数理解析研究所講究録
975 巻 1996 年 14-27

定常過程に対する 最大エントロピー法と大偏差定理

名大・情報文化　　井原 俊輔　(Shunsuke Ihara)

1　はじめに

定常過程のスペクトル推定のひとつの方法として最大エントロピー法を提唱したのは Burg [1] であった．最大エントロピー法そのものは最初 Jaynes [8] によって提唱され，Burg はそれを定常過程の場合に適用したものである．最大エントロピー法は結局のところ，Shannon によるエントロピーの意義付けに，その根拠をおいており，それ以上の理論的根拠が与えられてきたとは言えない．しかし，応用面からは有効な方法で，以来，最大エントロピー法はいろいろな分野で広く応用されている．最大エントロピー法を一般化したもの最小相対エントロピー法にがある．最小相対エントロピー法は大偏差定理と密接な関連をもっている．

この論文では定常ガウス過程に対してひとつの条件付極限定理を証明する．この条件付極限定理は最小相対エントロピー法を，したがって最大エントロピー法を，正当化するものとなっている．

2　主定理

$X=\{X_n\}$ は正則（純非決定的）な実定常ガウス過程でそのスペクトル密度関数（以下では SDF と略記する）を $g(\lambda)$ とする．一般性を失うことなく，平均値は $E[X_n]=0$ としてよい．このとき共分散は

$$\gamma_n \equiv E[X_kX_{k+n}] = \int_{-\pi}^{\pi} e^{in\lambda} g(\lambda)d\lambda$$

とスペクトル表現される．

1 次独立な関数 $u_k(\lambda)\in L^2[0,2\pi]$, $k=1,2,...,$ で $u_k(-\lambda)=\overline{u_k(\lambda)}$ なるものを考え，確率ベクトル $Z_n^{(K)}=(Z_{n1},...,Z_{nK})$ を

$$Z_{nk} = \int_{-\pi}^{\pi} u_k(\lambda) I_n(\lambda)d\lambda \tag{1}$$

で定める．ここで

$$I_n(\lambda) = \frac{1}{2n\pi}\left|\sum_{j=1}^{n} X_j e^{-ij\lambda}\right|^2$$

はペリオドグラムである．

我々の目的は n を大きくしたときの $Z_n^{(K)}$ の漸近挙動を調べることである．なお，

$$\frac{1}{f_\theta(\lambda)} - \frac{1}{g(\lambda)} = -2\pi\sum_{k=1}^{K}\theta_k\{u_k(\lambda) + \overline{u_k(\lambda)}\} \tag{2}$$

で与えられる SDF f_θ $(\theta = (\theta_1, ..., \theta_K) \in \boldsymbol{R}^K)$ が重要な役割を果たす．また，簡単のため $\gamma^{(K)} = (\gamma_1, ..., \gamma_K) \in \boldsymbol{R}^K$ と記す．我々の主要な結果は次の条件付極限定理である．

定理 1 $X = \{X_n\}$ は SDF g をもつ正則な定常ガウス過程とする．関数 g および u_k $(k = 1, 2, ...)$ は連続で，さらに $g(\lambda) > 0, \forall\lambda \in [-\pi, \pi]$, と仮定する．

$$\int_{-\pi}^{\pi} u_k(\lambda) f^*(\lambda)\, d\lambda = \hat{\gamma}_k, \quad k = 1, ..., K, \tag{3}$$

をみたす SDF $f^* = f_{\theta^*}$ $(\theta^* \in \boldsymbol{R}^K)$ が存在するような $\hat{\gamma}^{(K)} = (\hat{\gamma}_1, ..., \hat{\gamma}_K)$ が与えられたとする．このとき，任意の $L > K$ と $\delta > 0$ に対し，ある $\varepsilon_0 > 0$ があって， $0 < \varepsilon \le \varepsilon_0$ なら

$$\lim_{n\to\infty} P\left(Z_n^{(L)} \in B_L(\hat{\gamma}^{(L)}, \delta) \,\middle|\, Z_n^{(K)} \in B_K(\hat{\gamma}^{(K)}, \varepsilon)\right) = 1 \tag{4}$$

が成り立つ．ただし，$B_m(\hat{\gamma}^{(m)}, \delta)$ は $\boldsymbol{R}^m$ における中心 $\hat{\gamma}^{(m)}$ 半径 δ の開球で，$k > K$ に対しては $\hat{\gamma}_k$ は

$$\hat{\gamma}_k = \int_{-\pi}^{\pi} u_k(\lambda) f^*(\lambda)\, d\lambda, \quad k > K, \tag{5}$$

で与えられる．

上の定理のもつ意味を説明しよう．真の SDF が未知であるとし，観測データからスペクトルを推定する問題を考える．$(X_1, ..., X_n)$ を観測した結果

$$Z_n^{(K)} = \hat{\gamma}^{(K)} \tag{6}$$

がわかったとする．ラフにいえば，(6) は未知の SDF f が条件

$$\int_{-\pi}^{\pi} u_k(\lambda) f(\lambda)\, d\lambda = \hat{\gamma}_k, \quad k = 1, ..., K, \tag{7}$$

をみたすことを意味する．ところで，定理は条件 (6) の下では任意の L に対し

$$Z_n^{(L)} \approx \hat{\gamma}^{(L)} \tag{8}$$

が起こることを示している．条件 (7) をみたす SDF はたくさんあり，定理 1 の f^* もそのひとつである．一方 (8) は真の SDF が f^* のときのみに成り立つ．このことは条件 (6)

（あるいは (7) ）の下では $f^* = f_{\theta^*}$ を真の SDF と推定すべきであることを意味している．

それでは，この $f^* = f_{\theta^*}$ は如何にして決まる SDF か？ 答えは次のとうり．SDF のクラス $\mathcal{F}$ を

$$\mathcal{F} = \{f;\ f\text{は (7) をみたす SDF}\}$$

と定めるとき，f^* は $\mathcal{F}$ の中で事前に与えられた SDF g に対する相対エントロピー $\overline{H}(f;g)$ を最小にするものである：

$$\overline{H}(f^*;g) = \inf\{\overline{H}(f;g);\ f \in \mathcal{F}\}.$$

このように，我々の定理は最小相対エントロピー法を，したがって最大エントロピー法を，正当化するものである．

最大エントロピー法で最もよく知られているのはある次数までの自己共分散が与えられたときのスペクトル推定である [1]．この場合は我々の議論の中で関数系 $\{u_k\}$ として

$$u_k(\lambda) = e^{ik\lambda}, \quad k = 0, 1, 2, ...,$$

をとればよい．

我々の定理に対応する結果は，i.i.d. の場合とマルコフ連鎖の場合には既に知られている（[2, 3, 4] 参照）．

最大エントロピー法については 3 節でまとめておく．4 節に述べる大偏差定理を適用して定理 1 を証明する．定理 1 の証明は 5 節で与える．

3　最大エントロピー法

一般に，確率空間 Ω 上の確率測度 μ, ν に対して，ν に関する μ の相対エントロピー $H(\mu;\nu)$ は，μ が ν に関して絶対連続のとき

$$H(\mu;\nu) = \int_{\Omega} \log \frac{d\mu}{d\nu}(\omega) d\mu(\omega)$$

で定義され，絶対連続でないときは $H(\mu;\nu) = \infty$ とする．

定常過程 $X = \{X_n\}$ に対して，各 n について $(X_1, ..., X_n)$ が連続分布に従っているとき，X の（単位時間当りの）エントロピーは

$$\overline{h}(X) = \lim_{n\to\infty} \frac{1}{n} h_n(X)$$

で定義される．ここで $h_n(X)$ は $(X_1, ..., X_n)$ の differential entropy．二つの定常過程 $X = \{X_n\}, Y = \{Y_n\}$ に対し $(X_1, ..., X_n), (Y_1, ..., Y_n)$ の確率分布を各々 μ_X^n, μ_Y^n と記す．μ_Y^n に関する μ_X^n の相対エントロピー $H(\mu_X^n;\mu_Y^n)$ のことを $(Y_1, ..., Y_n)$ に関する $(X_1, ..., X_n)$ の相対エントロピーといい $H_n(X;Y)$ と書く．そして Y に関する X の（単位時間当りの）相対エントロピー $\overline{H}(X;Y)$ は

$$\overline{H}(X;Y) = \lim_{n\to\infty} \frac{1}{n} H_n(X;Y)$$

で定義される．

ガウス過程の場合にはエントロピー，相対エントロピーは SDF を使って具体的に計算できる．正則な定常過程の SDF の全体を $\mathcal{S}$ とする．

$$\mathcal{S} = \{f \in L^1[-\pi, \pi]: f(-\lambda) = f(\lambda) \geq 0, \int_{-\pi}^{\pi} |\log f(\lambda)| d\lambda < \infty\}$$

である．SDF $f, g \in \mathcal{S}$ に対し

$$\overline{h}(f) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log\{4\pi e f(\lambda)\}\, d\lambda,$$

$$\overline{H}(f; g) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left\{ \frac{f(\lambda)}{g(\lambda)} - 1 - \log \frac{f(\lambda)}{g(\lambda)} \right\} d\lambda$$

とおく．正則な定常ガウス過程 X の SDF を f とするそのエントロピーは

$$\overline{h}(X) = \overline{h}(f)$$

で与えられる．この場合エントロピー $\overline{h}(X)$ は有限である．さらに Y を SDF g をもつ正則な定常ガウス過程とすると相対エントロピーは

$$\overline{H}(X; Y) = \overline{H}(f; g)$$

で与えられる．

定常過程のスペクトル推定のひとつの方法として最大エントロピー法がある．与えられた条件をみたす定常過程のうちでエントロピー最大のものを求める定常過程と推定するのが最大エントロピー法である（Burg [1]）．これに対し，先ず定常過程 Y が与えられたとする．このとき与えられた条件をみたす定常過程のうちで相対エントロピー $\overline{H}(X; Y)$ を最小にする X を求める定常過程と推定するのが最小相対エントロピー法である．最大エントロピー法は最小相対エントロピー法の特別の場合と考えるることができる．

ここでは，共分散の言葉で条件が与えられる場合を考える．この場合，条件は SDF の言葉で表現でき，基本的には (7) の形の条件を扱っておけばよい．この場合，Y がガウス過程とすると，与えられた条件をみたす定常過程の中ではガウス過程の場合に $\overline{H}(X; Y)$ は最小になる．そこで以下ではガウス過程のみを扱うこととする．最大エントロピー法および最小相対エントロピー法は以下のように述べることができる．今 SDF のクラス $\mathcal{F}$ が与えられたとする．観測データから未知の SDF のみたすべき条件が与えられ，$\mathcal{F}$ はこの条件をみたす SDF の全体である．$\mathcal{F}$ の中でエントロピー $\overline{h}(f)$ を最大にする f^*，即ち

$$\overline{h}(f^*) \geq \overline{h}(f), \quad \forall f \in \mathcal{F}, \tag{9}$$

なる $f^* \in \mathcal{F}$ を求める SDF と推定するのが最大エントロピー法である．一方，事前情報として SDF g が与えられているとき，$\mathcal{F}$ の中で相対エントロピー $\overline{H}(f; g)$ を最小にする f^*，即ち

$$\overline{H}(f^*; g) \leq \overline{H}(f; g), \quad \forall f \in \mathcal{F}, \tag{10}$$

なる $f^* \in \mathcal{F}$ を求める SDF と推定するのが最小相対エントロピー法である．

$$\overline{H}(\mathcal{F};g) = \inf\{\overline{H}(f;g); f \in \mathcal{F}\}$$

とおくと，(10) は

$$\overline{H}(f^*;g) = \overline{H}(\mathcal{F};g)$$

と表現できる．

特に，

$$g_0(\lambda) = \frac{\sigma^2}{2\pi}, \quad -\pi \le \lambda \le \pi, \tag{11}$$

なる SDF を考える．g_0 は分散 σ^2 の ガウス型白色雑音 の SDF である．このとき

$$\overline{H}(f;g_0) = -\overline{h}(f) + \frac{1}{2}\log(2\pi\sigma^2) + \frac{1}{2\sigma^2}\int_{-\pi}^{\pi} f(\lambda)\,d\lambda$$

だから，もし分散 $\gamma_0 = \int_{-\pi}^{\pi} f(\lambda)\,d\lambda$ が与えられた場合には，(9) と (10) は同値である．したがって最大エントロピー法は最小相対エントロピー法の特別の場合と考えてよい．

相対エントロピーを最小にする SDF f^* を具体的に求めるにあたっては，下に述べる定理が有用である．二つの SDF $g, f^* \in \mathcal{S}$ に対し SDF のクラス $\mathcal{F}^*(f^*,g)$ を

$$\mathcal{F}^*(f^*,g) = \left\{ f \in \mathcal{S}; \int_{-\pi}^{\pi} \left(\frac{1}{f^*(\lambda)} - \frac{1}{g(\lambda)} \right) f(\lambda)\,d\lambda \le c^* \right\} \tag{12}$$

で定める．ここで

$$c^* = \int_{-\pi}^{\pi} \left(\frac{1}{f^*(\lambda)} - \frac{1}{g(\lambda)} \right) f^*(\lambda)\,d\lambda.$$

f^* に対しては (12) が等号で成り立つから $f^* \in \mathcal{F}^*(f^*,g)$ であることに注意しておく．

定理 2（[7] 参照） SDF $g \in \mathcal{S}$ と SDF のクラス $\mathcal{F} \subset \mathcal{S}$ が与えられたとする．$\mathcal{F}$ は凸集合で $\overline{H}(\mathcal{F};g) < \infty$ を仮定する．このとき $f^* \in \mathcal{F}$ に対し $\overline{H}(f^*;g) = \overline{H}(\mathcal{F};g)$ が成り立つための必要十分条件は

$$\mathcal{F} \subset \mathcal{F}^*(f^*,g) \tag{13}$$

が成り立つことである．特に，$\overline{H}(f^*;g) = \overline{H}(\mathcal{F}^*(f^*,g);g)$ である．

自己回帰過程（AR 過程）および自己回帰移動平均過程（ARMA 過程）は最も基本的な定常過程である．SDF が

$$f(\lambda) = \frac{1}{2\pi}\frac{\sigma_0^2}{|Q(e^{i\lambda})|^2}$$

の形のとき，AR(K) 過程という．ここで $Q(z)$ は K 次多項式．また，SDF が

$$f(\lambda) = \frac{1}{2\pi}\frac{|P(e^{i\lambda})|^2}{|Q(e^{i\lambda})|^2}$$

のとき，ARMA(K,L) 過程という．ここで $P(z)$，$Q(z)$ は各々 L，K 次多項式．

定常過程の最大エントロピー解析は，Burg [1] による次の例とともに始まった．

例 1（AR 過程） K 次までの自己共分散（の推定値） $\hat{\gamma}_0, \hat{\gamma}_1, ..., \hat{\gamma}_K$ が与えられたとする．このとき，求める SDF f は

$$\int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda = \hat{\gamma}_k, \quad k = 0, 1, ..., K, \tag{14}$$

をみたさなければならない．$(\hat{\gamma}_0, \hat{\gamma}_1, ..., \hat{\gamma}_K)$ が正定値ならば，(14) をみたす AR(K) 過程の SDF f^* がただひとつ存在する．この f^* を SDF にもつガウス AR 過程が条件 (14) の下でエントロピーを最大にする定常過程である．

例 1 で述べたことは，定理 2 を使って容易に示すことができる．さらに，ARMA 過程については次のことが知られている（ [7] 参照）．

例 2（ARMA 過程） 事前情報として SDF g をもつガウス AR(N_0) 過程 ($N_0 \le K$) が与えられたとする．求める SDF f のみたす条件が (14) と

$$\int_{-\pi}^{\pi} \frac{f(\lambda)}{|1 - \beta_m e^{i\lambda}|^2} d\lambda = \sigma_m^2, \quad m = 1, ..., M, \tag{15}$$

であるとする（β_m は定数）．条件 (14) と (15) をみたす ARMA (N, M) 過程（ただし $N = K + M$）の SDF f^* が存在すれば，それがこれらの条件をみたす定常過程の中で相対エントロピーを最小にするものである．定理 2 によれば，最小相対エントロピー法の帰結として ARMA 過程が得られるためには，条件 (14) とならんで (15) の型の条件が本質的であることがわかる．

4 大偏差定理

定理 1 の証明には Gärtner [6], Ellis [5] による大偏差定理（ Large Deviation Theorem ）を使う．K 次元確率ベクトルの列 $\{Z_n\}_{n=1,2,...}$ に対して

$$\Lambda_n(\theta) = \log E[\exp(\theta, Z_n)], \quad \theta \in \boldsymbol{R}^K, \tag{16}$$

とおく．ただし $(\cdot, \cdot)$ は $\boldsymbol{R}^K$ の内積．

定理 3（ [4, 5, 6] 参照） 極限

$$\Lambda(\theta) = \lim_{n \to \infty} \frac{1}{n} \Lambda_n(n\theta) \tag{17}$$

が存在 ($\Lambda(\theta) = \pm\infty$ を許す) し，$0 \in \mathcal{D}_\Lambda^\circ$ とする．ここで $\mathcal{D}_\Lambda^\circ$ は集合 $\mathcal{D}_\Lambda = \{\theta \in \boldsymbol{R}^K; \Lambda(\theta) < \infty\}$ の内点集合を表す．$\Lambda(\theta)$ に対する Fenchel-Legendre 変換 $\Lambda^*(x)$ を

$$\Lambda^*(x) = \sup_{\theta \in \boldsymbol{R}^K} \{(\theta, x) - \Lambda(\theta)\}, \quad x \in \boldsymbol{R}^K \tag{18}$$

と定義する．このとき，以下のことが成り立つ．
(a) 任意の閉集合 F に対し

$$\limsup_{n\to\infty}\frac{1}{n}\log P(Z_n\in F)\le -\inf_{x\in F}\Lambda^*(x). \tag{19}$$

(b) 任意の開集合 G に対し

$$\liminf_{n\to\infty}\frac{1}{n}\log P(Z_n\in G)\ge -\inf_{x\in G\cap H}\Lambda^*(x), \tag{20}$$

ただし，H はその exposing hyperplane が $\mathcal{D}_\Lambda^\circ$ に属すような Λ^* の exposed point の集合．なお，ある $\theta\in\boldsymbol{R}^K$ があって

$$(\theta,y)-\Lambda^*(y)>(\theta,x)-\Lambda^*(x),\quad \forall x\ne y,$$

のとき，$y\in\boldsymbol{R}^K$ は Λ^* の exposed point といい，点 θ を exposing hyperplane という．

5 主定理の証明

本節では定理 1 の証明を与える．定理 1 の関数系 $\{u_k\}$ としては，$\{u_k(\lambda)+\overline{u_k(\lambda)}\}/2$ をあらためて $u_k(\lambda)$ として議論してもよい．そこで，以下，$u_k(\lambda)$ は実関数ととする．したがって定義 (2) の f_θ は

$$f_\theta(\lambda)=\frac{g(\lambda)}{1-4\pi\sum_{k=1}^K\theta_k u_k(\lambda)g(\lambda)} \tag{21}$$

と書ける．

定理の証明に必要となる記号を導入しておこう．集合 $\Theta_K\subset\boldsymbol{R}^K$ を

$$\Theta_K=\left\{\theta\in\boldsymbol{R}^K;4\pi\sum_{k=1}^K\theta_k u_k(\lambda)g(\lambda)<1,\forall\lambda\in[-\pi,\pi]\right\}$$

と定める．$\boldsymbol{R}^K$ の部分集合 $\boldsymbol{P}_K,\boldsymbol{D}_K$ を

$$\boldsymbol{P}_K=\{\langle\mathbf{u},f\rangle_K;f\in\mathcal{S}\},\qquad \boldsymbol{D}_K=\{\langle\mathbf{u},f_\theta\rangle_K;\theta\in\Theta_K\} \tag{22}$$

と定める．ただし

$$\langle\mathbf{u},f\rangle_K=(\langle u_1,f\rangle,...,\langle u_K,f\rangle)\in\boldsymbol{R}^K,$$

そして $\langle u,v\rangle=\int_{-\pi}^{\pi}u(\lambda)\overline{v(\lambda)}d\lambda$ は $L^2[-\pi,\pi]$ の内積を表す．(21) より $\theta\in\Theta_K$ のとき $f_\theta\in\mathcal{S}$ がわかる．したがって $\boldsymbol{D}_K\subset\boldsymbol{P}_K$ である．次に，各 $\theta\in\Theta_K$ に対し，半平面 $\boldsymbol{Q}_K(\theta)$ を

$$\boldsymbol{Q}_K(\theta)=\{x\in\boldsymbol{R}^K;(\theta,x)\ge(\theta,\langle\mathbf{u},f_\theta\rangle_K)\} \tag{23}$$

と定義する．以下 $\theta^*\in\Theta_K$ を固定し，$f^*=f_{\theta^*}$ とおく．任意の $f\in\mathcal{S}$ に対して，(2) より，

$$(\theta^*,\langle\mathbf{u},f\rangle_K)=\sum_{k=1}^K\theta_k^*\int_{-\pi}^{\pi}u_k(\lambda)f(\lambda)d\lambda=-\frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{1}{f_{\theta^*}(\lambda)}-\frac{1}{g(\lambda)}\right)f(\lambda)d\lambda$$

である．よって，(12) に注意すれば，次の同値関係が成り立つことがわかる．

$$\langle \mathbf{u}, f\rangle_K \in \boldsymbol{Q}_K(\theta^*) \iff f \in \mathcal{F}^*(f^*, g) \tag{24}$$

したがって定理2より次のことが導ける．

$$\overline{H}(f^*; g) \leq \overline{H}(f; g) \quad \text{if} \quad \langle \mathbf{u}, f\rangle_K \in \boldsymbol{Q}_K(\theta^*). \tag{25}$$

関数 $v \in L^2[-\pi, \pi]$ に対する n 次元 Toeplitz 行列を $T_n(v)$ と記す．即ち，

$$T_n(v) = \Big[t_{pq}(v)\Big]_{p,q=1,\ldots,n}, \qquad t_{pq}(v) = \int_{-\pi}^{\pi} e^{i(p-q)\lambda} v(\lambda) d\lambda.$$

Toeplitz 行列を使えば，(1) の確率変数 Z_{nk} は

$$Z_{nk} = \frac{1}{2n\pi}(T_n(u_k)X', X), \quad k = 1, 2, \ldots, \tag{26}$$

と書ける．ここで $X = (X_1, \ldots, X_n)$ で X' は X を転置したもの．さらに，λ の関数として $|\sum_{j=1}^n X_j e^{ij\lambda}|^2 \in \mathcal{S}$ だから，(22) より

$$P(Z_n^{(K)} \in \boldsymbol{P}_K) = 1 \tag{27}$$

がわかる．よって定理3の (19), (20) において集合 F, G としては $\boldsymbol{P}_K$ の部分集合のみを考えればよい．

定理1の証明のためいくつかの補題を準備する．定理1の確率ベクトルの列 $\{Z_n^{(K)}\}$ に対し，$\Lambda_n(\theta) = \log E[\exp(\theta, Z_n^{(K)})]$ とし，$\Lambda(\theta)$, $\Lambda^*(x)$ を各々 (17), (18) で定義する．

補題 1 $\theta \in \Theta_K$ とする．十分大きな n に対し，行列 $E_n - 2T_n(g)T_n(\sum_{k=1}^K \theta_k u_k)$ は正則で $\Lambda_n(\theta)$ は

$$\Lambda_n(\theta) = -\frac{1}{2}\log\left|E_n - \frac{1}{n\pi}T_n(g)T_n\left(\sum_{k=1}^K \theta_k u_k\right)\right| \tag{28}$$

と書ける．ここで E_n は n 次元単位行列．

証明 $(X_1, \ldots, X_n)$ は n 次元ガウス分布 $N(0, T_n(g))$ に従うから，(26) に注意し，

$$\begin{aligned}
&E[\exp(\theta, Z_n^{(K)})] \\
&= \frac{1}{(2\pi)^{n/2}|T_n(g)|^{1/2}}\int_{\boldsymbol{R}^n} \exp\left\{\frac{\sum_{k=1}^K \theta_k (T_n(u_k)x, x)}{2n\pi} - \frac{1}{2}(T_n(g)^{-1}x, x)\right\} dx \\
&= \frac{\left|\left[T_n(g)^{-1} - \frac{1}{n\pi}T_n\left(\sum_{k=1}^K \theta_k u_k\right)\right]^{-1}\right|^{1/2}}{|T_n(g)|^{1/2}} \\
&= \left|E_n - \frac{1}{n\pi}T_n(g)T_n\left(\sum_{k=1}^K \theta_k u_k\right)\right|^{-1/2}
\end{aligned}$$

を得る．両辺の対数をとり (28) を得る． Q.E.D.

補題 2　$\theta \in \Theta_K$ のとき

$$\Lambda(\theta) = -\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left\{1 - 4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)\right\}d\lambda. \tag{29}$$

さらに $\Lambda(\theta)$ は Θ_K で2回偏微分可能で，その偏導関数は次で与えられる．

$$\frac{\partial\Lambda}{\partial\theta_m}(\theta) = \langle u_m, f_\theta\rangle, \tag{30}$$

$$\frac{\partial^2\Lambda}{\partial\theta_n\partial\theta_m}(\theta) = 4\pi\int_{-\pi}^{\pi}\frac{u_n(\lambda)u_m(\lambda)g(\lambda)^2}{\{1-4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)\}^2}d\lambda \tag{31}$$

$(m, n = 1, 2, \ldots)$. また次式も成り立つ．

$$\overline{H}(f_\theta; g) = \sum_{k=1}^{K}\theta_k\int_{-\pi}^{\pi}u_k(\lambda)f_\theta(\lambda)d\lambda - \Lambda(\theta). \tag{32}$$

証明　一般に，$v(\lambda) > 0$ なる連続関数 v に対し

$$\lim_{n\to\infty}\frac{1}{n}\log|T_n(v)| = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log\{2\pi v(\lambda)\}d\lambda \tag{33}$$

が成り立つ．したがって補題1より

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\Lambda_n(n\theta) &= -\lim_{n\to\infty}\frac{1}{2n}\log\left|E_n - \frac{1}{\pi}T_n(g)T_n(\textstyle\sum_{k=1}^{K}\theta_k u_k)\right| \\
&= -\lim_{n\to\infty}\frac{1}{2n}\log\left\{|T_n(g)|\left|T_n(g)^{-1} - \frac{1}{\pi}T_n(\textstyle\sum_{k=1}^{K}\theta_k u_k)\right|\right\} \\
&= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{\log[2\pi g(\lambda)] + \log\left[\frac{1}{2\pi g(\lambda)} - 2\sum_{k=1}^{K}\theta_k u_k(\lambda)\right]\right\}d\lambda \\
&= -\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left\{1 - 4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)\right\}d\lambda
\end{aligned}$$

となり (29) を得る．微分と積分の順序交換をしながら，(29) の両辺を微分して (30) および (31) を得る．

$$\overline{H}(f_\theta; g) = \frac{1}{4\pi}\int_{-\pi}^{\pi}\left(\frac{4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)}{1-4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)} - \log\frac{1}{1-4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)g(\lambda)}\right)d\lambda$$

だから，(21) と (29) より (32) を得る．　Q.E.D.

この補題から $\overline{H}(f_\theta; g)$ は $\theta \in \Theta_K$ の連続関数であることがわかる．

補題 3　任意の $f \in \mathcal{S}$ に対し，

$$\Lambda^*(\langle \mathbf{u}, f\rangle_K) \le \overline{H}(f; g) \tag{34}$$

である．特に，$f_\theta \in \mathcal{S}$ $(\theta \in \Theta_K)$ に対しては

$$\Lambda^*(\langle \mathbf{u}, f_\theta\rangle_K) = \overline{H}(f_\theta; g) \tag{35}$$

が成立ち，$y = \langle \mathbf{u}, f_\theta\rangle_K$ は Λ^* の exposed point で，θ が y に対する exposing hyperplane である．

証明 $y = \langle \mathbf{u}, f\rangle_K$ とおく．(34) を示すためには，任意の $\theta \in \boldsymbol{R}^K$ に対し

$$(\theta, y) - \Lambda(\theta) \leq \overline{H}(f; g) \tag{36}$$

を示せばよい．このことは

$$\begin{aligned}&(\theta, y) - \Lambda(\theta) - \overline{H}(f; g)\\&= \frac{1}{4\pi}\int_{-\pi}^{\pi}\left\{4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)f(\lambda) - \frac{f(\lambda)}{g(\lambda)} + 1 + \log\left(\frac{f(\lambda)}{g(\lambda)} - 4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)f(\lambda)\right)\right\}d\lambda\\&\leq 0\end{aligned}$$

からわかる．なお，最後の不等式を示すには $\log x \leq x - 1, \forall x > 0,$ を使えばよい．また等号は

$$\frac{f(\lambda)}{g(\lambda)} - 4\pi\sum_{k=1}^{K}\theta_k u_k(\lambda)f(\lambda) = 1, \quad -\pi \leq \lambda \leq \pi,$$

のとき，即ち f が (21) の f_θ に等しいときのみ成り立つ．次に，$\theta \in \Theta_K$ とし $y = \langle \mathbf{u}, f_\theta\rangle$ とする．このとき (30) より

$$\frac{\partial}{\partial\theta_m}((\theta, y) - \Lambda(\theta)) = \frac{\partial}{\partial\theta_m}\left(\sum_{k=1}^{K}\theta_k y_k - \Lambda(\theta)\right) = y_m - \frac{\partial\Lambda}{\partial\theta_m}(\theta) = 0, \quad m = 1, ..., K,$$

である．これより $\Lambda^*(y) = (\theta, y) - \Lambda(\theta))$ がわかる．したがって (32) より (35) が導ける．exposed point に関する主張は [4, Lemma 2.3.9] に示されている． Q.E.D.

補題 4 任意の $x \in \boldsymbol{Q}_K(\theta^*)$ に対し

$$\Lambda^*(x) \geq \overline{H}(f^*; g). \tag{37}$$

上式で等号が成立するのは $x = \langle \mathbf{u}, f^*\rangle_K$ のときに限る．さらに，

$$\Lambda^*(x) \geq \alpha, \quad x \in G^c \cap \boldsymbol{Q}_K(\theta^*), \tag{38}$$

をみたすような開集合 $G \ni \langle \mathbf{u}, f^*\rangle_K$ と定数 $\alpha > \overline{H}(f^*; g)$ が存在する．

証明 $x \in \boldsymbol{Q}_K(\theta^*)$ のとき，(32), (23) と $\Lambda^*(y)$ の定義より

$$\begin{aligned}\overline{H}(f^*; g) &= \sum_{k=1}^{K}\theta_k^*\int_{-\pi}^{\pi}u_k(\lambda)f^*(\lambda)d\lambda - \Lambda(\theta^*)\\&\leq (\theta^*, x) - \Lambda(\theta^*)\\&\leq \sup_\theta\{(\theta, x) - \Lambda(\theta)\}\\&= \Lambda^*(x)\end{aligned}$$

である．よって (37) が得られた．$x^* = \langle \mathbf{u}, f^*\rangle_K$ とおく．補題 3 に注意すれば，定理 2 より，

$$\Lambda^*(x) > \overline{H}(f^*; g), \quad x \in \boldsymbol{Q}_K(\theta^*) \cap \boldsymbol{D}_K,\ x \neq x^*, \tag{39}$$

がわかる．各点 $\theta \in \Theta_K$ において

$$\frac{\partial}{\partial \theta_m} \int_{-\pi}^{\pi} u_n(\lambda) f_\theta(\lambda) d\lambda = 4\pi \int_{-\pi}^{\pi} \frac{u_m(\lambda) u_n(\lambda) g(\lambda)^2}{\left(1 - 4\pi \sum_{k=1}^{K} \theta_k u_k(\lambda) g(\lambda)\right)^2} d\lambda$$

だから，$\boldsymbol{R}^K$ における Θ_K から $\boldsymbol{D}_K$ への写像

$$\varphi(\theta) = \langle \mathbf{u}, f_\theta \rangle_K$$

のヤコビー行列は狭義正定値である．したがって，θ^* を含む開集合 V に対し $\varphi(V)$ は x^* を含む開集合で，(39) より

$$\Lambda^*(x) > \overline{H}(f^*; g), \quad x \in \boldsymbol{Q}_K(\theta^*) \cap \varphi(V),\ x \neq x^*, \tag{40}$$

である．Λ^* は凸関数だから，(40) より

$$\Lambda^*(x) > \overline{H}(f^*; g), \quad x \in \boldsymbol{Q}_K(\theta^*),\ x \neq x^*, \tag{41}$$

がわかる．このことは (37) で等号が成り立つのは $x = x^*$ のときに限ることを示している．有界な開集合 G で $x^* \in G \subset \overline{G} \subset \varphi(V)$ なるものが存在する．G の境界を ∂G とすると，コンパクト集合 $\partial G \cap \boldsymbol{Q}_K(\theta^*)$ 上で連続な関数 $\Lambda^*(x)$ の最小値を α とすれば (40) より

$$\overline{H}(f^*; g) < \alpha$$

である．$\boldsymbol{Q}_K(\theta^*)$ は x^* を中心とする cone と考えられるから，上のことから (38) が成り立つ． Q.E.D.

準備ができたので定理 1 を証明する．

定理 1 の証明 集合 $W_d^{(m)}$ $(m = 1, 2, \ldots,\ d > 0)$ を

$$W_d^{(m)} = \{\langle \mathbf{u}, f_\theta \rangle_m ; \theta \in \overline{B}_K(\theta^*, d)\}$$

と定める．$h > 0$ を $\overline{B}_K(\theta^*, h) \subset \Theta_K$ かつ

$$\max_{\theta \in \overline{B}_K(\theta^*, h)} \sum_{k=1}^{L} \{\langle u_k, f_\theta \rangle - \langle u_k, f^* \rangle\}^2 < \frac{\delta^2}{4}$$

をみたすようにとれば，

$$W_h^{(L)} \subset B_L(\langle \mathbf{u}, f^* \rangle_L, \delta/2) \tag{42}$$

である．このとき，ある $\varepsilon_0 > 0$ があって

$$B_K(\langle \mathbf{u}, f^* \rangle_K, \varepsilon_0) \subset W_h^{(K)} \subset \boldsymbol{D}_K \tag{43}$$

が成り立つ．$0 < \varepsilon \leq \varepsilon_0$ とし，

$$\mathcal{F}_\varepsilon = \{f \in \mathcal{S}; \langle \mathbf{u}, f \rangle_K \in B_K(\langle \mathbf{u}, f^* \rangle_K, \varepsilon)\}$$

と定義すると，$\mathcal{F}_\varepsilon$ は凸集合である．ある $\tilde{\theta} \in \overline{B}_K(\theta^*, h)$ が存在し，$\tilde{f} = f_{\tilde{\theta}}$ は

$$\langle \mathbf{u}, \tilde{f} \rangle_K \in \overline{B}_K(\langle \mathbf{u}, f^* \rangle_K, \varepsilon)$$

および

$$\overline{H}(\tilde{f}; g) = \inf\{\overline{H}(f; g); f \in \mathcal{F}_\varepsilon\}$$

をみたす．以下 ε と δ を固定し，簡単のため

$$B_K = B_K(\langle \mathbf{u}, f^* \rangle_K, \varepsilon), \quad B_L = B_L(\langle \mathbf{u}, f^* \rangle_L, \delta)$$

と記す．閉球に対しては $\overline{B}_K, \overline{B}_L$ と略記する．補題 3 より，B_K の各点 $\langle \mathbf{u}, f_\theta \rangle_K$ は Λ^* の exposed point でその exposing hyperplane は $\theta \in \Theta_K \subset \mathcal{D}_\Lambda^\circ$ である．よって定理 3 (b) より

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log P\left(Z_n^{(K)} \in B_K\right) &\geq -\inf_{x\in B_K} \Lambda^*(x) \\ &= -\inf\{\overline{H}(f_\theta; g); f_\theta \in \mathcal{F}_\varepsilon\} \\ &= -\overline{H}(\tilde{f}; g) \end{aligned} \tag{44}$$

が成り立つ．

$$B_{L,K} = B_K \times \boldsymbol{R}^{L-K} = \{(x_1, ..., x_L) \in \boldsymbol{R}^L; (x_1, ..., x_K) \in B_K\}$$

とおき，その閉包を $\overline{B}_{L,K}$ と記す．定理 3 (a) より，$\{Z_n^{(L)}\}$ に対応する Fenchel-Legendre 変換を $\Lambda_L^*(x)$ とすると，

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log P\left(Z_n^{(L)} \in B_L^c,\, Z_n^{(K)} \in B_K\right) &\leq \limsup_{n\to\infty} \frac{1}{n} \log P\left(Z_n^{(L)} \in B_L^c \cap \overline{B}_{L,K}\right) \\ &\leq -\inf\{\Lambda_L^*(x); x \in B_L^c \cap \overline{B}_{L,K}\} \end{aligned} \tag{45}$$

である．上述の $\tilde{\theta} = (\tilde{\theta}_1, ..., \tilde{\theta}_K) \in \boldsymbol{R}^K$ に対し $\tilde{\theta}^{(L)} = (\tilde{\theta}_1, ..., \tilde{\theta}_K, 0, ..., 0) \in \boldsymbol{R}^L$ と定める．明らかに

$$\tilde{f} = f_{\tilde{\theta}} = f_{\tilde{\theta}^{(L)}}$$

である．

$$\boldsymbol{Q}_L(\tilde{\theta}^{(L)}) = \boldsymbol{Q}_K(\tilde{\theta}) \times \boldsymbol{R}^{L-K}$$

に注意すれば，

$$\langle \mathbf{u}, \tilde{f} \rangle_L \in \overline{B}_{L,K} \subset \boldsymbol{Q}_L(\tilde{\theta}^{(L)})$$

がわかる．したがって，補題 4 より

$$\Lambda_L^*(x) \geq \overline{H}(\tilde{f}; g), \quad \forall x \in \overline{B}_{L,K},$$

である．同じく補題 4 より

$$\langle \mathbf{u}, \tilde{f} \rangle_L \in G \subset B_L$$

なる開集合 G と定数 $\alpha > \overline{H}(\tilde{f};g)$ があって

$$\Lambda_L^*(x) \geq \alpha, \quad \forall x \in G^c \cap \boldsymbol{Q}_L(\tilde{\theta}^{(L)}),$$

とできる．$B_L^c \subset G^c$ かつ $\overline{B}_{L,K} \subset \boldsymbol{Q}_L(\tilde{\theta}^{(L)})$ だから，

$$\inf\{\Lambda_L^*(x); x \in B_L^c \cap \overline{B}_{L,K}\} \geq \alpha$$

である．よって (45) より

$$\limsup_{n\to\infty} \frac{1}{n} \log P\left(Z_n^{(L)} \in B_L^c,\, Z_n^{(K)} \in B_K\right) \leq -\alpha \tag{46}$$

を得る．$\overline{H}(\tilde{f};g) < \gamma < \beta < \alpha$ とすると，(44) と (46) よりある n_0 があって，$\forall n \geq n_0$ に対し

$$P\left(Z_n^{(K)} \in B_K\right) \geq e^{-\gamma n},$$
$$P\left(Z_n^{(L)} \in B_L^c,\, Z_n^{(K)} \in B_K\right) \leq e^{-\beta n}$$

が成り立つ．よって $n \geq n_0$ のとき

$$P\left(Z_n^{(L)} \in B_L^c \,\middle|\, Z_n^{(K)} \in B_K\right) \leq \frac{e^{-\beta n}}{e^{-\gamma n}} = e^{-(\beta-\gamma)n}$$

である．よって，$n \to \infty$ として (4) を得る． Q.E.D.

6 自己回帰モデル

我々の結果を自己回帰モデルの場合に適用する．関数系 $\{u_k\}$ として

$$u_k(\lambda) = e^{ik\lambda}$$

を考える．このとき (1) の確率変数は

$$Z_{nk} = \int_{-\pi}^{\pi} e^{ik\lambda} I_n(\lambda) d\lambda = \frac{1}{n} \sum_{j=1}^{n-k} X_j X_{j+k}$$

となり，これは k 次自己共分散に対する推定量である．$(X_1, ..., X_n)$ を観測した結果，

$$Z_{nk} = \hat{\gamma}_k, \quad k = 0, ..., K, \tag{47}$$

がわかったとする．問題は，条件 (47) の下で，真の SDF に対する最適の推定をすることである．あるいは，$K+1$ 次以上の自己共分散

$$\gamma_{K+1},\ \gamma_{K+2},\ \cdots\ ,$$

に対する最適な推定値を求めることといってもよい．事前情報として $X = \{X_n\}$ は AR(J) $(J \leq K)$ 過程とする．このとき，例 1，例 2 でみたように，最小相対エントロピー法は最適な推定 SDF として

$$\int_{-\pi}^{\pi} e^{ik\lambda} f^*(\lambda)\, d\lambda = \gamma_k, \quad k = 0, ..., K,$$

をみたす AR(K) 過程の SDF f^* を選ぶ．そして，γ_m $(m \geq K+1)$ に対する最適な推定値は

$$\gamma_m^* = \int_{-\pi}^{\pi} e^{im\lambda} f^*(\lambda) d\lambda, \quad m \geq K+1,$$

である．これに対し，定理 1 は任意の $m \geq K+1$ に対し，$\varepsilon > 0$ が十分小さければ，

$$\lim_{n\to\infty} P\left(|Z_{nm} - \gamma_m^*| < \delta \;\Big|\; \sum_{k=0}^{K} |Z_{nk} - \gamma_k|^2 < \varepsilon^2 \right) = 1$$

が成り立つことを主張している．このことから，我々の定理 1 が，標本の大きさが十分大きいときに，最大エントロピー法あるいは最小相対エントロピー法が最適な推定を行っているを保証しているがわかる．

参考文献

[1] Burg, J.P.: Maximum entropy spectral analysis. Presented at the 37th Ann. Int. Meet. Soc. Explor. Geophys., Oklahoma City, OK, 1967.; Modern Spectral Analysis, Childer, D.G. (Ed.), IEEE Press, New York, 1978, 34–41.

[2] Csiszár, I.: Sanov property, generalized I-projection and a conditional limit theorem. Ann. Prob., 12 (1984), 768–793.

[3] Cover, T.M. and Thomas, J.A.: Elememts of Information theory. Wiley, New York, 1991.

[4] Dembo, A. and Zeitouni, O.: Large Deviations Techniques and Applications. Jones and Bartlett Pub., Boston, MA, 1992.

[5] Ellis, R.S.: Large deviations for a genaral class of random vectors. Ann. Prob., 12 (1984), 1–12.

[6] Gärtner, J.: On large deviations from the invariant measure. Thory Prob. Appl., 22 (1977), 24–39.

[7] Ihara, S.: Information Theory for Continuous Systems. World Scientific, Singapore, 1993.

[8] Jaynes, E.T.: Information theory and statistical mechanics. Phys. Lev., 106 (1957), 620–630.