

ハミルトン「時系列解析」の前半から後半へ

目次

参考文献	1
ベイズ推測へ	3
カルマンフィルタ——みえないものを追いかける手続き	4

参考文献

- [1] James D. Hamilton. Time Series Analysis. Princeton University Press, 1994.
- [2] 沖本竜義. 経済・ファイナンスデータの計量時系列分析. 朝倉書店, 2010.
- [3] 渡辺澄夫. ベイズ統計の理論と方法. コロナ社, 2012.



ハミルトン「時系列解析」[1]の前半をふりかえってみます。

- 1章では差分方程式を導入しました，そもそも時系列というのは各時点の値がその時点より過去の値に左右されているような値の列ですね。「現時点の値」を「過去何ステップかの値」の式で表す差分方程式はそれを最も素朴に表現したモデルといえるでしょう——もっとも差分方程式といったときにはノイズのない決定論的なモデルになりますが，システムが安定（任意の時刻で値が有界）である条件も導出しましたね。つまり， p 次の差分方程式が安定である条件は，それを 1 次のベクトル差分方程式で表現したときの係数行列の全ての固有値の絶対値が 1 未満であることでした。
- 2章ではラグ演算子——いわば，入力された時系列の 1 ステップ昔を覗くメガネ——を導入しました。私たちはラグ演算子によって「現在の値」を「一番最初の値」の式でかくことが可能になりました。この形式もまた時系列の性質を知るのに有用です。1 章とは別のアプローチで同じ安定条件を導出しましたね。
- 3章では定常性，エルゴード性を導入し，ARMA 過程を導入しました。
- 4章では ARMA 過程の将来の値を予測しました。「この時系列の値は将来どうなるの？」というのは私たちが目の前の時系列に対して最も興味があることのひとつでしょう。この章ではいよいよその間に，過去の値を全て観測している理想的な場合，過去 m ステップしか観測できていない場合の順で取り組みました。何次のモデルを仮定すべきかに関連して，Box-Jenkins 法にも触れましたね。

- 5 章では観測値から ARMA 過程そのものを推定する方法を扱いました。正規ホワイトノイズを仮定して尤度を最大化するという手続きによりました。
- 6 章ではそれまで自己共分散関数によって時系列を特徴付けていたのとうってかわって、母集団スペクトルによって時系列を特徴付ける流儀を扱いましたね。もっとも両者はどちらかが特定されていればもう片方にかき換えられるわけですが。
- 7 章では漸近分布論を扱いました。これは標本サイズが大きくなったときに推測や検定の結果がが想定通りに正しくなっていくかを調べる時に必要な概念ですね。
- 8 章は既に扱っていた線形回帰モデルの最小 2 乗推定の性質を掘り下げました。まずオーソドックスな最小 2 乗法を扱い、一般化最小 2 乗法を扱いましたね。
- 10 章では単変量 ARMA 過程を多変量 ARMA 過程に拡張しました。
- 時系列が多変量になったとき私たちは「**この変数とこの変数は関係しているの？**」という興味を抱くでしょう。11 章では多変量 ARMA 過程を用いてそのような問いに答えていきました。グレンジャー因果検定やインパルス応答関数を取り扱いましたね。



——こうしてみると、前半を終えただけでも私たちは目の前の時系列に対して既に色々なことが推測できるように思えます。将来の値を予測したり、変数間にグレンジャー因果があるかを推測したり、あるとしたらその効果の大きさはどれくらいなのか推測したり——ただし往々にしてノイズが正規ホワイトノイズであるような ARMA 過程を仮定することにはなりますが。

となると、12 章からは何を扱うのでしょうか？やはり仮定を緩和していくとか？



それも後の章では出てくるけど、12 章でまずやることは違うかな。12 章でやることは、12 章の冒頭から天下りの的にいえば、「**推測に事前知識を反映させること**」みたいだからね。



事前知識？

12 章ベイズ推測へ



そうだな，12 章冒頭の例だけど，任意の時刻 t に $y_t \sim \text{i.i.d } N(\mu, \sigma^2)$ が観測されるとするよ． $y = (y_1, \dots, y_T)^\top$ を観測したときに $\theta = (\mu, \sigma^2)^\top$ をどのように推測すべきだろうか？



そのモデルはもはや各時刻の値が独立なので全く時系列らしくないですが……しかし，そのような推測は 5 章で扱いましたよ．そのモデルの下で時刻 $1, \dots, T$ にどんな値が観測されるかは確率ベクトル $Y = (Y_1, \dots, Y_T)^\top$ です． Y の確率密度関数は $\theta = (\mu, \sigma^2)^\top$ を用いてかけます．実際には $y = (y_1, \dots, y_T)^\top$ が観測されているのでこの点で Y の確率密度関数 (θ の関数と考えると尤度関数) が最大になってほしいです．そのような $\hat{\theta}$ を選びます…というものでしたよね．最尤推定です．



そうだね．その選び方がどれくらいよいかは，たくさんのパラレルワールドで y を観測してそれぞれで $\hat{\theta}$ を推定したときに，平均的にどれくらい真の θ とずれるか $E(\hat{\theta} - \theta)(\hat{\theta} - \theta)^\top$ で評価されるね．

じゃあ少し違う状況を考えるよ．いやモデルは同じなんだけど，観測者は事前には θ に対してある程度の知識というか信念をもっていたとする．「 μ がゼロ未満であることはほとんどないと思う」といった具合に．



なんと，事前知識が……いえでも，現実にはそのように事前にある程度の見通しがある状況もありそうですね．しかし，最尤推定ではそのような知識を取り込むことは難しそうですね．尤度関数の最大点を取るだけですから……「 μ がゼロ未満であることはほとんどないと思う」と考えているなら， μ がゼロ未満である場所で尤度関数を薄くしておく必要があります．しかし，どのように薄くすればよいのやら．それに，そうやって尤度関数の最大点が少しずれたとして，その点をどれくらい信用してよいのか……．



うん，もはや，「観測したデータをもとにただ 1 点の推測値を出す」という枠組みだと扱いづらくなってくるよね．だから，ここでは θ も確率変数であると考えよ． θ に対して事前想定している確率分布があって，観測したデータをも

とにその確率分布を更新する，という手続きをとる．



θ も確率変数であると考える？……これまでは神様がモデルというサイコロを振って観測値が出てくると考えていましたが，そのモデル (のパラメータ)すらもサイコロを振った結果で決まっていたということなのですか？では，モデルから観測値を出す神様は実は下請けの神様にパラメータを発注していて，下請けの神様がサイコロを振ってパラメータを出して納品していたと……神様界にもそんな下請け構造が……．



ちょっと意味がわからないかな．

ベイズ推測 [3]

ベイズ推測とは「真の確率分布 $q(x)$ は，およそ $p^*(x)$ であろう」と推測することである．

13 章カルマンフィルタ——みえないものを追いかける手続き



12 章でベイズ推測を導入して，13 章はカルマンフィルタ？ 何ですかそれは？



カルマンフィルタとは「各時刻の観測値が，観測できない『状態』から生成されるというモデルを仮定した下で，状態の分布をトラッキングする手続き」かな．仮定するモデルについてもっとちゃんと説明するよ．**状態空間モデル**というんだけど，各時刻の観測値 $y_t \in \mathbb{R}^n$ は以下のモデルから生成されると考えるよ．

$$\xi_{t+1} = F\xi_t + v_{t+1} \quad (13.1.1)$$

$$y_t = A^\top x_t + H^\top \xi_t + w_t \quad (13.1.2)$$

上式のうち $v_t \in \mathbb{R}^r$ と $w_t \in \mathbb{R}^n$ はベクトルホワイトノイズだ．ノイズ以外の部分は……2 つ目の (13.1.2) 式から説明すると，「 y_t は x_t の線形変換と ξ_t の線形変換の和からできる」という式だよね．このうち $x_t \in \mathbb{R}^k$ は予め決定している外生変数だ．例えば y_t が日ごとのデータだとしたら x_t は曜日を one-hot エンコ

ーディングしたベクトルとかね。この x_t はない教科書の方が多いと思うんだよね。曜日ごとの周期性とかが必要な場合でも ξ_t に押し付けることができると思うし。

それで肝心なのは $xi_t \in \mathbb{R}^r$ の方だ。これが観測できない「状態」。観測値を生み出す特徴ベクトルといった方がわかりやすいかな？ この ξ_t が時間発展すると考えるよ。1つ目の(13.1.1)式がそれだね。初期状態 ξ_0 は適当に決め打っておくことになるよ。



観測できない「状態」が観測値を生み出す……そうですね、例えば私が毎日動物園にパンダさんのようすをみにいくとします。私に観測できるのは、パンダさんが遊び回っているか、そっぽを向いているかのどちらくらいでしょう。しかし、パンダさんのそのようすの裏には「今日はいい気分だ」「今日は疲れがたまっている」「今日はいらいらしている」「今日は寂しい」などのもっと複雑な「状態」があるに違いありません。そのようなパンダさんの状態の分布の推移をこそモデリングしようというのが「状態空間モデル」なのですね？



う、うん……パンダの毎日のふるまいを分析する用途に適切かどうかは別として、状態空間モデルの考え方はそれで合っているよ。まあそれで、重要なのは、状態空間モデルは以下が満たされるとき状態 ξ_t を完全に解析的にトラッキングできるということなんだよね。そのときにトラッキングする手続きをカルマンフィルタとよぶよ。

- 初期状態 ξ_0 がガウス分布にしたがう。
- ベクトルホワイトノイズ v_t, w_t がガウス分布にしたがう。
- 変換 F, H が線形変換である(教科書の書き方では転置をとっているし最初から行列であることが前提になっているけどね)。