拡散過程の適合度検定

統計数理研究所　西山陽一

November 26, 2007

X は、ある初期値 X_0 を出発し、確率微分方程式

$$\mathrm{d}X_t = S(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad t > 0,$$

を満たすような 1 次元のエルゴード的拡散過程であるとする。我々は連続観測 $X^T = \{X_t : 0 \le t \le T\}$ に基づき未知のドリフト係数 S の適合度検定問題を考える。

I.I.D. データに対する適合度検定問題は古くから知られている。Kolmogorov-Smirnov 検定統計量は漸近的に distribution free であることが証明されている。極限は標準ブラウン橋の sup の分布である。ところが、拡散過程モデルにおいては、同様の結果は出てこない。すなわち、不変分布を F と表し、その自然な推定量

$$\widehat{F}_T(x) = \frac{1}{T}\int_0^T 1_{(-\infty,x]}(X_t)dt$$

を導入しても、

$$\sup_x \sqrt{T}|\widehat{F}_T(x) - F(x)|$$

は漸近的に distribution free にならない。実際、極限の確率過程の共分散は

$$E\eta(x)\eta(y) = 4E\left(\frac{[F(\xi\wedge x) - F(\xi)F(x)][F(\xi\wedge y) - F(\xi)F(y)]}{f(\xi)^2}\right)$$

となる。ただし ξ は不変分布に従う確率変数である。Kutoyants (2004) は彼の大著の最後の節で、この事実に触れ、さらなる研究が必要であると結んでいる。

我々は、重み付き経験過程

$$V_T(x) = \frac{1}{\sqrt{T}}\int_0^T \mathbf{1}_{(-\infty,x]}(X_t)\frac{1}{\sigma(X_t)}(\mathrm{d}X_t - S_0(X_t)\mathrm{d}t)$$

を導入し、統計量

$$\sup_x |V_T(x)|$$

の漸近挙動を考察する。結論は

- 帰無仮説のもとで漸近的に distribution free
- 対立仮説のもとで一致性をもつ

ということである。

証明の概略を述べる。まず、帰無仮説のもとで

$$V_T(x) = \frac{1}{\sqrt{T}}\int_0^T \mathbf{1}_{(-\infty,x]}(X_t)dW_t$$

であることに注意する。そこで、一般論として

$$M_T(\psi) = \frac{1}{\sqrt{T}} \int_0^T \psi(X_t) dW_t$$

という形の確率場 $\psi \rightsquigarrow M_T(\psi)$ に対する極限定理を用意し、それを使って帰無仮説のもとでの漸近分布を導出する。

極限定理自体、非常にシンプルな形をしているので、ここできちんと述べておく。

Theorem 1. *X, W は同一の確率基の上で定義されており、X はスピード測度 m が有限であるようなエルゴード的拡散過程であるとし、W は標準 Wiener 過程であるとする。可算個の関数族 $\Psi \subset \mathcal{L}^2(\mathbf{R}, m)$ が*

$$\int_0^1 \sqrt{\log N(\epsilon; \Psi, || \cdot ||_{\mathcal{L}^2(m)})} d\epsilon < \infty$$

を満たすとする。このとき、

$$M_T(\psi) = \frac{1}{\sqrt{T}} \int_0^T \psi(X_t) dW_t$$

は平均ゼロで共分散が

$$E\Gamma(\psi)\Gamma(\varphi) = \int_{\mathbf{R}} \psi(z)\varphi(z) \frac{m(dz)}{m(\mathbf{R})}$$

であるようなガウス確率場 $\psi \rightsquigarrow \Gamma(\psi)$ に $\ell^\infty(\Psi)$ 空間の中で分布収束する。

この定理の証明は、Nishiyama (2000) による連続マルチンゲールに対する緊密性判定条件と、Van der Vaart and Van Zanten (2005) による局所時間の一様有界性に関する結果を組み合わせることによる。より正確には、Nishiyama (1999) の提唱した quadratic modulus というランダムな量が確率有界であることを、局所時間のテクニックを用いることによって示す。

対立仮説のもとでの検定の一致性は

$$\begin{aligned} V_T(x) &= \sqrt{T} \cdot \frac{1}{T} \int_0^T \mathbf{1}_{(-\infty, x]}(X_t) \frac{1}{\sigma(X_t}) (S_0(X_t) - S_1(X_t)) dt \\ &\quad + \frac{1}{\sqrt{T}} \int_0^T \mathbf{1}_{(-\infty, x]}(X_t) dW_t \end{aligned}$$

という関係式から得られる。

References

Kutoyants, Y.A. (2004), *Statistical Inference for Ergodic Diffusion Processes*, Springer, New York.

Negri, I., Nishiyama, Y. (2006), Goodness of fit test for ergodic diffusion processes, *Research Memorandum*, **1019**, The Institute of Statistical Mathematics, Tokyo.

Nishiyama, Y. (1999), A maximal inequality for continuous martingales and M-estimators in a Gaussian white noise model, *The Annals of Statistics*, **27**, 2, 675-696.

Nishiyama, Y. (2000), *Entropy Methods for Martingales*, CWI Tract **128**, Centrum voor Wiskunde en Informatica, Amsterdam.

Van der Vaart, A.W., Van Zanten, H. (2005), Donsker theorems for diffusion: necessary and sufficient conditions, *The Annals of Probability*, **33**, No. 4, 1422-1451.