岩波データサイエンス（全6巻）

〈全巻の構成〉

Vol.1 特集「ベイズ推論と MCMC のフリーソフト」

Vol.2 特集「統計的自然言語処理 — ことばを扱う機械」

Vol.3 特集「因果推論 — 実世界のデータから因果を読む」

Vol.4 特集「地理空間情報処理」

Vol.5 特集「スパースモデリングと多変量データ解析」

Vol.6 特集「時系列解析 — 状態空間モデル・因果解析・ビジネス応用」

岩波データサイエンス Vol. 6

2017 年 6 月 23 日　第 1 刷発行

編　者　岩波データサイエンス刊行委員会

発行者　岡 本　厚

発行所　株式会社 岩波書店
〒 101-8002 東京都千代田区一ツ橋 2-5-5
電話案内 03-5210-4000
http://www.iwanami.co.jp/

印刷製本・三秀舎

© 岩波データサイエンス刊行委員会 2017
ISBN 978-4-00-029856-8　　Printed in Japan

IWANAMI
DATA
SCIENCE Vol. 6
岩波データサイエンス

特集「時系列解析——状態空間モデル・因果解析・ビジネス応用」
伊庭幸人 …… 4

状態空間モデル　**北川源四郎** …… 5

Rによる状態空間モデリング……dlmとKFASを用いて
伊東宏樹 …… 32

〈コラム〉状態空間モデルのパラメータ推定 …… 48

粒子フィルタを実装してみる
岩波データサイエンス刊行委員会 …… 49

〈コラム〉推定のための計算手法の比較 …… 56

状態空間モデルのマーケティングへの応用
佐藤忠彦 …… 57

VARモデルによる因果関係の推論
……内閣支持率と株価を例に　**川崎能典** …… 68

〈コラム〉隠れマルコフモデルと状態空間モデル …… 82

小特集「シミュレーションとデータサイエンス」
伊庭幸人 …… 97

シミュレーション，データ同化，
そしてエミュレーション　**樋口知之** …… 98

〈コラム〉エミュレータの活用 …… 106

天気予報とデータサイエンス
伊藤耕介……107

揺らぐタンパク質と老いる私
……ミスフォールディング時代のデータサイエンス
伊庭幸人……121

〈コラム〉分子シミュレーションの「逆問題」とは?
……131

リアルなSimCityの夢　**仙石裕明**……134

［**話題**］

時系列データから天体の形を再構成する
……時間軸天文学への誘い　**植村 誠**……86

夢と脳と機械学習……夢理論から夢のデータサイエンスへ
神谷之康……139

［**パズル・小説・漫画**］

計算機で作る面白いナンプレ⑥　**とん**……150

掌編小説《海に溺れて》❻流されて
円城 塔……84

掌編漫画 ②いつもの所で待ち合わせ　**panpanya**……132

シリーズ『岩波データサイエンス』の完結にあたって……152

表紙挿画＝蛯名優子
デザイン＝佐藤篤司

特集 時系列解析
――状態空間モデル・因果解析・ビジネス応用

『岩波データサイエンス』Vol. 6 の特集では，データサイエンスで最も重要なテーマのひとつである時系列解析を主題とする。

現在の流れとして「入門から状態空間モデルを教える」のが，トレンドになりつつある。そこで，最初の 4 つの解説は**状態空間モデルとその応用**に関するものとした。状態空間モデルは Vol. 1 でも扱っており，そこでは主に MCMC の応用を取り上げたが，本特集ではカルマンフィルタや粒子フィルタなどを含む全体像を提示したい。

冒頭は**北川**による総論である。この分野で著名な研究者である著者に新たに書き下ろしていただいたもので，状態空間モデルによる時系列解析入門から粒子フィルタの応用までがコンパクトにまとめられている。次の 2 編が「実践編」に相当する。**伊東**の解説では R のパッケージ dlm と KFAS を取り上げるが，両者で同じ問題を実装し，また Vol. 1 の特集とも連携するなど，相互の翻訳に役立つ「ロゼッタ・ストーン」的な面を持たせている。それに続く刊行委員会による解説では，粒子フィルタを R 言語で作ってみる。4 編目の**佐藤**の解説では「マーケティングの現場で状態空間モデルをどう生かすか」が論じられる。

特集の最後の**川崎**の解説では，先の 4 編とは趣を変えて，**VAR モデルによる時系列間の因果解析**を扱う。Vol. 3 の特集のコラムでは「時系列の因果関係は難物である」ことを述べた。VAR モデルによる解析では「非定常の変化は切り捨てて，定常的に繰り返し起きる揺らぎの部分の相関をみる」という立場をとるので，前処理での定常性の確保が重要になる。その上で「因果」については，時間的な前後関係を考慮して判断する（グレンジャー因果性，インパルス応答）。この解説では，実際のデータを用いて，こうした手順を懇切に説明している。

なお，この特集の外でも，植村による天文学の話題や小特集の一部で時系列データを扱っているので，あわせてご覧いただきたい。　（特集担当　伊庭幸人）

状態空間モデル

北川源四郎（明治大学）

状態空間モデルは多くの時系列モデルを統一的に取り扱うためのモデルで，時系列の予測，補間，成分分解やパラメータ推定など，時系列に関連した多くの問題を状態推定の問題として統一的に取り扱うことができる便利な枠組みである。この解説では，状態空間モデルとは何か，状態を逐次的に計算できるアルゴリズム，具体的な問題への応用を紹介する。

1　状態空間モデル

(1)　状態とは――時系列の予測と状態

気温や株価のように時間とともに変動する現象の記録を時系列と呼び，時刻 n の観測値を y_n と表すことにする。y_n は 1 変量でも，いくつかの変数を同時に観測した多変量でもよい。時系列 $Y_n=\{y_1, ..., y_n\}$ が与えられたとき，時刻 n までに得られた情報に基づいて将来の値 y_{n+k}, $k=1, 2, ...$ を推定する問題を予測という。

天気予報や経済予測が多くの人の関心事であるように，現実世界では予測は重要な問題である。通常のデータ解析と異なって，与えられたデータの範囲外のことを考えようとするところに，予測の問題のおもしろさと難しさがある。

予測の問題はモデリングにおいて，重要な課題を提起する。データを記述することだけが問題ならば，モデルのパラメータを増やせばいくらでもよいフィッティングを実現できる。しかし，そのような複雑すぎるモデルは予測においては，しばしば悲惨な結果をもたらす。一方，予測の役に立つモデルは，データの変動

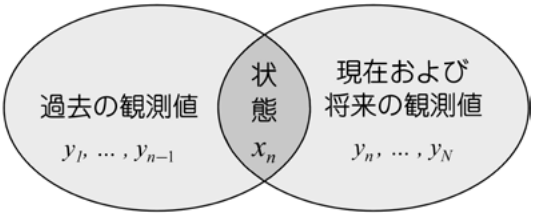

図 1―過去，現在，未来が張る空間の共通部分と状態

を外挿にも役に立つ本質的な変動と，一時的なランダムな変動に分離している。それによって，データに含まれる本質的な情報を取り出しているともいえる。

物理的なシステムでは，過去の情報は現在を経由して将来に伝えられる。したがって，過去の情報のうち将来の動きに関連する部分は何らかの形で現時点に集約されているものと考えることができる。この集約された情報のことを**状態**と呼ぶことにする(図 1)。言い換えれば，このシステムに関しては，過去のデータをすべて保存しておく必要はなく，この状態 x_n だけがあれば過去のデータは捨ててしまっても，時系列の将来の動きは予測できることになる。

(2) 状態空間モデル

次に，この状態を使って時系列がどのように表現できるかを考えてみよう。まず，状態 x_{n-1} を知っているとき，次の時刻 n の状態 x_n がどのように表現できるかを考えてみる。時系列の過去の変動を集約したものが状態 x_{n-1} なので，確定的な現象ならば適当な関数 $F(x)$ を用いて $x_n=F(x_{n-1})$ と表されるだろう。しかし，時系列の場合には不確定要因を伴うので，状態 x_{n-1} とシステムノイズと呼ばれる過去の変動には依存しない入力の関数として表現する必要がある。関数自体も時刻とともに変化する場合を考えると，状態 x_n は $x_n=F_n(x_{n-1})+G_n(v_n)$ と表現できる。このモデルのことを**システムモデル**と呼ぶことにする。

一方，状態 x_n が与えられたとき，時刻 n の時系列 y_n は，この状態の関数に観測時の外乱が加わったものと考えられるので，**観測モデル**として $y_n=H_n(x_n)+w_n$ が得られる。通常はこれらのモデルが使いやすいが，場合によっては，もっと一般的なモデル $x_n=F_n(x_{n-1}, v_n)$ や $y_n=H_n(x_n, w_n)$ を使うこともある。

次式のようにこの二つのモデルを統合したものを**状態空間モデル**と呼ぶ(図 2)。

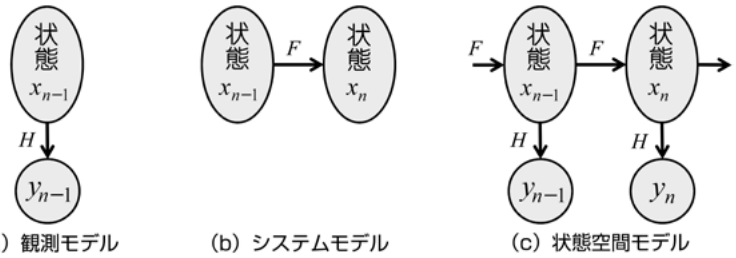

図 2—状態空間モデル。観測モデル(a)とシステムモデル(b)を連結したものが状態空間モデル(c)である。

$$x_n = F_n(x_{n-1}) + G_n(v_n) \quad \text{(システムモデル)} \tag{1}$$

$$y_n = H_n(x_n) + w_n \quad \text{(観測モデル)} \tag{2}$$

状態 x_n は k 次元，時系列 y_n は l 次元とすると，v_n と w_n はそれぞれ，確率密度関数 $q(v)$, $r(w)$ に従う m 次元と l 次元の白色雑音で，システムノイズ，観測ノイズと呼ばれる。また，F_n, G_n は m 次元，H_n は l 次元の非線形関数である。

状態空間モデル(1), (2)が与えられると，左辺の x_n および y_n の条件付き分布 $p(x_n|x_{n-1})$, $p(y_n|x_n)$ が定まる。観測モデル(2)から，$p(y_n|x_n) = r(y_n - H_n(x_n))$ となるので，観測ノイズの確率密度関数の w_n の代わりに $y_n - H(x_n)$ を代入すればよい。一方，システムモデルの場合は少し複雑になるが，$p(x_n|x_{n-1}) = \left|\frac{dG}{dv}\right|^{-1} q(G_n^{-1}(y_n - F_n(x_{n-1})))$ となる。実は，以下の節で示すフィルタや平滑化のアルゴリズムの公式の導出に用いるのは，これらの条件付き分布だけである。したがって，二つの条件付き分布の形を直接規定する

$$x_n \sim q_n(\cdot|x_{n-1}) \quad \text{(システムモデル)} \tag{3}$$

$$y_n \sim r_n(\cdot|x_n) \quad \text{(観測モデル)} \tag{4}$$

を状態空間モデルの一般形として用いることもできる([2]の 1060 頁)。

時系列解析で用いられるモデルの多くは，状態空間モデルの形で表現し，さまざまな形の統計的な問題を状態の推定に関する問題として統一的に取り扱うことができる。以下に，代表的な例をいくつか示しておく。

例 1——線形・ガウス型状態空間モデル　関数 F_n, G_n, H_n が線形関数で，それぞれ $F_n(x) = F_n x$, $G_n(v) = G_n v$, $H_n(x) = H_n x$ と表現でき，外乱 v_n と w_n はそれぞれ，m 次元および l 次元の正規分布に従い，$v_n \sim N(0, Q_n)$ と $w_n \sim N(0, R_n)$

となるとき，上記の状態空間モデルは，

$$x_n = F_n x_{n-1} + G_n v_n \quad \text{（システムモデル）} \tag{5}$$

$$y_n = H_n x_n + w_n \quad \text{（観測モデル）} \tag{6}$$

となり，線形・ガウス型状態空間モデルと呼ばれる。ただし，F_n, G_n, H_n はそれぞれ $k \times k, k \times m, l \times k$ の行列である。このモデルを単に状態空間モデルと呼ぶことが多い。

例 2——AR モデル　2 次の AR（自己回帰）モデル

$$y_n = a_1 y_{n-1} + a_2 y_{n-2} + v_n, \qquad v_n \sim N(0, \sigma^2) \tag{7}$$

は線形・ガウス型状態空間モデル

$$x_n = \begin{bmatrix} y_n \\ y_{n-1} \end{bmatrix}, \quad F = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$H = [1 \quad 0], \quad Q = \sigma^2, \quad R = 0 \tag{8}$$

によって表現できる。これは x_{n-1} と x_n の関係および x_n と y_n の関係を考えてみると，簡単に確かめることができる。また，より一般の自己回帰・移動平均（ARMA）モデルもこのような形の状態空間モデルを用いて表現できるが，状態ベクトル x_n は単に過去の時系列を並べただけではなく，一工夫が必要である[9]。

例 3——成分分解モデル　時系列 y_n を次のように J 個の成分 $x_n^{(j)}, j=1, ..., J$ と観測ノイズ w_n に分解するものとする。

$$y_n = x_n^{(1)} + \cdots + x_n^{(J)} + w_n \tag{9}$$

さらに，それぞれの成分 $x_n^{(j)}$ は状態空間モデルで

$$x_n^{(j)} = F^{(j)} x_{n-1} + G^{(j)} v_n \quad \text{（システムモデル）} \tag{10}$$

$$y_n = H^{(j)} x_n + w_n \quad \text{（観測モデル）} \tag{11}$$

と表現できるものとする。このとき，

$$F = \begin{bmatrix} F^{(1)} & & \\ & \ddots & \\ & & F^{(J)} \end{bmatrix}, \quad G = \begin{bmatrix} G^{(1)} & & \\ & \ddots & \\ & & G^{(J)} \end{bmatrix}, \quad x_n = \begin{bmatrix} x_n^{(1)} \\ \vdots \\ x_n^{(J)} \end{bmatrix}$$

$$H = [H^{(1)}, ..., H^{(J)}] \tag{12}$$

とおくと，時系列の変動を J 個の成分の和として表現する状態空間モデル

$$x_n = F x_{n-1} + G v_n$$

定常時系列と AR モデル

時系列は時間とともに不確実性を伴って変動するが，その平均，分散，共分散関数などの確率的構造自体は時間的に変化しない場合は定常時系列と呼ばれる。一方，定常性が成り立たないものが非定常時系列である。有限次元の線形定常時系列は ARMA モデルで表現できるが，実際の解析では簡潔で安定な AR モデルを用いることのほうが多い。

AR モデルを用いると時系列の予測分布を簡単に求めることができる。たとえば，時刻 $n-1$ までの時系列が与えられているとき，(7) を仮定すると y_n の予測分布は，平均 $a_1 y_{n-1}+a_n y_{n-2}$，分散 σ^2 の正規分布となる。2 期先以上の予測分布についても，繰り返し代入により計算できるが，複雑な式になるので，カルマンフィルタの一般式を用いるほうがよい。

時系列の変動特性を直感的に捉えるためには，周波数成分に分けて表現するパワースペクトルを考えるのが便利である。ただし，パワースペクトルを推定するために，時系列をフーリエ変換してペリオドグラムを求めると多数のピークが現れるため，どれが意味のあるピークであるかの判断が難しい。したがって，ペリオドグラムは平滑化が必要になるが，経験的な要素が多く，客観的な推定が難しい。一方，AR モデルの次数 m が適切に選択できると，周波数 f のパワースペクトルは

$$p(f)=\sigma^2\left|1-\sum_{j=1}^{m} a_j \exp\{-2\pi ijf\}\right|^{-2} \tag{13}$$

で自動的に求められる [9]。ただし，スペクトルが k 個のピークを持つためには AR 次数 m は $2k$ 以上が必要であることが知られている [9]。

$$y_n = Hx_n + w_n \tag{14}$$

が得られる。

逆にこのモデルを用いて，時系列を J 個の成分 $x_n^{(1)}, \ldots, x_n^{(J)}$ と観測ノイズ w_n に分解することができる。季節調整モデルはこの方法により，時系列を，トレンド，季節成分のほか，場合によっては曜日効果項や定常サイクル成分などに分解することができる。

例 4——時変係数モデル　自己回帰モデル

$$y_n = \sum_{j=1}^{m} a_j y_{n-j} + w_n, \qquad v_n \sim N(0, \sigma^2) \tag{15}$$

において自己回帰係数 a_j を時間とともに変化する係数 $a_{n,j}$ で置き換えた場合を考える。ここで係数変化のモデルとしてランダムウォークモデルを導入する。

$$a_{n,j} = a_{n-1,j} + v_{n,j}, \qquad v_{n,j} \sim N(0, \tau^2) \tag{16}$$

このとき，m 個の時変 AR 係数を縦に並べて状態ベクトルを

$$x_n = [a_{n,1}, \ldots, a_{n,m}]^T \tag{17}$$

と定義すると，自己回帰モデルと係数の変化を表すランダムウォークモデルは状態空間モデルを用いて

$$x_n = Fx_{n-1} + Gv_n \tag{18}$$

$$y_n = H_n x_n + w_n \tag{19}$$

と表現することができる。(16)と(17)からシステムモデル(18)が，また，(15)と(17)から観測モデル(19)が得られることは簡単に確かめられる。ただし，$w_n \sim N(0, \sigma^2)$, $v_n \sim N(0, \tau^2)$,

$$F = G = I_m, \qquad H_n = (y_{n-1}, \ldots, y_{n-m}) \tag{20}$$

である。ここで，H_n に時系列が入っていることを奇異に感じる読者もいると思うが，時刻 n の時点では $y_{n-1}, \ldots, y_{n-m}$ はすべて既知なので H_n に過去の時系列が含まれていても問題はない。

例 5——時変分散モデル　時系列 y_n は正規分布に従うが，その分散は時々刻々と変化し，$y_n = \varepsilon_n$, $\varepsilon_n \sim N(0, \sigma_n^2)$ となる場合を考える。ここで，さらに分散 σ_n^2 の時間変化が

$$\log \sigma_n^2 = \log \sigma_{n-1}^2 + v_n \tag{21}$$

と表されるものとする。このとき，状態を $x_n = \log \sigma_n^2$ と定義すると，時系列の状態空間モデルは

$$\begin{aligned} x_n &= x_{n-1} + v_n \\ y_n &= \exp(x_n/2) w_n \end{aligned} \tag{22}$$

と表される。ただし，w_n は標準正規分布に従うものとする。このように，時変分散モデルは(1), (2)式を拡張した一般型の非線形状態空間モデルにおいて

$F(x, v)=x+v, H(x, w)=\exp(x/2)w$ とおけばよい。

2　フィルタリング―状態の逐次推定

(1)　状態推定の問題

時系列の観測値 $Y_j=\{y_1, \ldots, y_n\}$ と状態空間モデルが与えられたとき，状態 x_n を推定する問題は，状態推定と呼ばれる。観測値の最終時点 n と推定する状態の時刻 j の大小関係により，以下のように 3 つの場合に分けられる(図 3)。

$n < j$ の場合　　予測

$n = j$ の場合　　フィルタ

$n > j$ の場合　　平滑化

状態空間モデルに関して，このような状態推定の問題を考えるのは，後に示すように時系列の予測，補間，パラメータ推定，成分分解などの現実の課題解決に必要な問題のほとんどがこの状態推定を利用することによって統一的に解決できるからである。リアルタイム処理が要求されるオンライン制御や逐次予測などは別にして，一般に情報抽出や推定においては，平滑化は注目する時点の前後の情報を用いるので，予測やフィルタよりも精度のよい推定を行うことができる。

状態推定においては，推定値だけでなく，少なくともその分散，一般には，分布を求める必要がある。たとえば，予測においては状態の予測値だけが注目されることが多いが，実際にはその分散または標準偏差もわからなければ合理的な意思決定には役に立たない。明日の気温の予測値が同じ 10 度であっても，その標準偏差が 1 度の場合と 10 度の場合では，判断は異なる可能性が高いからである。

パラメータ推定に使われる尤度も，予測値だけでは計算できない。正規分布を

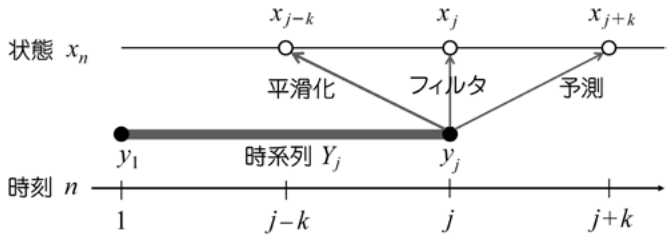

図 3―状態推定の 3 類型

仮定したモデルの場合でも，予測値とその分散，一般のモデルの場合には予測分布が必要である。

(2) 逐次フィルタと平滑化

このように，観測値 Y_j が与えられたとき，状態推定の問題は Y_j のもとで状態 x_n の条件付き分布 $p(x_n|Y_j)$ を求める問題となる。このとき，予測分布 $p(x_n|Y_{n-1})$ とフィルタ分布 $p(x_n|Y_n)$ は状態の定義から，$p(x_n|x_{n-1}, Y_{n-1})=p(x_n|x_{n-1}), p(y_n|x_n, Y_{n-1})=p(y_n|x_n)$ が成り立つので，以下の式により逐次的に計算することができる。

一期先予測

$$p(x_n|Y_{n-1}) = \int_{-\infty}^{\infty} p(x_n|x_{n-1})p(x_{n-1}|Y_{n-1})dx_{n-1} \tag{23}$$

フィルタ

$$p(x_n|Y_n) = \frac{p(y_n|x_n)p(x_n|Y_{n-1})}{p(y_n|Y_{n-1})} \tag{24}$$

ただし，$p(y_n|Y_{n-1})=\int p(y_n|x_n)p(x_n|Y_{n-1})dx_n$ である。

予測公式(23)の右辺の $p(x_n|x_{n-1})$ はシステムモデル(1)によって決まるものである。したがって，時刻 $n-1$ のフィルタ分布 $p(x_{n-1}|Y_{n-1})$ が与えられると，時刻 n の予測分布 $p(x_n|Y_{n-1})$ が計算できる。ただし，この分布を具体的にどうやって求めるかは重要な課題であり，その計算法についてはこの後，順次紹介する。

同様に，フィルタ公式(24)の右辺の $p(y_n|x_n)$ は観測モデル(2)によって決まるので，予測公式で求めた予測分布 $p(x_n|Y_{n-1})$ と最新の観測値 y_n から，フィルタ分布 $p(x_n|Y_n)$ を求めることができる。

図 4 に示すように，このフィルタ分布は，次のステップで予測分布 $p(x_{n+1}|Y_n)$ の計算に利用できる。このように予測とフィルタの操作を繰り返すことによって逐次フィルタが構成できる。

一方，平滑化のアルゴリズムとしては，固定区間平滑化，固定ラグ平滑化，固定点平滑化の 3 種類があるが，ここでは，逆方向の逐次計算によって平滑化分布 $p(x_n|Y_N)$ を計算できる固定区間平滑化のアルゴリズムを示しておく。

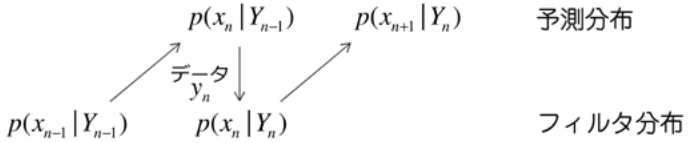

図 4―逐次フィルタのアルゴリズム。予測とフィルタの繰り返し計算により実現できる。

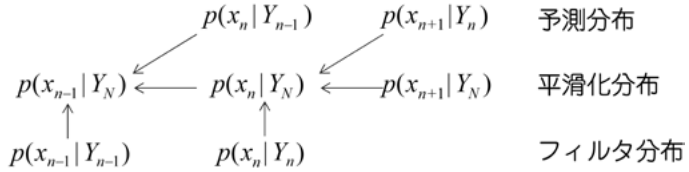

図 5―平滑化アルゴリズム。後ろ向きの逐次計算となる。

平滑化

$$p(x_n|Y_N) = p(x_n|Y_n)\int_{-\infty}^{\infty}\frac{p(x_{n+1}|Y_N)\,p(x_{n+1}|x_n)}{p(x_{n+1}|Y_n)}dx_{n+1} \tag{25}$$

この平滑化公式は，状態 x_n のフィルタ分布 $p(x_n|Y_n)$ と状態 x_{n+1} の予測分布 $p(x_{n+1}|Y_n)$ および平滑化分布 $p(x_{n+1}|Y_N)$ が得られると，状態 x_n の平滑化分布 $p(x_n|Y_N)$ が計算できることを示している。なお，データの終端では平滑化分布 $p(x_N|Y_N)$ は，フィルタ分布と一致することに注意されたい。

したがって，逐次フィルタにより予測分布とフィルタ分布をデータの終端まですべて求めておくと，図 5 に示すように，時間軸と逆方向に平滑化分布を計算することができる。

(3) 線形・ガウス型状態空間モデルとカルマンフィルタ

(1)と(2)の状態空間モデルにおいて，関数 F_n と H_n が線形で，白色雑音 v_n および w_n が正規分布に従う場合，線形・ガウス型状態空間モデルと呼ばれる。以下では

$$x_n = F_n x_{n-1} + G_n v_n, \tag{26}$$

$$y_n = H_n x_n + w_n \tag{27}$$

を考えることにする。ただし，初期状態 x_0 の分布が $p(x_0|Y_0)$ で与えられるものとする。ここで，時刻 0 には時系列の観測値は得られていないので，$Y_0=\phi$ は