Pergamon

Computers Math. Applic. Vol. 27, No. 7, pp. 15–21, 1994

0898-1221/94 $6.00 + 0.00

0898-1221(94)E0022-C

# Hellinger Transform of Gaussian Autoregressive Processes

I. FAZEKAS
Institute of Mathematics and Informatics, Kossuth University
P.O. Box 12, H-4010 Debrecen, Hungary

*(Received August 1993; accepted September 1993)*

**Abstract**—The limit of the Hellinger transform of measures related to multidimensional stationary Gaussian autoregressive processes is obtained

## 1. INTRODUCTION AND MAIN RESULT

Let $P_1, \dots, P_N$ be probability measures defined on a common measurable space $(\Omega, \mathcal{F})$. Let $P$ be a $\sigma$-finite measure dominating each $P_j$. Then the Hellinger transform of $P_1, \dots, P_N$ is defined by

$$\mathbb{H}_\alpha(P_1, \dots, P_N) = \int_\Omega \prod_{j=1}^N \left(\frac{dP_j}{dP}\right)^{\alpha_j} dP, \tag{1}$$

where $\alpha = (\alpha_1, \dots, \alpha_N)$, $\alpha_1 > 0, \dots, \alpha_N > 0$, $\sum_{j=1}^N \alpha_j = 1$.

We remark that for $N = 2$ the quantity $\mathbb{H}_\alpha(P_1, P_2)$ is called the Hellinger integral of measures $P_1$ and $P_2$. The Hellinger integral is widely used to study absolute continuity of measures (see e.g., [1–3]).

Let $\mathcal{N}_p(b, R)$ denote the $p$-dimensional normal distribution with mean vector $b$ and covariance matrix $R$. Let $P_j$ denote the distribution $\mathcal{N}_p(0, R_j)$, $j = 1, \dots, N$. Suppose that $R_j$ is invertible for $j = 1, \dots, N$ and $\sum_{j=1}^N \alpha_j R_j^{-1}$ is also invertible. Then the Hellinger transform of $P_1, \dots, P_N$ is the following

$$\mathbb{H}_\alpha(P_1, \dots, P_N) = \left\{ \prod_{j=1}^N (\det R_j)^{\alpha_j} \det \left( \sum_{j=1}^N \alpha_j R_j^{-1} \right) \right\}^{-1/2}. \tag{2}$$

Let $\{\xi_t, t = 0, \pm 1, \pm 2, \dots\}$ be a (multidimensional) Gaussian process with mean 0. Let $H_j$ be the hypothesis that $\{\xi_t\}$ has covariance matrix $C_j$. Let $P_j^{(t)}$ denote the distribution of $\{\xi_0, \dots, \xi_t\}$ if $H_j$ is true, $j = 1, \dots, N$. Then

$$\mathbb{H}_\alpha(P_1^{(t)}, \dots, P_N^{(t)}) = \left\{ \prod_{j=1}^N \left( \det \left( [C_j]_0^t \right) \right)^{\alpha_j} \det \left( \sum_{j=1}^N \alpha_j \left( [C_j]_0^t \right)^{-1} \right) \right\}^{-1/2}, \tag{3}$$

I wish to thank F. Liese and Y. S. Sathe for helpful discussions.
This research was supported by the Hungarian Foundation for Scientific Researches under Grant No. OTKA-T4047/1992 and Grant No. OTKA-1648/1991.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

where $[C_j]_0^t$ denotes an appropriate 'section' of $C_j$ (i.e., $[C_j]_0^t = \left((C_j)_{k,l}\right)_{k,l=0}^t$). We suppose that the inverses involved exist.

Our aim is to study $\lim_{t\to\infty}(1/(t+1))\log \mathbb{H}_\alpha(P_1^{(t)},\dots,P_N^{(t)})$ when the process $\{\xi_t\}$ is stationary (under each hypothesis $H_j$, $j=1,\dots,N$). If $N=2$ then $\lim_{t\to\infty}(1/(t+1))\log \mathrm{H}_{\alpha,1-\alpha}(P_1^{(t)},P_2^{(t)})$ is Rényi's information (see [4]). Suppose that the (multidimensional) stationary process $\{\xi_t\}$ has spectral density (matrix) function $f_j(\lambda)$, $\lambda\in[-\pi,\pi]$, under hypothesis $H_j$

$$(C_j)_{k,l} = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \exp(i(k-l)\lambda) f_j(\lambda)\,d\lambda, \qquad \forall\, k,l. \tag{4}$$

It is known that a stationary process with spectral density $f_j(\lambda)f_h(\lambda)$ has covariance matrix $C_jC_h$. Moreover, if $f_j(\lambda)$ is invertible for all $\lambda$ then a stationary process with spectral density $(f_j(\lambda))^{-1}$ has covariance matrix $(C_j)^{-1}$.

We need the following result due to [5].

LEMMA 1. *Suppose that $f_j(\lambda)$ is positive definite for $\lambda\in[-\pi,\pi]$. Then*

$$\lim_{t\to\infty}\frac{1}{t+1}(\log\det([C_j]_0^t)) = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\log\det(f_j(\lambda))\,d\lambda. \tag{5}$$

However, the above mentioned facts cannot be used directly to calculate

$$\lim_{t\to\infty}\frac{1}{t+1}\log \mathbb{H}_\alpha\left(P_1^{(t)},\dots,P_N^{(t)}\right)$$

for arbitrary stationary processes because formula (3) involves $\left([C_j]_0^t\right)^{-1}$ and not $[C_j^{-1}]_0^t$.

For multidimensional Gaussian AR(1) processes we have the following theorem.

THEOREM 1. Let $\{y_t\}$ be an $m$-dimensional stationary Gaussian AR(1) process with zero mean

$$y_{t+1} = Ay_t + \varepsilon_{t+1}, \quad t = 0,\pm1,\pm2,\dots, \tag{6}$$

where $\{\varepsilon_t\}$ is a sequence of independent standard Gaussian random vectors. Let $H_j$ be the hypothesis that $A = A_j$, $j=1,\dots,N$. Suppose that the eigenvalues of $A_j$ are inside the unit circle for all $j$. Let

$$f_j(\lambda) = \left([I - A_j\exp(i\lambda)]^\top [I - A_j\exp(i\lambda)]\right)^{-1} \tag{7}$$

be thespectral density of the process and let $P_j^{(t)}$ denote the distribution of $\{y_0,\dots,y_t\}$ if $H_j$ is true, $j=1,\dots,N$. Then

$$\begin{aligned}\lim_{t\to\infty}&\frac{1}{t+1}\log \mathrm{H}_\alpha(P_1^{(t)},\dots,P_N^{(t)})\\ &= \frac{-1}{4\pi}\left[\sum_{j=1}^N \alpha_j\int_{-\pi}^{+\pi}\log\det f_j(\lambda)\,d\lambda + \int_{-\pi}^{+\pi}\log\det\sum_{j=1}^N \alpha_j (f_j(\lambda))^{-1}\,d\lambda\right],\end{aligned} \tag{8}$$

where $\alpha = (\alpha_1,\dots,\alpha_N)$, $\alpha_1>0,\dots,\alpha_N>0$, $\sum_{j=1}^N\alpha_j = 1$.

Here, and in what follows, $^\top$ denotes the transposed of a matrix while $I$ denotes the unit matrix (of appropriate size). If $N=2$, Theorem 1 is known for any multidimensional stationary Gaussian process (see [6]).

The proof of Theorem 1 is given in Section 3. In Section 2, a stronger version of Theorem 1 is proved for scalar prosesses (Theorem 2).

## 2. SCALAR AR(1) PROCESSES

Let $\{\xi_t,\ t = 0, \pm 1, \pm 2, \dots\}$ be a real stationary Gaussian AR(1) process. Suppose that under hypothesis $H_j$

$$\xi_{t+1} = a_j\, \xi_t + \varepsilon_{t+1}, \qquad t = 0, \pm 1, \pm 2, \dots, \tag{9}$$

where $\{\varepsilon_t\}$ is a sequence of independent standard Gaussian random variables, $|a_j| < 1$. Suppose that the mean of the process is 0. The covariance matrix and its inverse are of the form (see [7, p. 131])

$$C_j = \frac{1}{1-a_j^2} \begin{pmatrix} \ddots & \ddots & \ddots & & \\ \ddots & 1 & a_j & a_j^2 & \\ \ddots & a_j & 1 & a_j & \ddots \\ & a_j^2 & a_j & 1 & \ddots \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \quad C_j^{-1} = \begin{pmatrix} \ddots & \ddots & \ddots & & \\ \ddots & 1+a_j^2 & -a_j & 0 & \\ \ddots & -a_j & 1+a_j^2 & -a_j & \ddots \\ & 0 & -a_j & 1+a_j^2 & \ddots \\ & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{10}$$

The covariance matrix of $\{\xi_0, \dots, \xi_t\}$ and the inverse covariance matrix have the form

$$\left[C_j\right]_0^t = \frac{1}{1-a_j^2} \begin{pmatrix} 1 & a_j & \cdots & a_j^{t-1} & a_j^t \\ a_j & 1 & \cdots & a_j^{t-2} & a_j^{t-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_j^{t-1} & a_j^{t-2} & \cdots & 1 & a_j \\ a_j^t & a_j^{t-1} & \cdots & a_j & 1 \end{pmatrix}, \tag{11}$$

$$\left(\left[C_j\right]_0^t\right)^{-1} = \begin{pmatrix} 1 & -a_j & 0 & \cdots & 0 \\ -a_j & 1+a_j^2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+a_j^2 & -a_j \\ 0 & \cdots & 0 & -a_j & 1 \end{pmatrix}. \tag{12}$$

It is easy to see that $\det\left(\left[C_j\right]_0^t\right)^{-1} = 1 - a_j^2$, but according to (3), we need

$$d = \det\left(\sum_{j=1}^{N} \alpha_j \left(\left[C_j\right]_0^t\right)^{-1}\right).$$

By (12)

$$d = \det\left[C\right]_0^{t-2} - 2a^2 \det\left[C\right]_0^{t-3} - 4a^4 \det\left[C\right]_0^{t-4}, \tag{13}$$

where $C = \sum_{j=1}^{N} \alpha_j C_j^{-1}$ and $a = \sum_{j=1}^{N} \alpha_j a_j$.

The spectral density of the stationary process $\{\xi_t\}$ defined by (9) is

$$f_j(\lambda) = \left(1 - 2a_j \cos\lambda + a_j^2\right)^{-1}. \tag{14}$$

The stationary process with spectral density

$$g_j(\lambda) = \frac{1}{f_j(\lambda)} = 1 - 2a_j \cos\lambda + a_j^2 \tag{15}$$

has covariance matrix $C_j^{-1}$. Furthermore, the stationary process with spectral density

$$g(\lambda) = \sum_{j=1}^{N} \alpha_j\, g_j(\lambda) = \sum_{j=1}^{N} \frac{\alpha_j}{f_j(\lambda)} \tag{16}$$

has covariance matrix $C = \sum_{j=1}^{N} \alpha_j C_j^{-1}$.

We need the following lemma (see [8, p. 76]).

LEMMA 2. *Let $f(\lambda)$ be a spectral density function on $[-\pi, +\pi]$ and let $C$ be the covariance matrix corresponding to $f$. Suppose that $f'$ exists and satisfies the Lipschitz condition*

$$|f'(\lambda_1) - f'(\lambda_2)| < K|\lambda_1 - \lambda_2|^{\beta}, \tag{17}$$

*where $K > 0$, $0 < \beta < 1$. Then*

$$\lim_{t\to\infty} \det([C]_0^t) \Big/ \left\{\exp\left(\frac{1}{2\pi}\int_{-\pi}^{+\pi} \log f(\lambda)\,d\lambda\right)\right\}^{t+1} = L_f \tag{18}$$

*exists and $1 < L_f < \infty$.*

As spectral densities defined by (14) and (15) satisfy Lipschitz condition (17), we can apply Lemma 2. Using (18) and (13), we obtain

$$\lim_{t\to\infty} \det\left(\sum_{j=1}^{N} \alpha_j \left([C_j]_0^t\right)^{-1}\right) \Big/ \left\{\exp\left(\frac{1}{2\pi}\int_{-\pi}^{+\pi} \log g(\lambda)\,d\lambda\right)\right\}^{t+1} = \hat{L}_g, \tag{19}$$

where $\hat{L}_g$ is a finite constant. For the first term on the right hand side of (3), we have

$$\lim_{t\to\infty} \prod_{j=1}^{N} \left(\det([C_j]_0^t)\right)^{\alpha_j} \Big/ \prod_{j=1}^{N} \left\{\left(\exp\left(\frac{1}{2\pi}\int_{-\pi}^{+\pi} \log f_j(\lambda)\,d\lambda\right)\right)^{t+1}\right\}^{\alpha_j} = \prod_{j=1}^{N} L_{f_j}^{\alpha_j}. \tag{20}$$

Thus, we have proved the following theorem.

THEOREM 2. *Let $\{\xi_t\}$ be a zero mean stationary Gaussian AR(1) process, $\xi_{t+1} = a\xi_t + \varepsilon_{t+1}$, $t = 0, \pm 1, \pm 2, \dots$, where $\{\varepsilon_t\}$ is a sequence of independent standard Gaussian random variables. Let $H_j$ be the hypothesis that $a = a_j$, where $|a_j| < 1$, $j = 1, \dots, N$. Let $f_j(\lambda) = \left(1 - 2a_j \cos\lambda + a_j^2\right)^{-1}$ be the spectral density of the process and let $P_j^{(t)}$ denote the distribution of $\{\xi_0, \dots, \xi_t\}$ if $H_j$ is true, $j = 1, \dots, N$. Then*

$$\lim_{t\to\infty} \frac{\mathbb{H}_\alpha\left(P_1^{(t)}, \dots, P_N^{(t)}\right)}{\left\{\exp\left((-1/4\pi)\int_{-\pi}^{+\pi} \log\left(\prod_{j=1}^{N} f_j^{\alpha_j}(\lambda) \sum_{j=1}^{N} \alpha_j f_j^{-1}(\lambda)\right) d\lambda\right)\right\}^{t+1}} = L, \tag{21}$$

*where $L$ is a finite constant and $\alpha = (\alpha_1, \dots, \alpha_N)$, $\alpha_1 > 0, \dots, \alpha_N > 0$, $\sum_{j=1}^{N} \alpha_j = 1$.*

Relation (21) can be rewritten into the following form

$$\lim_{t\to\infty} \frac{1}{t+1} \log \mathrm{H}_\alpha\left(P_1^{(t)}, \dots, P_N^{(t)}\right) = \frac{-1}{4\pi}\int_{-\pi}^{+\pi} \log\left(\prod_{j=1}^{N} f_j^{\alpha_j}(\lambda) \sum_{j=1}^{N} \alpha_j f_j^{-1}(\lambda)\right) d\lambda. \tag{22}$$

## 3. MULTIDIMENSIONAL AR(1) PROCESSES

Let $\{y_t,\ t = 0, \pm 1, \pm 2, \dots\}$ be an $m$-dimensional stationary Gaussian AR(1) process. Suppose that under hypothesis $H_j$

$$y_{t+1} = A_j y_t + \varepsilon_{t+1}, \quad t = 0, \pm 1, \pm 2, \dots, \tag{23}$$

where $\{\varepsilon_t\}$ is a sequence of independent standard Gaussian random vectors. We assume that the eigenvalues of $A_j$ are inside the unit circle. Suppose that the mean of the process is 0. The covariance matrix and its inverse are the following block matrices

$$C_j = \begin{pmatrix} \ddots & \ddots & \ddots & & \\ \ddots & S_j & S_j A_j^\top & S_j A_j^{\top 2} & \\ \ddots & A_j S_j & S_j & S_j A_j^\top & \ddots \\ & A_j^2 S_j & A_j S_j & S_j & \ddots \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{24}$$

$$C_j^{-1} = \begin{pmatrix} \ddots & \ddots & \ddots & & \\ \ddots & I + A_j^\top A_j & -A_j^\top & 0 & \\ \ddots & -A_j & I + A_j^\top A_j & -A_j^\top & \ddots \\ & 0 & -A_j & I + A_j^\top A_j & \ddots \\ & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{25}$$

where $S_j = \sum_{\nu=0}^{\infty} A_j^\nu A_j^{\top \nu}$.

The covariance matrix of $\{y_0, \dots, y_t\}$ and the inverse covariance matrix have the form

$$[C_j]_0^t = \begin{pmatrix} S_j & S_j A_j^\top & \cdots & S_j A_j^{\top(t-1)} & S_j A_j^{\top t} \\ A_j S_j & S_j & \cdots & S_j A_j^{\top(t-2)} & S_j A_j^{\top(t-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_j^{t-1} S_j & A_j^{t-2} S_j & \cdots & S_j & S_j A_j^\top \\ A_j^t S_j & A_j^{t-1} S_j & \cdots & A_j S_j & S_j \end{pmatrix}, \tag{26}$$

$$\left([C_j]_0^t\right)^{-1} = \begin{pmatrix} S_j^{-1} + A_j^\top A_j & -A_j^\top & 0 & \cdots & 0 \\ -A_j & I + A_j^\top A_j & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & I + A_j^\top A_j & -A_j^\top \\ 0 & \cdots & 0 & -A_j & I \end{pmatrix}. \tag{27}$$

PROOF OF THEOREM 1. First we shall prove that eigenvalues of $\left([C_j]_0^t\right)^{-1}$ are bounded from zero. Let $\lambda$ denote an eigenvalue of $[C_j]_0^t$. By Gersgorin's theorem (see [9, Chapter XIV, Theorem 6])

$$\left\|\left(S_j - \lambda I\right)^{-1}\right\|^{-1} \le \sum_k \left(\|S_j A_j^{\top k}\| + \|A_j^k S_j\|\right) \le 2\|S_j\| \sum_k \|A_j^k\|. \tag{28}$$

As the spectral radius of $A_j$ is less than 1 the series $\sum_{k=0}^{\infty} A_j^k$ is absolute convergent. Therefore, $2\|S_j\| \sum_k \|A_j^k\| \le K$, where $K$ is a finite constant not depending on $t$. As $\left(S_j - \lambda I\right)^{-1}$ is symmetric its spectral norm is equal to its spectral radius. Therefore, (28) implies that there exists a constant $K_0$ (not depending on $t$) such that $|\lambda| \le K_0$ for an arbitrary eigenvalue of $[C_j]_0^t$. So eigenvalues of $\left([C_j]_0^t\right)^{-1}$ are bounded from zero. As $\left([C_j]_0^t\right)^{-1}$ are positive definite there exists an $L > 0$ such that $|\lambda| > L$ for each $\lambda$ being an eigenvalue of $\sum_{j=1}^{N} \alpha_j \left([C_j]_0^t\right)^{-1}$ (see [10, Theorem 3.6.3]).

Let us consider the following partition

$$\sum_{j=1}^{N} \alpha_j \left( [C_j]_0^t \right)^{-1} = \begin{pmatrix} U & V \\ V^\top & W \end{pmatrix}, \tag{29}$$

where $W$ is of size $m \times m$ while $U$ is of size $tm \times tm$. Adding $\sum_{j=1}^{N} \alpha_j A_j^\top A_j$ to the bottom right hand block of $\sum_{j=1}^{N} \alpha_j \left( [C_j]_0^{t-1} \right)^{-1}$ we obtain $U$. Hence, $|\lambda| > L$ for each eigenvalue of $U$, whence $|\lambda| < 1/L < \infty$ for each eigenvalue of $U^{-1}$, where the upper bound does not depend on $t$. Taking into account the special structure of matrices in (29) we can see that the eigenvalues of $W - V^\top U^{-1} V$ have an absolute bound. Therefore,

$$\lim_{t\to\infty} \frac{1}{t} \log\det(W - V^\top U^{-1} V) = 0. \tag{30}$$

Now, consider the following partition

$$U = \begin{pmatrix} U_0 & V_0 \\ V_0^\top & W_0 \end{pmatrix}, \tag{31}$$

where $U_0$ is of size $m \times m$ while $W_0$ is of size $(t-1)m \times (t-1)m$. We know that $|\lambda| > L$ for each eigenvalue of $U$. As $U$ is symmetric, we can apply [11, Theorem 9.C.1]. We obtain that $|\lambda| > L$ for each eigenvalue of $W_0$. Therefore,

$$\lim_{t\to\infty} \frac{1}{t} \log\det(U_0 - V_0 W_0^{-1} V_0^\top) = 0. \tag{32}$$

Now

$$\begin{aligned} \det\left( \sum_{j=1}^{N} \alpha_j \left( [C_j]_0^t \right)^{-1} \right) &= \det U \det(W - V^\top U^{-1} V) \\ &= \det W_0 \det(U_0 - V_0 W_0^{-1} V_0^\top) \det(W - V^\top U^{-1} V). \end{aligned} \tag{33}$$

By (30) and (32),

$$\lim_{t\to\infty} \frac{1}{t+1} \log\det\left( \sum_{j=1}^{N} \alpha_j \left( [C_j]_0^t \right)^{-1} \right) = \lim_{t\to\infty} \frac{1}{t+1} \log\det W_0, \tag{34}$$

if the limit on the right hand side exists. But

$$W_0 = \sum_{j=1}^{N} \alpha_j \left( [C_j^{-1}]_0^{t-2} \right), \tag{35}$$

that is $W_0$ contains sections of the inverse covariances $C_j^{-1}$. Equations (3), (5), (34), and (35) imply Theorem 1.

## REFERENCES


1. A.N. Shiryayev, *Probability* (in Russian), Nauka, Moscow, (1989).
2. M. Arató, Some remarks on the absolute continuity of measures (in Russian), *Mat. Kut. Közl.* **VI** (2), 123–126 (1961).
3. F. Liese and I. Vaida, *Convex Statistical Distances*, Teubner-Texte, Leibzig, (1987).
4. S. Sugimoto and T. Wada, Spectral expressions of information measures, *IEEE Transactions of Information Theory* **34** (4), 625–631 (1988).
5. M. Rosenblatt, Asymptotic distribution of eigenvalues of block Toeplitz matrices, *Bull. Amer. Math. Soc.* **66**, 320–321 (1960).

6. D. Kazakos and P. Papantoni-Kazakos, Spectral distance measures between Gaussian processes, *IEEE Transactions on Automatic Control* **25** (5), 950–958 (1980).
7. M. Arató, Linear stochastic systems with constant coefficients. A statistical approach (in Russian), Nauka, Moscow; English edition: Lecture Notes in Control and Information Sciences, Vol. 45, Springer-Verlag, Berlin, (1982).
8. U. Grenander and G. Szegö (1958), *Toeplitz Forms and Their Applications*, University of California Press, Berkeley and Los Angeles, (1958).
9. F.R. Gantmacher, *Theory of Matrices*, Chelsea, Bronx, New York, (1959).
10. P. Lancaster, *Theory of Matrices*, Academic Press, New York, (1969).
11. A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, (1979).