

ベイズアンモデルを用いた時系列予測合成

慶應義塾大学経済学部四年

中妻照雄研究会

渡邊 直斗

Abstract

近年、ベイズ的手法を用いてのモデルや予測の比較,測定や結合に関する研究が行われている. その中で多くの, 予測密度のプール化の手法が考案されてきた.

Contents

1	イントロダクション	3
2	手法	3
2.1	Agent Opinion Analysisとその拡張モデル	3
2.1.1	Agent Opinion Analysis	3
2.1.2	キャリブレーション関数の設計	4
2.2	動的連続設計	4
2.2.1	Agent Opinion Analysisの時系列設定	4
2.2.2	Dynamic Linear Models	5

1 イントロダクション

2 手法

今回の論文では、Agent Opinion Analysisと呼ばれる統計的手法を動的に使えるように拡張した手法(McAlinn and West)を用いることによって分析を行う。

2.1 Agent Opinion Analysisとその拡張モデル

ここでは、基本的なAgent Opinion Analysis(West 1992)とその拡張モデルについてまとめる。この段階では、時系列解析には対応していない。

2.1.1 Agent Opinion Analysis

ベイジアン意思決定者 \mathcal{D} は J のエージェント $\mathcal{A}_j, (j = 1 : J)$ が提供する情報をもとに結果 y を予測する(ここでのエージェントは、モデル、予測者などを意味する)。まず、 \mathcal{D} は事前確率 $p(y)$ を持ち、エージェントは予測情報である $h_j(y)$ を提供する。提供された情報セットを $\mathcal{H} = \{h_1(\cdot), \dots, h_J(\cdot)\}$ とすると、 \mathcal{D} は自身の事後確率 $p(y|\mathcal{H})$ を用いて y を予測するようになる。これがAgent Opinion Analysisになる。

しかし、各々のバイアスに加え、エージェント間の相互依存などが存在するため、簡単に $p(y, \mathcal{H}) = p(y)p(\mathcal{H}|y)$ を概念化できない。そこで、概念化のための関連するAgent Opinion Analysis理論(Gunest and Schervish 1985; West and Crosse 1992)についてまとめる。この理論では、 \mathcal{D} は情報セット \mathcal{H} を観測する前に $p(y)$ を指定すると仮定し、その期待値を $m(\mathbf{x}) = E[\prod_{j=1:J} h_j(\mathbf{x}_j)]$ とする(ここで $\mathbf{x} = (x_1, \dots, x_J)'$ はエージェントのダミー変数ベクトル)。すると、以下のベイジアンモデル $p(y, \mathcal{H})$ の部分集合が導かれる。

$$p(y|\mathcal{H}) = \int_{\mathbf{x}} \alpha(y|\mathbf{x}) \prod_{j=1:J} h_j(x_j) d\mathbf{x} \quad (1)$$

ここで、 $\mathbf{x} = (x_1, \dots, x_J)'$ は J 次元状態ベクトルであり、 $\alpha(y|\mathbf{x})$ は \mathbf{x} が与えられたときの y の条件付き確率である。

このように表示したことによって、次の3つのように解釈された状態変数 \mathbf{x} の存在がわかる。まず、エージェント j による予測分布は、その状態変数 x と同様であること。2つ目は状態変数は、各エージェントのモデル $h_j(\cdot)$ から独立に生成されること。3つ目は状態変数は、結果 y に対して条件付き分布 $\alpha(y|\mathbf{x})$ を通して関連していることである。つまり、 x_j を状態変数、 $\alpha(y|\mathbf{x})$ を意思決定者 \mathcal{D} のキャリブレーション関数であると考えられる。

この理論によって、問題であった $p(y, \mathcal{H})$ について考える必要がなくなった。また、 $\alpha(y|\mathbf{x})$ の関数の形を特定していないので、様々な設計を行うことによりモデルの拡張が可能になったことがわかる。

2.1.2 キャリブレーション関数の設計

さらに、(Aastveit et al. 2015)の例のように、予測 $\pi_0(y)$ を状態依存変数 $a_{0:J}(\mathbf{x})$ を用いて行おうとし、 $\alpha(y|\mathbf{x}) = a_0(\mathbf{x})\pi_0(y) + \sum_{1:J} a_j(\mathbf{x})\delta_{x_j}(y)$ とおけば、キャリブレーションを明示的に示した(1)に書き換えられる。それは、以下の式になる。

$$p(y|\mathcal{H}) = q_0(y|\mathcal{H})\pi_0(y) + \sum_{1:J} q_j(y|\mathcal{H})h_j(y) \quad (2)$$

ここで、重み $q_j(\mathcal{H}) = \int a_j(\mathbf{x})\delta_{x_j}(y) \prod_j h_j(x_j)d\mathbf{x}$ は、 \mathcal{H} に依存して、キャリブレーションできるように動く。

これをもとに、(McAlinn and West)より次のように書き換えられる。

$$\alpha(y|\mathbf{x}) = N(y|\mathbf{F}'\boldsymbol{\theta}, v) \quad \text{with} \quad \mathbf{F} = (1, \mathbf{x}')' \quad \text{and} \quad \boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_J)' \quad (3)$$

ここで、 $\theta_0 = f - \sum_{j=1:J} \theta_j m_j$ である。この定義により、時系列分析への拡張のスタート地点に立つことができた。

2.2 動的連続設計

ここで、本題である時系列の連続的な分析に対応できるようにする。

2.2.1 Agent Opinion Analysisの時系列設定

意思決定者 \mathcal{D} は、各時系列 $t = 1, 2, \dots$ において、エージェントから予測密度を受け取り、連続的に y_t を予測している。例えば、 $t-1$ の時、意思決定者 \mathcal{D} は、エージェントから、予測密度のデータセット $\mathcal{H} = \{h_{t1}(y_t), \dots, h_{tJ}(y_t)\}$ を受け取り、 y_t を予測しようとする。意思決定者が、使う情報をまとめると $\{y_{1:t-1}, \mathcal{H}_{1:t}\}$ と表せる。前述のモデルを動的に対応できるように拡張すると、意思決定者 \mathcal{D} が、 y_t を予測するための $t-1$ 時点での分布は以下のような形になる。

$$p(y_t|\Phi_t, y_{1:t-1}, \mathcal{H}_{1:t}) \equiv p(y_t|\Phi_t, \mathcal{H}_t) = \int \alpha_t(y_t|\mathbf{x}_t, \Phi_t) \prod_{j=1:J} h_{tj}(x_{tj})dx_{tj} \quad (4)$$

ここで、 $\mathbf{x}_t = x_{t,1:J}$ は、時間 t におけるエージェント \mathcal{J} の状態ベクトルで、 $\alpha_t(y_t|\mathbf{x}_t, \Phi_t)$ は意思決定者 \mathcal{D} が、 y_t を予測するため \mathbf{x}_t を与えられた時の、条件付きキャリブレーション

の確率密度関数である。そして Φ_t は、 \mathcal{D} の $t-1$ 時点での事後分布 $p(\Phi_t|y_{1:t-1}, \mathcal{H}_{1:t-1})$ のパラメーターをキャリブレーションする確率密度関数を定義する時変パラメーターを表す。

次に考えることは、様々な形を仮定できる $\alpha_t(y_t|x_t, \Phi_t)$ とそれを定義する動的状態変数 Φ_t をどう扱うかである。そこで、考えられるのは扱いやすい動的線形回帰モデルや、DLMs(West and Harrison 1997; Prado and West 2010)である。

2.2.2 Dynamic Linear Models

動的回帰のために、(3)で示したキャリブレーション関数を時系列設定に拡張することを考える。(3)は次のように書き換えられる。

$$\alpha_t(y_t|x_t, \Phi_t) = N(y_t|\mathbf{F}_t'\boldsymbol{\theta}_t, v_t) \quad \text{with} \quad \mathbf{F}_t = (1, x_t')' \quad \text{and} \quad \boldsymbol{\theta}_t = (\theta_{t0}, \theta_{t1}, \dots, \theta_{tJ})' \quad (5)$$

この際、残差変動を定義する条件付き分散 v_t が加わる。つまり、各時間 t における Φ_t は今 $\Phi_t = (\boldsymbol{\theta}_t, v_t)$ となっている。この設計によって、Dynamic Linear Model(DLM)(West and Harrison 1997)の一般的な活用形を定義することができる。以下のような形になる。

$$\begin{aligned} y_t &= \mathbf{F}_t'\boldsymbol{\theta}_t + \nu_t, \nu_t \sim N(0, v_t), \\ \boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \omega_t, \omega_t \sim N(0, v_t \mathbf{W}_t) \end{aligned} \quad (6)$$

今回の論文では、パラメータ推定はL-BFGS-B法を用いることによって計算する。