 2001.

[5] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications : With R Examples*, Springer, 2006.

[6] M. J. Wainwright and M. I. Jordan, Graphical Models, Exponential Families, and Variational Inference, *Foundations and Trends in Machine Learning*, 1 : 1-305, 2008.

OR 事典 Wiki のページに掲載された用語解説は，順次，OR 学会ホームページの OR 事典 Wiki のコーナーでも公開されます．ほかにも多くの OR 用語の解説が掲載されていますので，是非ご活用下さい．
http://www.orsj.or.jp/~wiki/wiki/