PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 29, No. 2, July 1971

ON DETERMINATION OF THE OPTIMAL FACTOR OF A NONNEGATIVE MATRIX-VALUED FUNCTION

HABIB SALEHI[1]

ABSTRACT. Let $F=[f_{ij}]$, $1<i, j\leqq q$, be a measurable, nonnegative definite $q\times q$ matrix-valued function defined on the unit circle C. It is known that when F and log det F are in $L_1(C)$, F admits a factorization of the form $F=\Phi\Phi^*$, where Φ is an optimal, full rank function in $L_2^{0+}(C)$. Under the additional assumption that $\{(\prod_{i=1}^{q} f_{ii})/\det F\}$ is in $L_1(C)$, an iterative procedure which yields an infinite series for Φ in terms of F is given. The optimal function Φ plays a significant role in the multivariate prediction theory of stochastic processes. The present work generalizes the results of several authors concerning the determination of the optimal factor Φ.

1. **Introduction.** Let $F=[f_{ij}]$, $1\leqq i, j\leqq q<\infty$, be a measurable, nonnegative definite $q\times q$ matrix-valued function defined on the unit circle C. An important problem in multivariate prediction theory is, given that $F\in L_1$ and log det $F\in L_1$, to find a measurable $q\times q$ matrix-valued function Φ defined on C such that $F=\Phi\Phi^*$ a.e. on C, $\Phi\in L_2^{0+}$, whose 0th Fourier coefficient C_0 has the property that

$$C_0>0 \quad \text{and} \quad \det C_0=\exp\frac{1}{2\pi}\int_0^{2\pi}\log\det F\,d\theta>0.$$

An interesting iterative procedure which yields an infinite series for Φ in terms of F was given by Wiener and Masani in [5] under the hypothesis that there exist constants c_1, c_2; $0\leqq c_1\leqq c_2<\infty$, such that:

$$c_1 I\leqq F\leqq c_2 I. \tag{1.1}$$

In [2] Masani was able to improve the results he and Wiener gave in [5] by assuming in lieu of condition (1.1) that:

(1.2)
(i) $F\in L_1$,
(ii) F^{-1} exists a.e. on C and $F^{-1}\in L_1$,
(iii) if λ and μ denote the smallest and largest eigenvalues of F, then $\mu/\lambda\in L_1$.

Received by the editors August 14, 1970.
AMS 1970 *subject classifications*. Primary 60G25; Secondary 47A50.
Key words and phrases. Factorization problem, optimal factor, prediction theory, iterative procedure, eigenvalue, determinant, Fourier coefficient, bounded operator.
[1] This research was supported by NSF GP-11626 and NONR N0014-67-A-0112-9015.

In his work [2] Masani repeatedly made use of the fact that $F^{-1} \in L_1$. However for many situations this condition is not necessarily satisfied as was pointed out by the author in [4].

The purpose here is to give an algorithm for finding $\mathbf{\Phi}$, similar to the one obtained in [5], [2] and [4], under a weaker and a more readily verifiable assumption on F. The main result of this paper is:

1.3 THEOREM. *Let $F=[f_{ij}]$, $1 \leqq i, j \leqq q$, be a measurable, nonnegative $q \times q$ matrix-valued function on C such that*

$$\text{(1.4)} \qquad \begin{aligned} &\text{(i)} \qquad F \in L_1 \quad \textit{and} \quad \log\det F \in L_1, \\ &\text{(ii)} \qquad \left\{ \left(\prod_{i=1}^{q} f_{ii} \right) \Big/ \det F \right\} \in L_1. \end{aligned}$$

Then

(a) *$f_{ii} = \phi_i \bar{\phi}_i$, $\phi_i \in L_2^{0+}$ (see §2), $|\phi_i| \neq 0$ a.e. on C, $1 \leqq i \leqq q$.*

(b) *Let $\hat{F}=[\hat{f}_{ij}]$, $\hat{f}_{ij}=f_{ij}/(\phi_i\bar{\phi}_j)$, $1 \leqq i, j \leqq q$, and λ, μ denote the smallest and the largest eigenvalues of $\hat{F}$. If $f=\frac{1}{2}\{\mu+\lambda\}$, then $f=\phi\bar{\phi}$, $\phi \in L_2^{0+}$, $|\phi| \neq 0$ a.e. on C.*

(c) *Let M be defined by $\hat{F}=f(I+M)$. Then $I+M=\chi\chi^*$, $\chi \in L_2^{0+}$, χ optimal of full rank in L_2^{0+} (see Definition 2.1).*

(d) *F is factorable, $F=\mathbf{\Phi}\mathbf{\Phi}^*$, $\mathbf{\Phi} \in L_2^{0+}$, $\mathbf{\Phi}$ optimal of full rank in L_2^{0+}. If $\mathbf{\Sigma}$ is defined by*

$$\mathbf{\Sigma} = \begin{bmatrix} \phi_1 & & & & & & 0 \\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & \phi_i & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ 0 & & & & & & \phi_q \end{bmatrix},$$

then the optimal factor $\mathbf{\Phi}$ is given by $\mathbf{\Phi}=\phi\mathbf{\Sigma}\chi$.

The proof of this theorem essentially consists of an initial factorization of the diagonal entries of F, followed by the application of the known results to the remaining factor. After determining the factors ϕ, $\mathbf{\Sigma}$ and χ in §2 and §3, a proof of Theorem 1.3 will be given near the end of §3.

1.5 REMARK. It is easy to see that (1.1) implies (1.4). We will show that the condition $(\mu/\lambda)^{q-1} \in L_1$ implies (1.4) (ii) (for $q=2$ this condition reduces to (1.2) (iii)), and that (1.2) (ii) is not necessary for the determination of $\mathbf{\Phi}$. We will also give a simple example where (1.2) (ii), (iii) fail to hold, yet (1.4) is satisfied.

2. **Preliminary results.** As in [2] bold face letters $\boldsymbol{A}$, $\boldsymbol{B}$, etc. will

denote $q \times q$ matrices and bold face letters $\boldsymbol{F}$, $\boldsymbol{\Phi}$, etc. will denote functions whose values are such matrices. tr, det, *, will be reserved for the trace, determinant and adjoint of matrices. $|A|_B$, will denote the Banach norms of A.

We shall be concerned with the class $\boldsymbol{L}_p$ $(1 \leqq p \leqq \infty)$ of $q \times q$ matrix-valued functions $\boldsymbol{F}$ on C whose entries are in L_p in the usual sense. $\boldsymbol{L}_2^+$, $\boldsymbol{L}_2^{0+}$, $\boldsymbol{L}_2^-$, $\boldsymbol{L}_2^{0-}$ will denote the subspace of functions in $\boldsymbol{L}_2$ whose nth Fourier coefficients vanish for $n \leqq 0$, $n < 0$, $n \geqq 0$, $n > 0$, respectively. If $\boldsymbol{F} \in \boldsymbol{L}_2$ and has Fourier coefficients A_k, $-\infty < k < \infty$, then $\boldsymbol{F}_+$, $\boldsymbol{F}_{0+}$, $\boldsymbol{F}_-$, $\boldsymbol{F}_{0-}$, will denote the functions in $\boldsymbol{L}_2^+$, $\boldsymbol{L}_2^{0+}$, $\boldsymbol{L}_2^-$, $\boldsymbol{L}_2^{0-}$, whose nth Fourier coefficients are A_n, for $n > 0$, $n \geqq 0$, $n < 0$, $n \leqq 0$, respectively (and zero for the remaining n). In $\boldsymbol{L}_2$ we introduce the Gramian, inner product and norm

$$(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \frac{1}{2\pi} \int_0^{2\pi} \boldsymbol{\Phi}(e^{i\theta}) \boldsymbol{\Psi}^*(e^{i\theta})\, d\theta, \qquad ((\boldsymbol{\Phi}, \boldsymbol{\Psi})) = \operatorname{tr}(\boldsymbol{\Phi}, \boldsymbol{\Psi}),$$

$$\|\boldsymbol{\Phi}\| = (\operatorname{tr}(\boldsymbol{\Phi}, \boldsymbol{\Phi}))^{1/2}.$$

2.1 Definition. (a) $\boldsymbol{\Phi} \in \boldsymbol{L}_2$ is said to be of full-rank if $|\det \boldsymbol{\Phi}(e^{i\theta})| > 0$ a.e. on C. (b) $\boldsymbol{\Phi}$ is called an optimal function in $\boldsymbol{L}_2^{0+}$ if $\boldsymbol{\Phi} \in \boldsymbol{L}_2^{0+}$, $\boldsymbol{\Phi}(0)^2 \geqq 0$ and $\boldsymbol{\Psi} \in \boldsymbol{L}_2^{0+}$, $\boldsymbol{\Psi}\boldsymbol{\Psi}^* = \boldsymbol{\Phi}\boldsymbol{\Phi}^*$ implies $\{\boldsymbol{\Psi}(0)\boldsymbol{\Psi}^*(0)\}^{1/2} \leqq \boldsymbol{\Phi}(0)$.

§4 of [2] contains several interesting results pertaining to the determination of the generating function of a multivariate stochastic process. Below we summarize those results which are mostly relevant to our work as will be needed in §3.

2.2 Theorem. *Let* (i) $\boldsymbol{M}$ *be a hermitian* $q \times q$ *matrix-valued function on* C, *and* $\boldsymbol{M} \in \boldsymbol{L}_\infty$,

(ii) $|\boldsymbol{M}(e^{i\theta})|_B < 1$ *a.e. on* C,

(iii) $(\boldsymbol{I} + \boldsymbol{M})^{-1} \in \boldsymbol{L}_1$,

(iv) *for each* $\boldsymbol{\Psi} \in \boldsymbol{L}_2$, $\mathcal{P}(\boldsymbol{\Psi}) = (\boldsymbol{\Psi}\boldsymbol{M})_+$.

Then (a) $\mathcal{P}$ *is a bounded linear operator on* $\boldsymbol{L}_2$ *into* $\boldsymbol{L}_2^+$.

(b) $\mathcal{P}(\boldsymbol{I}) = \boldsymbol{M}_+$, $\mathcal{P}^2(\boldsymbol{I}) = (\boldsymbol{M}_+\boldsymbol{M})_+$, *and so on.*

(c) *The series* $\boldsymbol{I} - \boldsymbol{M}_+ + (\boldsymbol{M}_+\boldsymbol{M})_+ - \{(\boldsymbol{M}_+\boldsymbol{M})_+\boldsymbol{M}\}_+ + \cdots$ *is mean convergent.*

(d) *If* χ *is the optimal factor of* $(\boldsymbol{I} + \boldsymbol{M})$[3] *with* $\boldsymbol{C}_0$ *as its* 0*th Fourier coefficient*, $\boldsymbol{\Psi}$ *the sum of the series in* (c), *then*

$$\boldsymbol{\Psi} = \boldsymbol{C}_0 \chi^{-1} \in \boldsymbol{L}_2^{0+}, \qquad \boldsymbol{C}_0^2 = \boldsymbol{\Psi}(\boldsymbol{I} + \boldsymbol{M})\boldsymbol{\Psi}^*. \tag{2.3}$$

[2] For a function $\boldsymbol{\Phi} \in \boldsymbol{L}_2^{0+}$, $\boldsymbol{\Phi}(0)$ denotes the value of the analytic extension of $\boldsymbol{\Phi}$ into the interior of C. Obviously $\boldsymbol{\Phi}(0)$ is the 0th Fourier coefficient of $\boldsymbol{F}$.

[3] We note that under our conditions on $\boldsymbol{M}$, $\boldsymbol{I} + \boldsymbol{M}$ is factorable.

Hence, as in [**2**], letting

$$\Psi(e^{i\theta}) = \sum_0^\infty A_k e^{ki\theta}, \qquad \Psi^{-1}(e^{i\theta}) = \sum_0^\infty B_k e^{ki\theta}, \tag{2.4}$$

we get, $A_0 = I$, and, for $m > 0$,

$$A_m = -\Gamma_m + \sum_n \Gamma_n \Gamma_{m-n} - \sum_n \sum_p \Gamma_p \Gamma_{n-p} \Gamma_{m-n} + \cdots, \tag{2.5}$$

where Γ_k is the kth Fourier coefficient of M, and all subscripts run from 1 to ∞. The kth Fourier coefficients of the optimal factor χ and χ^{-1} can then be had from the relations

$$C_k = B_k C_0, \qquad D_k = C_0^{-1} A_k. \tag{2.6}$$

3. **Determination of the optimal factor.** Let $F = [f_{ij}]$, $1 \leqq i, j \leqq q$, be a measurable nonnegative $q \times q$ matrix-valued function defined on C satisfying the condition (1.4).

It is well known that condition (1.4) (i) implies that F is factorable and that the optimal factor is of full-rank. Condition (1.4) (i) also implies the following lemma.

3.1 LEMMA. *Let $F = [f_{ij}]$, $1 \leqq i, j \leqq q$, satisfy condition* (1.4) (i). *Then each $f_{ii} \in L_1$ and* $\log f_{ii} \in L_1$, $1 \leqq i \leqq q$.

From Lemma 3.1 and the usual factorization problem for nonnegative scalar-valued functions follows that there exist optimal factors of rank 1 such that

$$f_{ii} = \phi_i \bar{\phi}_i, \qquad \phi_i \in L_2^{0+}, \qquad |\phi_i| > 0 \text{ a.e. on } C, \quad 1 \leqq i \leqq q. \tag{3.2}$$

The Fourier coefficients of ϕ_i, as usual, can be obtained from the Fourier coefficients of $\log f_{ii}$, so that we may regard ϕ_i, $1 \leqq i \leqq q$, as known. Now let the diagonal matrix Σ be defined by

$$\Sigma = \begin{bmatrix} \phi_1 & & & & 0 \\ & \ddots & & & \\ & & \phi_i & & \\ & & & \ddots & \\ 0 & & & & \phi_q \end{bmatrix}. \tag{3.3}$$

With this setting we have

$$F = \Sigma \hat{F} \Sigma^*, \qquad \hat{F} = [\hat{f}_{ij}], \qquad \hat{f}_{ij} = f_{ij}/(\phi_i \bar{\phi}_j). \tag{3.4}$$

In (3.4) the first and the third matrices on the right are in $\boldsymbol{L}_2^{0+}$ and $\boldsymbol{L}_2^{0-}$, and the one in the middle, $\hat{\boldsymbol{F}}$, is well defined and is in $\boldsymbol{L}_\infty$. Moreover $\hat{\boldsymbol{F}}$ is nonnegative and log det $\hat{\boldsymbol{F}} \in \boldsymbol{L}_1$. In brief

$$(3.5) \qquad \hat{\boldsymbol{F}} \in L_\infty, \qquad \log \det \hat{\boldsymbol{F}} \in \boldsymbol{L}_1.$$

By (3.5) $\hat{\boldsymbol{F}}$ is factorable:

$$(3.6) \quad \hat{\boldsymbol{F}} = \boldsymbol{\Omega}\boldsymbol{\Omega}^*, \qquad \boldsymbol{\Omega} \in \boldsymbol{L}_\infty^{0+}, \quad \boldsymbol{\Omega}(0) > 0, \quad \boldsymbol{\Omega} \text{ optimal in } \boldsymbol{L}_2^{0+}.$$

From (3.4) and (3.6) we obtain

$$(3.7) \qquad \boldsymbol{F} = (\boldsymbol{\Sigma\Omega})(\boldsymbol{\Sigma\Omega})^*.$$

Now let $\boldsymbol{\Phi} = \boldsymbol{\Sigma\Omega}$. Since $\boldsymbol{\Sigma} \in \boldsymbol{L}_2^{0+}$ and $\boldsymbol{\Omega} \in \boldsymbol{L}_\infty^{0+}$, we conclude that

$$(3.8) \qquad \boldsymbol{\Phi} = \boldsymbol{\Sigma\Omega} \in \boldsymbol{L}_2^{0+}.$$

Moreover since $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ are full-rank, optimal factors by Lemma 2.3 [**4**] it follows that $\boldsymbol{\Phi}$ is the full-rank, optimal factor of $\boldsymbol{F}$.

In view of these results, for the determination of $\boldsymbol{\Phi}$, it suffices to determine the Fourier coefficients of the optimal factor $\boldsymbol{\Omega}$ of $\hat{\boldsymbol{F}}$.

3.9 THEOREM. *Let* $\boldsymbol{F} = [f_{ij}]$, $1 \leq i, j \leq q$, *be a measurable nonnegative* $q \times q$ *matrix-valued function satisfying condition* (5.1), *and let* $\hat{\boldsymbol{F}}$ *be defined as in* (3.4). *Then*

(a) $\hat{\boldsymbol{F}} \in \boldsymbol{L}_\infty \subseteq \boldsymbol{L}_1$.

(b) $\hat{\boldsymbol{F}}^{-1} \in \boldsymbol{L}_1$.

(c) *If* λ, μ *denote the smallest and largest eigenvalues of* $\hat{\boldsymbol{F}}$, *then* $\mu/\lambda \in \boldsymbol{L}_1$.

PROOF. (a) Since $|f_{ij}/\phi_i\bar{\phi}_j| \leq 1$, $\hat{f}_{ij} = f_{ij}/\phi_i\bar{\phi}_j$, the (i, j)th entry of $\hat{\boldsymbol{F}}$ is in $\boldsymbol{L}_\infty$ and hence $\hat{\boldsymbol{F}} \in \boldsymbol{L}_\infty$.

(b) Since the entries of $\hat{\boldsymbol{F}}$ are bounded by 1 it follows that the entries of $\hat{\boldsymbol{F}}^{-1}$ are bounded by $\{(q-1)!\}^2/\det \hat{\boldsymbol{F}}$.[4] But

$$1/\det \hat{\boldsymbol{F}} = \left(\prod_{i=1}^{q} f_{ii}\right) \Big/ \det \boldsymbol{F} \in \boldsymbol{L}_1.$$

Therefore $\hat{\boldsymbol{F}}^{-1} \in \boldsymbol{L}_1$.

(c) follows from (a) and (b) (see [**2**, p. 158]).

Now, as in [**2**], we let

$$(3.10) \qquad f = \frac{1}{2}\{\mu + \lambda\}, \qquad \boldsymbol{M} = \frac{1}{f}\hat{\boldsymbol{F}} - \boldsymbol{I}.$$

[4] One can show that this bound can be replaced by $1/\det \hat{\boldsymbol{F}}$.

Then as shown in [2], we have

$$(3.11)\qquad \begin{aligned} &\text{(i)} && \hat{F} = f(I+M),\\ &\text{(ii)} && |M|_B < 1 \text{ a.e. on } C,\\ &\text{(iii)} && I+M,\ (I+M)^{-1} \in L_1,\\ &\text{(iv)} && f,\ 1/f \in L_1. \end{aligned}$$

By (3.11) (iv), since f is scalar valued, its optimal factor ϕ can be found and we have

$$(3.12)\qquad f = \phi\bar{\phi}, \qquad \phi \in L_2^{0+}, \quad |\phi| \neq 0 \text{ a.e. on } C.$$

By (3.11) (ii)–(iii) and Theorem 2.2 the optimal factor χ of $I+M$ can be found by an iterative procedure (cf. (2.3)–(2.6)). We write

$$(3.13)\qquad I+M = \chi\chi^*, \qquad \chi \in L_2^{0+}, \quad \chi(0) > 0, \quad \chi \text{ optimal}.$$

Summing up we have proved our main result Theorem 1.3.

The following lemma and corollary shed some light on the relation between the hypothesis (1.2) (iii) made in [2] and our condition (1.4).

3.14 LEMMA. *Let* $F=[f_{ij}]$, $1 \leqq i, j \leqq q$, *be a measurable, nonnegative* $q \times q$ *matrix-valued function on* C. *Let* λ, μ *be the smallest and largest eigenvalues of* F. *If* $(\mu/\lambda)^{q-1} \in L_1$, *then* $(\prod_{i=1}^{q} f_{ii})/\det F \in L_1$.

3.15 COROLLARY. *Let* $q=2$. *Then hypothesis* (1.2) (iii) *alone implies our condition* (1.4) (ii).

Next we give an example for which the hypothesis (1.4) (ii) is satisfied, but condition (1.2) (ii)–(iii) is not satisfied.

3.16 EXAMPLE. Let $F = \left[\begin{smallmatrix} f & 0 \\ 0 & g \end{smallmatrix}\right]$, where

(i) $0 \leqq f, g \in L_1$; $\log f$, $\log g \in L_1$,

(ii) $1/f$, $1/g \notin L_1$,

(iii) $(g/f) \notin L_1$.[5]

Obviously (1.2) (ii)–(iii) are not satisfied, but (1.4) (ii) is satisfied, because

$$\left(\prod_{i=1}^{2} f_{ii}\right) \Big/ \det F = fg/fg = 1 \in L_1.$$

We note that our algorithm remains valid with $\hat{F} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$.

REFERENCES

1. H. Helson, *Lectures on invariant subspaces*, Academic Press, New York, 1964. MR 30 #1409.

[5] Take for example $f = \exp\{-1/|\theta-1|^\alpha\}$, $g = \exp\{-1/|\theta+1|^\alpha\}$, $0<\alpha<1$.

2. P. Masani, *The prediction theory of multivariate stochastic processes.* III. *Unbounded spectral densities*, Acta Math. **104** (1960), 141–162. MR **22** #12679.

3. ———, *Recent trends in multivariate prediction theory*, Proc. Sympos. Multivariate Analysis (Dayton, Ohio, 1965), Academic Press, New York, 1966, pp. 351–382. MR **35** #5079.

4. H. Salehi, *A factorization algorithm for $q \times q$ matrix-valued functions on the real line R*, Trans. Amer. Math. Soc. **124** (1966), 468–479. MR **33** #8039.

5. N. Wiener and P. Masani, *The prediction theory of multivariate stochastic processes.* II. *The linear predictor*, Acta Math. **99** (1958), 93–137. MR **20** #4325.

Michigan State University, East Lansing, Michigan 48823