

距離尺度による時系列パターンの分類

正員 石井 直宏[†] 正員 方野 健次[†]
正員 岩田 彰[†] 正員 鈴村 宣夫[†]

Classification of the Time Series by Distance Measures

Naohiro ISHII[†], Kenji KATANO[†], Akira IWATA[†] and
Nobuo SUZUMURA[†], *Regular Members*

あらまし 非定常時系列から情報を抽出するため、まず非定常時系列を定常小区間に分割し、その上で情報の抽出を行い、時系列パターンを分類する方法がある。このとき、時系列を比較し、分類するための距離尺度が必要となる。本論文では、統計的パターンの分類で導入されているKullback-Leibler情報量、Kullback-Leibler divergence, Bhattacharyya距離およびChernoff距離を平均値が零の場合の時系列について比較検討したものである。具体的な時系列としては、脳波時系列などの零平均値で振幅および周波数成分が変化する場合について、Kullback-Leibler, Kullback-Leibler divergenceの距離尺度の感度がBhattacharyya及びChernoffの距離尺度の感度より敏感であることを明らかにした。次にKullback-Leibler情報量を睡眠脳波時系列へ適用し、7グループの睡眠脳波パターンを分類した。この分類には、階層的クラスタリングと正準判別分析を適用し検討した。

1. ま え が き

非定常時系列パターンを分類する問題は地震、音声、経済変動および生体電気活動など多くの分野で見い出される。これらの非定常時系列から意味のある情報を抽出するために、定常小区間に分割し、時系列パターンを分類するような距離尺度が必要となる⁽¹⁾。距離尺度としては、時系列へ自己回帰モデルを適合させ^{(2)~(4)}、その予測残差から、時系列の変化を検出するようなスペクトル誤差量がある⁽⁵⁾。又、音声の情報処理などの距離尺度として、最ゆう法による残差のゆう度の展開がある⁽⁶⁾。この距離尺度は最ゆう距離とも呼ばれている。さて本論文で注目する脳波時系列は、その情報が交流成分にあり、通常直流成分は除去される。このことは零平均の時系列となることを意味する。そこで零平均値の二つの時系列のパワースペクトル密度が互いに近い条件では、Kullback-Leibler情報量(K-L情報量)⁽⁷⁾が、距離尺度としてのスペクトル誤差量と

なることを示した。次に時系列を自己回帰モデルへ適合させた条件のもとで、K-L情報量は最ゆう距離と等価となることを示した。K-L情報量を他の距離尺度と比較するため、K-L divergence, Bhattacharyya距離⁽⁸⁾、およびChernoff距離を導入した。時系列の変化に対する、これらの距離尺度の感度を比較するため、はじめに振幅の度合いの変化、次に周波数の変化について検討した。ここで周波数の変化は時間軸上の系列相関係数の変化として、とらえ比較した。その結果、振幅および周波数の変化を検出する感度として、K-L divergence, K-L情報量が優れており、Bhattacharyya距離およびChernoff距離の感度の低いことが示された。

更に、K-L情報量と等価な最ゆう距離を使って睡眠脳波時系列パターンを分類した。その分類の方法として、階層的クラスタリングを適用した。その結果、最ゆう距離は二つの時系列のパワースペクトル密度の差を全周波数にわたって積分した値であり、ある周波数帯のわずかな特徴の差違の抽出としては、必ずしもよい結果を示さない。そこでK-L情報量を各周波数帯に分割した値を各々変数とする正準判別分析を繰り返し適用した。この分析法により、睡眠脳波時系列パ

[†]名古屋工業大学工学部情報工学科, 名古屋市
Faculty of Engineering, Nagoya Institute of Technology,
Nagoya-shi, 466 Japan
論文番号: 昭 56-95[D-16]

ターンが分類できるようになり、有効な方法であることが示された。

2 K-L情報量の導入

はじめに自己回帰過程およびKullback-Leibler 情報量(K-L情報量)を導入する。\$m\$ 次の自己回帰過程は式(1)で表される^{(2),(3)}。

$$X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_m X_{t-m} = \xi_t \quad (1)$$

ここで \$\{X_t\}\$ は定常確率過程、\$\{\alpha_i\}\$ は係数の集合を示す。\$\{X_t\}\$ が定常正規確率過程なら、時刻 \$t = n\$ の最良推定値は、

$$\begin{aligned} \hat{x}_n &= E[x_n | x_{n-m}, x_{n-m+1}, \dots, x_{n-1}] \\ &= -\{\alpha_1 x_{n-1} + \alpha_2 x_{n-2} + \dots + \alpha_m x_{n-m}\} \end{aligned} \quad (2)$$

と表現できる。ここで \$E[a|b]\$ は事象 \$b\$ の生起の下で \$a\$ の生起する条件付確率の期待値を表す。時刻 \$t = n\$ の予測誤差を \$e_n\$ とすると、

$$e_n = x_n - \hat{x}_n \quad (3)$$

式(2)を式(3)へ代入して \$z\$ 変換すると、

$$E(z) = \left(1 + \sum_{k=1}^m \alpha_k z^{-k}\right) X(z) \equiv A(z) X(z) \quad (4)$$

を得る。ここで \$E(z)\$ は \$e_n\$ の \$z\$ 変換を表す。更に、\$\{X_t\}\$ のパワースペクトル密度を \$p(\omega)\$ で表すと、

$$p(\omega) = |X(e^{j\omega})|^2 \quad (5)$$

で与えられる。式(4)の \$z = \exp(j\omega)\$ とおくと、\$\{X_t\}\$ のパワースペクトル密度 \$p(\omega)\$ は

$$p(\omega) = \frac{|E[\exp(j\omega)]|^2}{|A[\exp(j\omega)]|^2} \quad (6)$$

となる。さて、二つの正規密度関数を各々、\$N(\mu_1, \Sigma_1)\$ 及び \$N(\mu_2, \Sigma_2)\$ で表す。

ここで \$\mu_i\$ (\$i=1, 2\$) は平均値、\$\Sigma_i\$ (\$i=1, 2\$) は共分散行列を表す。このとき K-L 情報量は

$$\begin{aligned} I(1:2) &= \frac{1}{2} \log \frac{|\Sigma_2|}{|\Sigma_1|} + \frac{1}{2} \text{tr} \Sigma_1 (\Sigma_2^{-1} - \Sigma_1^{-1}) \\ &\quad + \frac{1}{2} \text{tr} \Sigma_2^{-1} (\mu_1 - \mu_2)(\mu_1 - \mu_2)^t \end{aligned} \quad (7)$$

となる⁽⁷⁾。ここで \$\Sigma_i^{-1}\$ (\$i=1, 2\$) は \$\Sigma_i\$ の逆行列を表し、“tr” は行列のトレースを表す。“\$t\$” は転置を示す。

更に、データ長 \$T\$ の K-L 情報量は

$$\begin{aligned} I_T(1:2) &= \frac{1}{2} T^{-1} \left[\log \frac{|\Sigma_2|}{|\Sigma_1|} + \text{tr} \Sigma_1 (\Sigma_2^{-1} - \Sigma_1^{-1}) \right] \\ &\quad + \frac{1}{2} T^{-1} \delta^t \Sigma_2^{-1} \delta \end{aligned} \quad (8)$$

ここで

$$\mu_j = \{\mu_j(0), \dots, \mu_j(T-1), \mu_j(T)\} \quad (j=1, 2)$$

$$\left. \begin{aligned} \Sigma_j &= \{\sigma_j(s-t), s, t=0, 1, 2, \dots, (T-1)\} \\ \delta &= \mu_1 - \mu_2 \end{aligned} \right\} \quad (9)$$

である。データ長 \$T\$ が大なら

$$\lim_{T \rightarrow \infty} T^{-1} \log |\Sigma_j| = \int_{-\pi}^{\pi} \log p_j(\omega) \frac{d\omega}{2\pi} \quad (10)$$

及び

$$\lim_{T \rightarrow \infty} T^{-1} \text{tr} \Sigma_1 \Sigma_2^{-1} = \int_{-\pi}^{\pi} \frac{p_2(\omega)}{p_1(\omega)} \frac{d\omega}{2\pi} \quad (11)$$

が成立する⁽¹⁰⁾。時間軸上のデータ \$(x_{t+1}, x_{t+2}, \dots, x_{t+n})\$ の各々を多変量 \$(y_1, y_2, \dots, y_n)\$ へ写像することによって、式(8)の右辺のデータ長 \$T\$ を除去した形で表現できる。以下断らないかぎり、\$I_T(1:2)\$ の式より \$T\$ を省く。式(10)、(11)を使い、周波数上での K-L 情報量は

$$\begin{aligned} I_\omega(1:2) &= \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{p_2(\omega)}{p_1(\omega)} + \log \frac{p_2(\omega)}{p_1(\omega)} - 1 \right) \frac{d\omega}{2\pi} \\ &\quad + \frac{1}{2} \int_{-\pi}^{\pi} \frac{|D(\omega)|^2}{p_2(\omega)} \cdot \frac{d\omega}{2\pi} \end{aligned} \quad (12)$$

ここで

$$D(\omega) = \int_{-\infty}^{\infty} \delta(t) e^{-j\omega t} dt$$

は平均値関数

$$\delta(t) = \mu_1(t) - \mu_2(t)$$

のフーリエ変換を表す⁽¹¹⁾。本研究の対象とする脳波時系列では、交流成分のみに注目し、直流成分が除去される。従って、常に次式の成立されることを前提条件としている。

$$\delta(t) = \mu_1(t) - \mu_2(t) = 0$$

このとき、K-L 情報量は

$$I_\omega(1:2) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{p_2(\omega)}{p_1(\omega)} + \log \frac{p_2(\omega)}{p_1(\omega)} - 1 \right) \frac{d\omega}{2\pi} \quad (13)$$

となる。

(A) 二つの時系列の周波数構造がほとんど等しい場合の K-L 情報量

K-L 情報量を表す式(13)で時系列のスペクトル密度 \$p_1(\omega)\$ が、他の系列のスペクトル密度 \$p_2(\omega)\$ に近い場合は、式(13)の第 2 項目は

$$\begin{aligned} \log \left\{ 1 + \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right) \right\} \\ = \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right) - \frac{1}{2} \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right)^2 + \text{高次の項} \end{aligned} \quad (14)$$

のようにテイラー展開される。式(14)を式(13)に代入し、

$$\int_{-\pi}^{\pi} \left(\frac{p_2(\omega)}{p_1(\omega)} + \log \frac{p_2(\omega)}{p_1(\omega)} - 1 \right) \frac{d\omega}{2\pi}$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right)^2 \frac{d\omega}{2\pi} \quad (15)$$

いま、与えられた第1の時系列を自己回帰モデルとして表現すると、そのパワースペクトル密度は

$$p_1(\omega) = \frac{\sigma_1^2}{|A_1(e^{j\omega})|^2} \quad (16)$$

で与えられる。式(16)に対して、対比する他の時系列の自己回帰モデルのパワースペクトル密度は

$$p_2(\omega) = \frac{|E_2(e^{j\omega})|^2}{|A_2(e^{j\omega})|^2} \quad (17)$$

とする。ここで、二つの時系列の周波数成分がほとんど等しい構造を有している。すなわち

$$|A_1(e^{j\omega})|^2 \simeq |A_2(e^{j\omega})|^2 \quad (18)$$

と仮定すれば、式(15)の関数項が

$$\int_{-\pi}^{\pi} \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right)^2 \frac{d\omega}{2\pi} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 - \frac{E_2^2(e^{j\omega})}{\sigma_1^2} \right)^2 d\omega$$

となる。更に式(16)~(18)より

$$\begin{aligned} \frac{p_2(\omega)}{p_1(\omega)} \sigma_1^2 &= |E_2(e^{j\omega})|^2 \\ &\simeq r(0) + \sum_{k=1}^m r(k) \cos(\omega k) \end{aligned} \quad (19)$$

が成立する。ここで $\{r(k)\}$ ($k=0, 1, 2, \dots, m$) は自己回帰モデルの残差系列の共分散関数となる。

式(15)の関数項は式(19)の関係を使うと、

$$(\sigma_1^2 - r(0))^2 + 2 \sum_{k=1}^m r(k)^2 \quad (20)$$

となる。式(19)、(20)で m は次数を表す。すなわち、二つの時系列の周波数構造がほとんど等しい場合、K-L情報量は Bodenstein らのスペクトル誤差量⁽⁵⁾と等しくなる。

(B) 二つの時系列の周波数構造が異なる場合のK-L情報量

脳波時系列の小区間では、定常正規過程とみなされることから⁽⁸⁾、自己回帰モデルの残差系列のパワースペクトルは平坦とみなされる。従って式(17)でも、その分子が

$$|E_2(e^{j\omega})|^2 = \sigma_2^2$$

とおけることから、K-L情報量は

$$\begin{aligned} I_T(1:2) &= \frac{1}{2} \left\{ \left(\frac{\sigma_1}{\sigma_2} \right)^2 \left\{ 1 + \int_{-\pi}^{\pi} \frac{|A_2(e^{j\omega}) - A_1(e^{j\omega})|^2 d\omega}{|A_1(e^{j\omega})|^2 2\pi} \right\} \right. \\ &\quad \left. - \log \left(\frac{\sigma_1}{\sigma_2} \right)^2 - 1 \right\} \end{aligned} \quad (21)$$

となる⁽¹⁾。ここでは、二つの時系列の周波数構造が異なる場合、すなわち式(18)の仮定はしていない。このと

き、式(21)の中括弧の項は、残差系列のゆう度比、すなわち

$$1 + \int_{-\pi}^{\pi} \frac{|A_2(e^{j\omega}) - A_1(e^{j\omega})|^2 d\omega}{|A_1(e^{j\omega})|^2 2\pi} = \frac{\delta}{\alpha} \quad (22)$$

となる。ここで

$$\alpha = \sum_{n=-M}^M r_{1\alpha}(n) r_{1x}(n)$$

$$\delta = \sum_{n=-M}^M r_{2\alpha}(n) r_{1x}(n)$$

である。 $\{r_{1\alpha}(n)\}$ は i 系列の時系列の自己回帰モデルの係数の自己相関係数列であり、 $\{r_{ix}(n)\}$ は i 系列の時系列の自己相関係数列である。更に

$$\log \frac{\sigma_1^2}{\sigma_2^2} = \log \left(1 + \frac{\sigma_1^2 - \sigma_2^2}{\sigma_2^2} \right) \quad (23)$$

とおき、テイラー展開し、式(21)に代入すれば

$$I_T(1:2) = \frac{1}{2} \left\{ \frac{\sigma_1^2}{\sigma_2^2} \left(\frac{\delta}{\alpha} - \frac{\sigma_1^2}{\sigma_2^2} \right) \right\} \quad (24)$$

となる。一方、自己回帰モデルによる残差系列の分散は、時系列の共分散の系列 $\{r_{ix}(n)\}$ と自己回帰モデルの係数 $\{\alpha(n)\}$ を使って、 p 次のモデルなら

$$\sigma_i^2 = r_{ix}(0) + \sum_{n=1}^p \alpha(n) r_{ix}(n) \quad (25)$$

と表される。すなわち、K-L情報量の式(24)は自己回帰モデルの係数と、時系列の共分散の系列で表現できる。更に、定常正規過程の自己回帰モデルへの適合が十分ならば、残差系列の分散のオーダーと値が等しい。すなわち

$$\sigma_1^2 \simeq \sigma_2^2 \quad (26)$$

の場合

$$I_T(1:2) = \frac{1}{2} \left(\frac{\delta}{\alpha} - 1 \right) \quad (27)$$

となる。式(27)は板倉らが導出した最ゆう法によるゆう度比⁽⁶⁾の結果と等価である。

3. K-L情報量, K-L divergence と Bhattacharyya 距離との比較

二つの正規密度関数が $N(\mu_1, \Sigma_1)$ 及び $N(\mu_2, \Sigma_2)$ のとき、Bhattacharyya 距離 (B-距離) は次式で⁽⁹⁾与えられる。

$$\frac{1}{8} \mu^t \Sigma^{-1} \mu + \frac{1}{2} \log \left(\frac{|\Sigma|}{|\Sigma_1|^{1/2} |\Sigma_2|^{1/2}} \right) \quad (28)$$

ここで、 $\mu = \mu_1 - \mu_2$ 及び

$$\Sigma = \frac{1}{2} \Sigma_1 + \frac{1}{2} \Sigma_2$$

が成立する。さて、B-距離を時間領域で表現すると

$$B_T(1:2) = \frac{1}{T} \left\{ \frac{1}{8} \delta^2 \Sigma^{-1} \delta + \frac{1}{2} \log \left(\frac{|\Sigma|}{|\Sigma_1|^{1/2} |\Sigma_2|^{1/2}} \right) \right\} \quad (29)$$

ここで、 $\Sigma = \frac{1}{2} \Sigma_1 + \frac{1}{2} \Sigma_2$

を示し、且つ Σ_j 及び δ は式(9)の定義と同じである。

ここで式(10)の結果を使い、B-距離を周波数領域で表現すると、

$$B_\omega(1:2) = \frac{1}{2} \int_{-\pi}^{\pi} \log \frac{p(\omega)}{p_1(\omega)^{1/2} p_2(\omega)^{1/2}} \frac{d\omega}{2\pi} + \frac{1}{8} \int_{-\pi}^{\pi} \frac{|D(\omega)|^2}{p(\omega)} \frac{d\omega}{2\pi} \quad (30)$$

ここで、 $p(\omega)$ 、 $p_1(\omega)$ 、 $p_2(\omega)$ は Σ 、 Σ_1 、 Σ_2 の各々に対応するパワースペクトル密度を表す。 $D(\omega)$ は式(12)のそれと等しい。さて、与えられた時系列を自己回帰モデルで表現すると、脳波時系列の $\delta(t) = 0$ の条件および、

$$\int_{-\pi}^{\pi} \log \{ |A_i(e^{j\omega})|^2 \} \frac{d\omega}{2\pi} = 0 \quad (31)$$

より¹²⁾、式(30)は

$$B_\omega(1:2) = \frac{1}{2} \int_{-\pi}^{\pi} \log \frac{\sigma^2}{\sigma_1 \sigma_2} \frac{d\omega}{2\pi} \quad (32)$$

ここで、 σ^2 、 σ_1^2 、 σ_2^2 は各々自己回帰モデルの残差系列の分散を表す。式(32)は

$$B_\omega(1:2) = \frac{1}{2} \log \frac{\sigma^2}{\sigma_1 \sigma_2} \quad (33)$$

となる。ここで σ^2 、 σ_1^2 、 σ_2^2 は各々

$$\begin{aligned} \sigma^2 &= r(0) + \Sigma \alpha(n) \sigma_x(n) \\ \sigma_1^2 &= r_1(0) + \Sigma \alpha_1(n) r_{1x}(n) \\ \sigma_2^2 &= r_2(0) + \Sigma \alpha_2(n) r_{2x}(n) \end{aligned}$$

で与えられる。

はじめに、K-L情報量、K-L divergence、及びB-距離について、時系列の振幅の差の感度の比較を行う。これらの距離尺度は時間領域で共分散行列で表されるが、いま系列の自己相関数値はほとんど零と仮定すると、K-L情報量は

$$I_T(1:2) = \frac{1}{2} \left(\log \frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} - 1 \right) \quad (34)$$

となる。ここで σ_i^2 は i 時系列の分散である。更にK-L divergence は $I_T(1:2)$ と $I_T(2:1)$ の両者の和として計算される。いまK-L divergence を $J_T(1:2)$ とおくと、

$$\begin{aligned} J_T(1:2) &= I_T(1:2) + I_T(2:1) \\ &= \frac{1}{2} \left\{ \left(\frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2} \right) - 2 \right\} \end{aligned} \quad (35)$$

又、B-距離は

$$\begin{aligned} B_T(1:2) &= \frac{1}{2} \left\{ \log \left\{ \frac{1}{2} (\sigma_1'^2 + \sigma_2'^2) \right\} - \frac{1}{2} \log \sigma_1'^2 \right. \\ &\quad \left. - \frac{1}{2} \log \sigma_2'^2 \right\} \end{aligned} \quad (36)$$

式(34)、(35)及び式(36)の距離尺度を比較するため、いま

$$\sigma_i'^2 = K \sigma_i^2 \quad (37)$$

とおくと、式(34)、(35)及び式(36)は各々次式で与えられる。

$$\left. \begin{aligned} I_T(1:2) &= \frac{1}{2} \left(\log K + \frac{1}{K} - 1 \right) \\ J_T(1:2) &= \frac{1}{2} \left(\frac{1}{K} + K - 2 \right) \\ B_T(1:2) &= \frac{1}{2} \left[\log \frac{1}{2} (1+K) - \frac{1}{2} \log K \right] \end{aligned} \right\} \quad (38)$$

となる。式(38)の各々の距離尺度を横軸に K の値をとり、プロットしたものを図1に示してある。

ここで時系列の振幅の変化に対する感度を調べるため式(38)の距離尺度を K で微分した場合の値を表1に示す。

表1は K の値に対する各距離尺度の感度として、表中の絶対値の大きなものがより高い感度であることを示す。

さて、B-距離は、それを一般化した Chernoff 距離の特別な場合とみなすことができる。Chernoff 距離

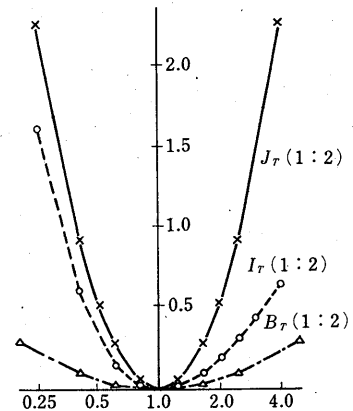


図1 $J_T(1:2)$ 、 $I_T(1:2)$ 及び $B_T(1:2)$ による振幅の変化に対する各々の値

Fig.1-Values to the change of amplitude component calculated by $J_T(1:2)$, $I_T(1:2)$ and $B_T(1:2)$.

表 1 K の値に対する各尺度の感度を表す微係数の値

K	$\frac{dI_T(1:2)}{dK}$	$\frac{dJ_T(1:2)}{dK}$	$\frac{dB_T(1:2)}{dK}$
0.25	-6.00	-7.50	-0.60
0.50	-1.00	-1.50	-0.17
1.00	0.00	0.00	0.00
2.00	0.13	0.28	0.03
4.00	0.09	0.47	0.04

(C-距離) は式(39)で定義⁽⁹⁾される。

$$C(1:2;\tau) = \frac{1}{2} \tau(1-\tau) \mu^t \Sigma^{-1} \mu + \frac{1}{2} \log \left(\frac{|\Sigma|}{|\Sigma_1|^{1-\tau} |\Sigma_2|^\tau} \right) \quad (39)$$

ここで、 $\mu = \mu_1 - \mu_2$ であり、

$$\Sigma = (1-\tau) \Sigma_1 + \tau \Sigma_2, \quad 0 \leq \tau \leq 1$$

である。さて、いままでと同じ条件、 $\mu = 0$ の条件で、式(39)の時間領域での値を考える。時系列に相関のない場合、式(39)を代入すると、

$$C_T(1:2;\tau) = \frac{1}{2} \log \frac{1+\tau(K-1)}{K^\tau} \quad (40)$$

となる。式(40)を最大にするような τ の最適値を τ_{op} とすると、各 K の値に対する τ_{op} とそのときの C-距離、 $C_T(1:2;\tau_{op})$ は表 2 のように与えられる。

表 2 より、 τ_{op} のときの C-距離の値は、B-距離の値とほとんど等しいことが分かる。

次に K-L 情報、K-L divergence、B-距離の尺度が時系列の周波数構造の変化に対し、いかなる感度の変化を示すかを明らかにする。二つの時系列の振幅値については同じ、すなわち

$$\sigma'_1 = \sigma'_2$$

と仮定して周波数に対する変化の感度を調べる。いま、 S_i を時系列 i の相関行列、 $|S_i|$ をその行列式とすると、

K-L 情報量は $\sigma'_1 = \sigma'_2$ の条件の下で、

$$I_T(1:2) = \frac{1}{2} \log \frac{|S_2|}{|S_1|} + \frac{1}{2} \text{tr} \{S_1 S_2^{-1} - I\} \quad (41)$$

となる。ここで I は恒等行列である。同様に K-L di-

表 2 K の値に対する Chernoff 距離尺度の最適な τ の値

K	τ_{op}	$C_T(1:2;\tau_{op})$	$B_T(1:2)$
0.20	0.629	0.313	0.294
0.40	0.575	0.104	0.101
0.60	0.542	0.033	0.032
0.80	0.519	0.006	0.059
1.25	0.481	0.006	0.006
2.50	0.458	0.104	0.101

vergence は

$$J_T(1:2) = \frac{1}{2} \text{tr} \{S_1 S_2^{-1} + S_2 S_1^{-1} - 2I\} \quad (42)$$

更に B-距離は

$$B_T(1:2) = \frac{1}{2} \log \frac{\left| \frac{1}{2} S_1 + \frac{1}{2} S_2 \right|}{|S_1|^{\frac{1}{2}} |S_2|^{\frac{1}{2}}} \quad (43)$$

となる。時系列の周波数の変化に対する感度を式(41)、(42)及び式(43)で比較するため、1 次自己回帰モデルで表される二つの時系列を考える。その一つの時系列の相関行列を S_1 、他の一つを S_2 とおいて、各々次式で与えられるものとする。

$$S_1 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 1 & \rho' \\ \rho' & 1 \end{pmatrix} \quad (44)$$

ここで ρ 及び ρ' は各々の時系列の系列相関係数である。ここで

$$\rho' = K\rho \quad (45)$$

とおいて、式(44)を式(41)~(43)へ代入し、それらの値を図 2 に示す。

さて、相関行列の変化に対する距離尺度の感度は、式(45)を式(41)~(43)へ代入し、 K で微分した値で与えられる。 $\rho = 0.2$ とおいたときの、各尺度の微分値を表 3 に示した。表 3 で、絶対値の大きいもの程、感度が良

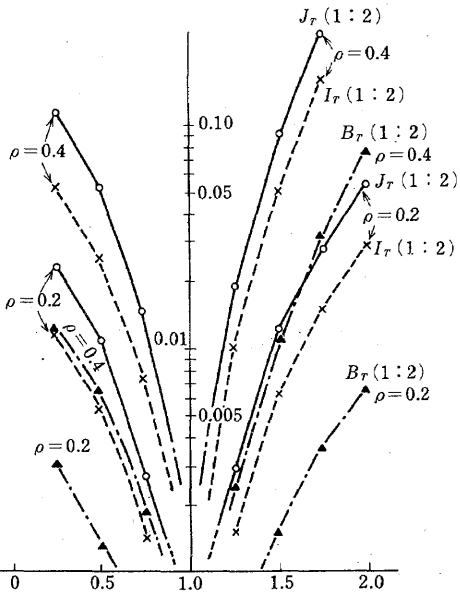


図 2 $J_T(1:2)$ 、 $I_T(1:2)$ 及び $B_T(1:2)$ による系列相関係数値の変化による各々の値

Fig.2-Values of measures to the change of correlation values.

表3 Kの値に対する各尺度の感度を表す微係数の値

K	$\frac{dI_T(1:2)}{dK}$	$\frac{dJ_T(1:2)}{dK}$	$\frac{dB_T(1:2)}{dK}$
	dK	dK	dK
0.25	-0.030	-0.062	-0.008
0.50	-0.021	-0.042	-0.005
0.75	-0.011	-0.022	-0.003
1.00	0.000	0.000	0.000
1.25	0.012	0.024	0.003
1.50	0.026	0.051	0.006

いことを示す。

4. K-L情報量によるクラスタリング

本章ではK-L情報量, 特に式(22)の最ゆう距離による睡眠脳波パターンの階層的クラスタリングを行う。前もって医師の視察により各グループ20個ずつ, 全体で7グループのデータ⁽¹⁾を取り上げ, 下記のように印した。

- ① 高振幅δ波(D), ② 低振幅δ波(H)
- ③ θ波(T), ④ Stage 1の波形(L)
- ⑤ Stage REMの波形(R),
- ⑥ α波(A), ⑦ 紡錘波(S)

以上のデータが最ゆう距離により, どの程度まで分類できるかを階層的クラスタリングを行い, そのデンドログラムの結果の一部を図3に示した。

図3が示すように, 十分なグループ化がなされていない。最ゆう距離は二つのパワースペクトル密度の差の全周波数帯域の積分値である。そこで局所的なスベ

表4 脳波パターンの周波数帯域への分割

波名	帯域	波名	帯域
δ_0 波	$0 \leq \delta_0 < 1.5$	α_2 波	$10 \leq \alpha_2 < 11.5$
δ_1 波	$1.5 \leq \delta_1 < 3.5$	α_3 波	$11.5 \leq \alpha_3 < 15$
θ_1 波	$3.5 \leq \theta_1 < 5.5$	β_1 波	$15 \leq \beta_1 < 18$
θ_2 波	$5.5 \leq \theta_2 < 7.5$	β_2 波	$18 \leq \beta_2 < 21$
α_1 波	$7.5 \leq \alpha_1 < 10$	β_3 波	$21 \leq \beta_3 < 25$

クトル密度の差を検出するため, 表4の10個の周波数帯域を各変数の領域とする正準分析を行った。

表4の波名に対する変量は, まずK-L情報量の式(15)を式(46)の右辺の離散量で表す。

$$\int_{-\pi}^{\pi} \left(\frac{p_2(\omega) - p_1(\omega)}{p_1(\omega)} \right)^2 \frac{d\omega}{2\pi} \simeq \frac{1}{10} \sum_{i=1}^{10} \left(\frac{p_{2i}(\omega) - p_{1i}(\omega)}{p_{1i}(\omega)} \right)^2 \quad (46)$$

ここで,

$$p_{ji}(\omega) \simeq \sum_{l=1}^{B_i} \log p_{jil}(\omega)$$

とおいた。

l は0.25 Hzでサンプリングを行い, 例えば B_1 は δ_0 波帯で8個のサンプル値となり $B_1=8$ となる。以下同様。従って各変量 x_i は

$$x_i = \frac{(p_{Bi} - \overline{p_{Bi}})^2}{\overline{p_{Bi}^2}}, \quad (i=1 \sim 10) \quad (47)$$

となる。 $\overline{p_{Bi}}$ は B_i 波帯のデータの平均値である。

正準判別分析は, 以上の10個の変量 $\{x_i\}$ ($i=1 \sim 10$) に対し, 重みを与えて特徴づけを行い, その上で

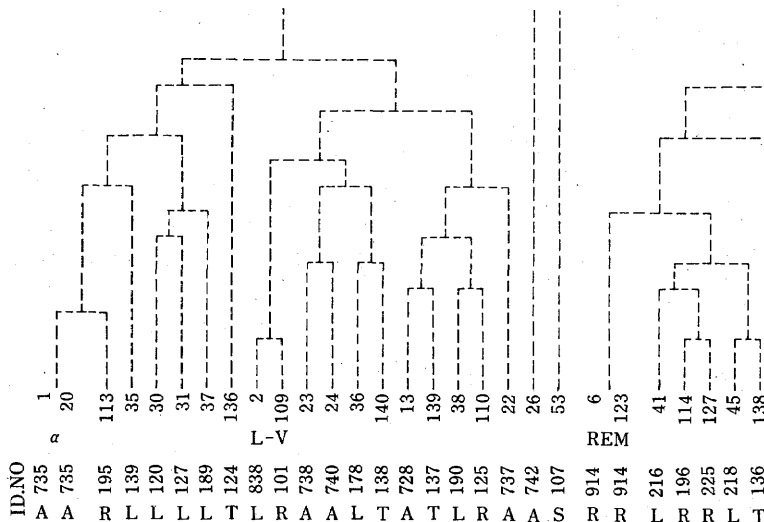


図3 クラスタリングによる樹状図(数字はデータ番号を示し, アルファベットはグループを示す)

Fig.3-Dendrogram by clustering method.

グループの母集団間の変動 Σ_B を大きくすることである。各変量の重みベクトル

$$W = (w_1, w_2, \dots, w_{10})$$

に対し、パラメータの変換を行うと、

$$y = W^t x = (w_1 x_1, w_2 x_2, \dots, w_{10} x_{10})$$

となる。実際には

$$y = W^t (x - \bar{x}) \quad (48)$$

とする。このような変換のもとでグループ間の母集団の変動 Σ_B と各グループ内の母集団の変動 Σ_W は、各々

$$W^t \Sigma_W W \text{ と } W^t \Sigma_B W$$

となる。規格化条件 $W^t \Sigma_W W = \text{定数}$ とおいて、

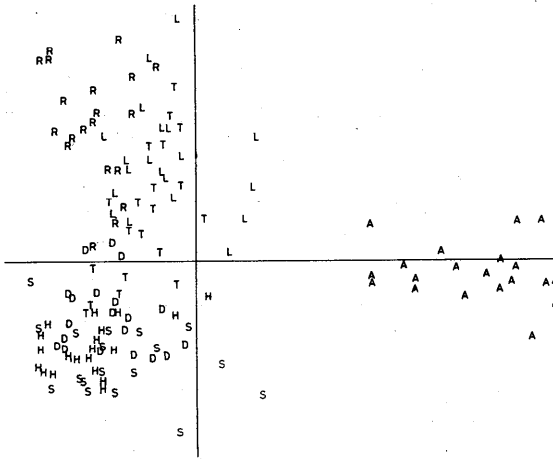


図4 正準分析法による睡眠脳波パターンのグループ化

Fig.4-Clustering by canonical analysis of sleeping EEG patterns.

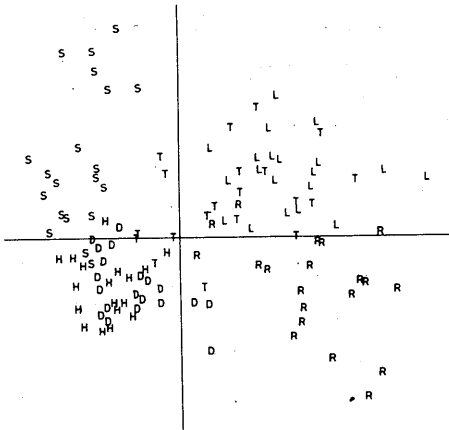


図5 第2回目の正準分析法によるグループ化
Fig.5-Clustering by canonical analysis.

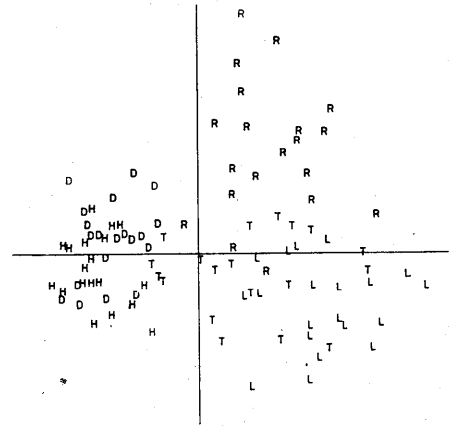


図6 第3回目の正準分析法によるグループ化
Fig.6-Clustering by canonical analysis.

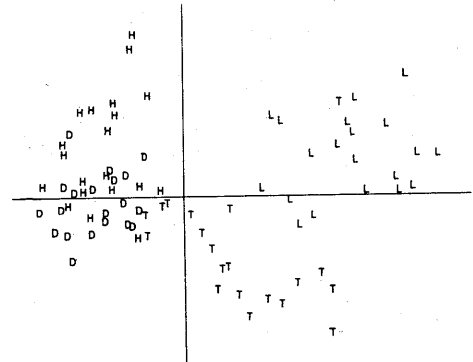


図7 第4回目の正準分析法によるグループ化
Fig.7-Clustering by canonical analysis.

($W^t \Sigma_B W / W^t \Sigma_W W$) の比を最大にするように W をラグランジュの乗数法によって計算し、式(48)を各データごとに算出した。図4には7グループのすべての脳波パターンに正準判別分析を適用し、第1及び第2の正準軸となる重みを求め、各々の重みを使い式(48)によって、各データをプロットしたものである。図4は α 波グループ (A) が分離できることを示す。すなわち、第1の正準軸である横軸に、あるしきい値を設定して、 α 波グループ (A) と他のグループを分離できる。次に α 波グループ (A) を除いたデータで、再び正準判別を行うと図5となり紡錘波グループ (S) が分離できる。

図6から Stage REMに特有な波のグループ (R) が分離できる。図7から Stage 1に特有な波のグループ (L) が分離できる。同様に図8から θ 波のグループ (T) が分離できる。次に図9から高振幅- δ 波のグループ

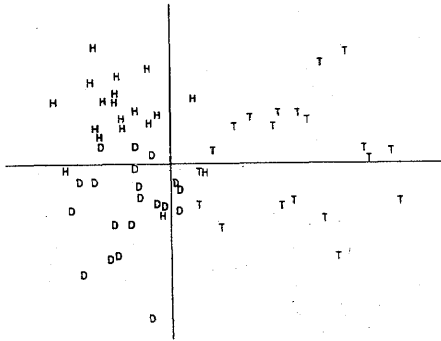


図8 第5回目の正準分析法によるグループ化
Fig.8-Clustering by canonical analysis.

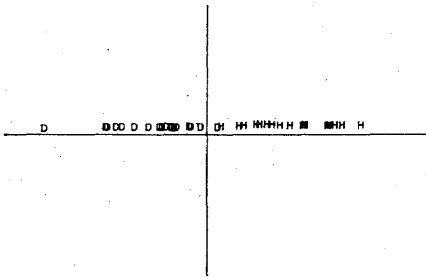


図9 第6回目の正準分析法によるグループ化
Fig.9-Clustering by canonical analysis.

(D)と低振幅- δ 波のグループ(H)が分離できる。

5. む す び

非定常時系列の分類のために、距離尺度としてK-L情報量を導入し、従来の距離尺度が統一的に表されることを示した。次にK-L情報量、K-L divergenceが他の尺度のB-距離およびC-距離と比較し、時系列の構造の変化を検出する際に、感度の良い尺度であることを示した。最後にK-L情報量を睡眠脳波時系列の分類へ適用したところ、K-L情報量の周波数帯別の値を変数とする正準判別分析法が有効であることを示した。本研究の次の段階は、本論文のK-L情報量を脳波による睡眠段階の判定のアルゴリズムに組み入れることである。

謝辞 本研究に関し、日ごろ御指導頂く、岐阜大学医学部渡辺教授および名古屋大学医学部寺嶋博士に深く感謝致します。

文 献

- (1) 石井, 岩田, 鈴村, 杉本: "自己回帰モデルによる睡眠脳波時系列の分類", 信学論(D), J61-D, 12, pp.901-908(昭53-12).
- (2) 赤池, 中川: "ダイナミックシステムの統計的解析と制御", サイエンス社(昭52).
- (3) 石井, 鈴村: "自己回帰モデルの次数の推定について", 信学論(A), J60-A, 6, pp.513-520(昭52-06).
- (4) 石井, 岩田, 鈴村: "情報量による自己回帰モデルの評価と非正規過程への適用", 信学論(A), J61-A, 1, pp.25-32(昭53-01).
- (5) Bodenstein, G. and Praetorius, H. M.: "Feature extraction from the EEG by adaptive segmentation", Proc. IEEE, 65, 5, pp.642-652(1977).
- (6) 板倉, 斉藤: "統計的手法による音声スペクトル密度とホルモント周波数の推定", 信学論(A), 53-A, 1, pp.35-42(昭45-01).
- (7) Kullback, S. and Leibler, R. A.: "On information and sufficiency", Annals of Math. Statistics, 22, pp.79-86(1951).
- (8) Sugimoto, H. et al.: "On the stationarity and normality of EEG data during sleep stages", Comput. Program Biomed., 8, pp.224-234(1978).
- (9) Kailath, T.: "The divergence and Bhattacharyya distance measures in signal selection", IEEE Trans. Commun. Technol., COM-15, 1, pp.52-60(1967).
- (10) Grenander and Szego: "Toeplitz Forms and Their Application", Univ. of Calif., p.58(1958).
- (11) Shunway, R. H. and Unger, N.: "Linear discriminant functions for stationary time series", Ame. Stat. Association, 69, pp.948-956(1974).
- (12) Gray, A. H. and Markel, J.: "Distance measures for speech processing", IEEE Trans. Acoust., Speech & Signal Process. ASSP-24, 5, pp.380-391(1976).

(昭和55年6月26日受付)