

揺動散逸定理に基づく定常性の検定

大濱巖 積山功 草場康司 柳川堯
九州大学大学院数理学研究科

1 序

時系列解析では多くの場合、モデルに定常性を仮定する。したがって実際にデータを解析するさい、そのデータが定常モデルに当てはまっているかどうか重要になってくる。以下では与えられたデータが定常確率過程の実現値であることを検定する方法を述べる。

2 基本理論

次の帰無仮説に対する検定を構成する。

$H_0(S) : d$ 次元時系列 $\{X_j\}_{j=1}^N$ は定常である。

Def. $\{X_j\}$ が定常 $\stackrel{\text{def}}{\iff} E[X_j], \text{Cov}(X_j, X_{j+h})$ ($1-j \leq h \leq N-h$) が j に依存しない。

以下、 $E[X_j] = 0$ と仮定する。 $H_0(S)$ の下で 1 Step Predictor :

$$\widehat{X}_j = (\widehat{X}_j^{(1)}, \dots, \widehat{X}_j^{(d)})' \quad \text{s.t.}$$

$$\widehat{X}_j^{(k)} = \begin{cases} 0 & (j=1) \\ P_{M_1^{j-1}(X)} X_j^{(k)} & (j \geq 2) \end{cases}$$

ここに $M_1^{j-1}(X) = \{ \sum_{i=1}^{j-1} \sum_{k=1}^d a_{ik} X_i^{(k)} \mid \forall a_{ik} \in \mathbb{R} \}$
P : 射影作用素

は次の性質を持つ。

(i) $E[X_j - \widehat{X}_j] = 0$

(ii) $\text{Cov}[(X_j - \widehat{X}_j), (X_i - \widehat{X}_i)] = 0$ if $i \neq j$

今、 $V_{j-1} = \text{Var}[X_j - \widehat{X}_j]$ とおくと

$$\exists W_j : \text{正則行列} \quad \text{s.t.} \quad W_j W_j' = V_j$$

よって $\xi_j = W_{j-1}^{-1}(X_j - \widehat{X}_j)$ とすると

$$\{\xi_j\}_{j=1}^N \sim WN(0, I_d) \quad I_d : d \times d \text{ 単位行列}$$

$$\text{Def. } \{Z_j\}_{j=1}^n \sim WN(0, V) \quad \stackrel{dej}{\iff} \quad E[Z_j] = 0, \text{Cov}[Z_i, Z_j] = \delta_{i,j}V$$

したがって、 $\xi_j = (\xi_j^{(1)}, \dots, \xi_j^{(d)})'$ とすると

$$\{\xi_l\}_{l=1}^{dN} = \{\xi_j^{(k)}\}_{k=1, \dots, d : j=1, \dots, N} = \{\xi_1^{(1)}, \dots, \xi_1^{(d)}, \dots, \xi_N^{(1)}, \dots, \xi_N^{(d)}\} \sim WN(0, 1)$$

岡部[1]より

$$\{X_j\}_{j=1}^N \text{が定常} \iff \{\xi_l\}_{l=1}^{dN} \sim WN(0, 1)$$

したがってこの検定は、帰無仮説

$$H_0(WN) : \{\xi_l\}_{l=1}^{dN} \sim WN(0, 1)$$

に対する検定と同値になる。

3 データの分割

$R(h) = \text{Cov}(X_j, X_{j+h})$ の推定量として、

$$\hat{R}(h) = N^{-1} \sum_{j=1}^{N-h} (X_j - \bar{X})(X_{j+h} - \bar{X})'$$

を用いる。ここに $\bar{X} = N^{-1} \sum_{j=1}^N X_j$ である。今、信頼できる $\hat{R}(h)$ は岡部[1]より

$$0 \leq h \leq M = [3\sqrt{N}/d] - 1, \quad [\cdot]: \text{ガウス記号}$$

この $\hat{R}(h)$ ($0 \leq h \leq M$) を真の値とみなし、Blockwell-Davis [2] の Innovation Algorithm を用いて 1 Step Predictor を求める。しかし $\hat{R}(h)$ ($0 \leq h \leq M$) からは \widehat{X}_j ($1 \leq j \leq M+1$) しか求められないので、データ $\{X_1, X_2, \dots, X_N\}$ を

$$\begin{aligned} X[1] &= \{X_1, \dots, X_{M+1}\}, X[2] = \{X_2, \dots, X_{M+2}\}, \dots \\ \dots, X[i] &= \{X_i, \dots, X_{M+i}\}, \dots, X[N-M] = \{X_{N-M}, \dots, X_N\} \end{aligned}$$

と分割する。

この分割した各データ $X[i]$ に対し、Innovation Algorithm を用いて $\hat{R}(h)$ ($0 \leq h \leq M$) の値から $\xi[i] = \{\xi_{i,1}, \dots, \xi_{i,M+1}\}$ を計算する。ここで $\xi_{i,j} = (\xi_{i,j}^{(1)}, \dots, \xi_{i,j}^{(d)})'$ とすれば、2. で述べたように、

$$X[i] \text{が定常である。} \iff \xi[i] \sim WN(0, 1)$$

$$\begin{aligned} \text{ここに } \xi[i] &= \{\xi_{[i]j}\}_{j=1}^a = \{\xi_{i,j}^{(k)}\}_{k=1, \dots, d : j=1, \dots, M+1} \\ a &= d(M+1) \end{aligned}$$

ところで、 $\{X_j\}_{j=1}^N$ が定常ならば $\{X_j\}_{j=l}^m$ ($1 \leq l < m \leq N$) も定常であるから、 $(1 \leq i \leq N - M)$ に対し帰無仮説 $H_0(S)$ は帰無仮説

$$H_0^{(i)}(S) : X[i] = \{X_j\}_{j=i}^{M+i} \text{ は定常である。}$$

を imply する。さらに岡部[1]より $H_0^{(i)}(S)$ は帰無仮説

$$H_0^{(i)}(WN) : \xi[i] = \{\xi_{[i]j}\}_{j=1}^a \sim WN(0, 1)$$

と同値である。

4 検定の構成

有意水準 α で $H_0(S)$ を検定する。3. で述べたように $H_0(S)$ は $H_0^{(i)}(S)$ を imply し、 $H_0^{(i)}(S)$ は $H_0^{(i)}(WN)$ と同値であったから、各 i に対し $H_0^{(i)}(WN)$ を検定することで、 $H_0(S)$ に対する検定とする。さらにここでは $H_0^{(i)}(WN)$ そのものに対する検定ではなく、 $H_0^{(i)}(WN)$ を強くした 帰無仮説

$$H_0^{(i)} : \xi[i] = \{\xi_{[i]j}\}_{j=1}^a \sim IID(0, 1)$$

に対する検定を構成する。

Def. $\{X_j\}_{j=1}^n \sim IID(0, 1) \stackrel{def}{\iff}$ 任意の j に対し X_j は独立に平均0、分散1の同一分布に従う。

THEOREM $\mu_4 = E[\xi_{[i]j}^4] < \infty$ のとき $H_0^{(i)}$ の下で

$$Z^{(i)}_\alpha = \sqrt{a} \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 - 1 \\ \hat{\gamma}_\epsilon(1) \\ \vdots \\ \hat{\gamma}_\epsilon(L) \end{pmatrix} \xrightarrow{L} N(0, \Sigma) \quad \text{as } N \rightarrow \infty$$

ここに

$$\hat{\mu}_k = \hat{\mu}_{[i]k} = a^{-1} \sum_{j=1}^a \xi_{[i]j}^k \quad (k = 1, 2)$$

$$\hat{\gamma}_\epsilon(h) = \hat{\gamma}_{\xi_{[i]}}(h) = a^{-1} \sum_{j=1}^{a-h} \xi_{[i]j} \xi_{[i]j+h}$$

$$L = [2\sqrt{a}] - 1$$

$$\Sigma = \begin{pmatrix} 1 & \mu_3 & 0' \\ \mu_3 & \mu_4 - 1 & 0' \\ 0 & 0 & I_L \end{pmatrix}, \quad \mu_3 = E[\xi_{[i]j}^3]$$

[証明の概略]

- $\hat{\gamma}_\epsilon(h)$ を $\tilde{\gamma}_\epsilon(h) = a^{-1} \sum_{j=1}^a \xi_{[i]j} \xi_{[i]j+h}$ で置き換えた $Y^{(i)}_a$ が、 $Y^{(i)}_a \xrightarrow{L} N(0, \Sigma)$ を満たす。
 - Cramér-Wold 分解
 - m 従属強定常過程に関する中心極限定理
- $Z^{(i)}_a - Y^{(i)}_a \xrightarrow{P} 0$ as $N \rightarrow \infty$

COROLLARY $\mu_4 < \infty, \mu_4 - 1 - \mu_3^2 \neq 0$ のとき、 $H_0^{(i)}$ の下で

$$T_{(i)} = Z^{(i)'}_a \hat{\Sigma}_a^{(i)-1} Z^{(i)}_a \xrightarrow{L} \chi_{L+2}^2 \text{ as } N \rightarrow \infty$$

$$\text{ここに } \hat{\Sigma}_a^{(i)} = \begin{pmatrix} 1 & \hat{\mu}_3 & 0' \\ \hat{\mu}_3 & \hat{\mu}_4 - 2\hat{\mu}_2 + 1 & 0' \\ 0 & 0 & I_L \end{pmatrix}$$

$$\hat{\mu}_k = \hat{\mu}_{[i]k} = a^{-1} \sum_{j=1}^a \xi_{[i]j}^k \quad (k = 3, 4)$$

したがって、帰無仮説 $H_0^{(i)'}$ を

$$H_0^{(i)'} : \mu_4 < \infty, \mu_4 - 1 - \mu_3^2 \neq 0 \text{ かつ } H_0^{(i)}$$

として、この Corollary から

$$\Pr(T_{(i)} \geq C \mid H_0^{(i)'}) = \alpha / (N - M)$$

なる C を自由度 $L + 2$ のカイ二乗分布表より求め、

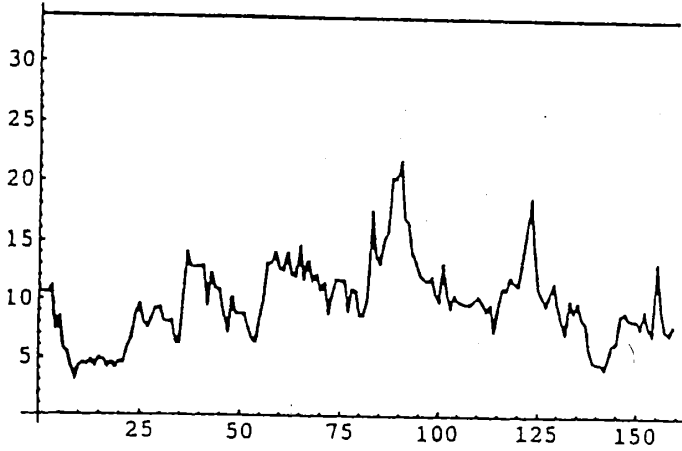
$$T_{(i)} \geq C \implies H_0^{(i)'} \text{ を棄却}$$

全ての i について $H_0^{(i)'}$ が棄却されなければ、 $H_0(S)$ を採択する。この方法は Type I FWE をほぼ α でおさえる。即ち

$$\Pr(T_{(1)} < C, \dots, T_{(N-M)} < C \mid \bigcap_{j=1}^{N-M} H_0^{(j)'}) \geq 1 - \alpha$$

5 シミュレーション

400 個の物理乱数を用いて有意水準 $\alpha = 0.20$ でシミュレーションを試みた。



横軸 : i
 縦軸 : $T(i)$
 $C = 33.8737$

Fig. 5.1 simulation

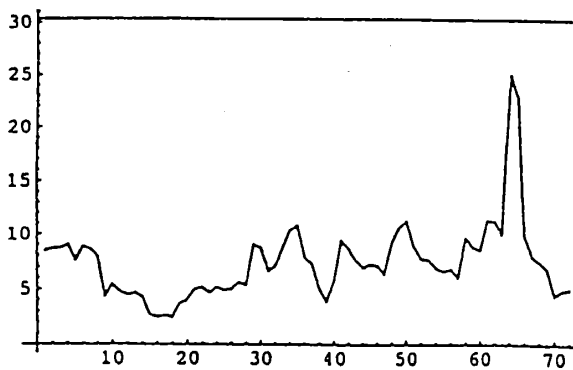
6 データ解析

6.1 札幌の年間降水量

1889 年から 1989 年までの札幌の年間降水量を、有意水準 0.10 で定常性検定した。

	0	1	2	3	4	5	6	7	8	9
1880										1100
1890	1101	977	1053	895	987	844	1151	725	1079	1069
1900	855	1155	929	896	1033	884	841	1121	903	1107
1910	841	1004	1168	907	1321	1115	889	1194	1041	1041
1920	1251	882	1339	1157	1003	1082	1359	1050	923	1109
1930	1152	1173	1369	1046	1002	1291	1231	1094	1108	1174
1940	938	1250	1095	977	1012	1176	1196	1177	891	1023
1950	1305	1112	963	1342	1058	1349	1141	1242	1215	1065
1960	1067	1095	1227	1138	1274	1344	1328	1104	981	958
1970	1119	953	1559	1173	1063	1432	1000	1103	1090	1079
1980	1179	1672	1045	885	725	1054	1128	998	1121	998
1990										

Table 6.1.1 札幌の年間降水量 (mm)



横軸 : i
 縦軸 : $T(i)$
 $C = 30.3253$

Fig. 6.1.1 統計量 $T(i)$

6.2 太陽黒点

太陽黒点の発生個数に対して、生データおよびログ変換したものを有意水準 0.10 で定常性検定した。

1770	101	1790	90	1810	0	1830	71	1850	66
1771	82	1791	67	1811	1	1831	48	1851	64
1772	66	1792	60	1812	5	1832	28	1852	54
1773	35	1793	47	1813	12	1833	8	1853	39
1774	31	1794	41	1814	14	1834	13	1854	21
1775	7	1795	21	1815	35	1835	57	1855	7
1776	20	1796	16	1816	46	1836	122	1856	4
1777	92	1797	6	1817	41	1837	138	1857	23
1778	154	1798	4	1818	30	1838	103	1858	55
1779	125	1799	7	1819	24	1839	86	1859	94
1780	85	1800	14	1820	16	1840	63	1860	96
1781	68	1801	34	1821	7	1841	37	1861	77
1782	38	1802	45	1822	4	1842	24	1862	59
1783	23	1803	43	1823	2	1843	11	1863	44
1784	10	1804	48	1824	8	1844	15	1864	47
1785	24	1805	42	1825	17	1845	40	1865	30
1786	83	1806	28	1826	36	1846	62	1866	16
1787	132	1807	10	1827	50	1847	98	1867	7
1788	131	1808	8	1828	62	1848	124	1868	37
1789	118	1809	2	1829	67	1849	96	1869	74

Table 6.2.1 Wölfer Sunspot Numbers (1770 - 1869)

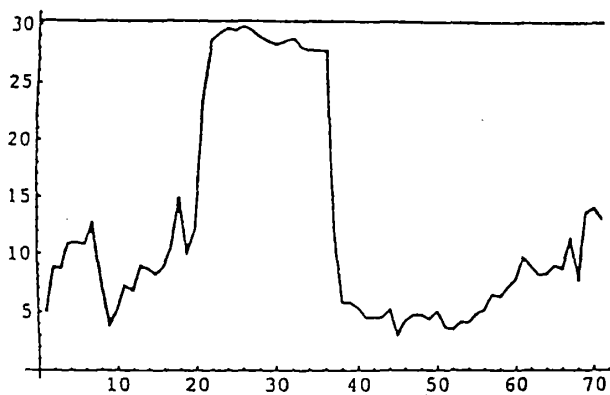


Fig. 6.2.1 統計量 $T(i)$ (生データ)

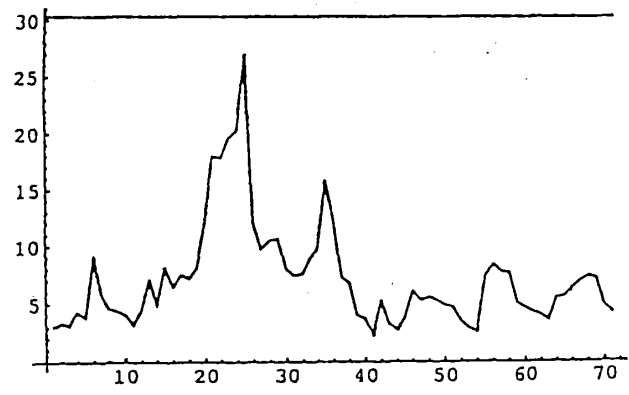


Fig. 6.2.2 統計量 $T(i)$ (ログ変換)

参考文献

- [1] 岡部靖憲. KM_2O - ランジュヴァン方程式と揺動散逸定理...
- 実証科学としての数学をめざして -. Pre-Print.
- [2] Blockwell, P.J. and R.A. Davis. *Time Series : Theory and Methods*.
Springer, New York. (1987)

† 812 福岡市東区箱崎 6 - 10 - 1 九州大学大学院数理学研究科
大濱 滋 ohama@math.kyushu-u.ac.jp
柳川 亮 yanagawa@math.kyushu-u.ac.jp