SP XIV, 324-331, SLN 784, 1980

WIENER-HOPF FACTORIZATION FOR MATRICES

by

M.T. Barlow, L.C.G. Rogers, and David Williams

1. The main results. Let E be a finite set. Let $\mathcal{Q}(E)$ denote the set of (real) $E \times E$ matrices Q such that, for $i, j \in E$,

$$Q(i,j) \geq 0 \quad (i \neq j), \quad \sum_{k \in E} Q(i,k) \leq 0 .$$

Let Q now denote some fixed element of $\mathcal{Q}(E)$. Let v be a function from E to $\mathbb{R} \setminus \{0\}$, and let V be the diagonal $E \times E$ matrix $\operatorname{diag}\{v(i) : i \in E\}$. Let $E^+ = \{i \in E : v(i) > 0\}$, and $E^- = \{i \in E : v(i) < 0\}$. Let I denote the identity $E \times E$ matrix, I^+ the identity $E^+ \times E^+$ matrix, and I^- the identity $E^- \times E^-$ matrix.

Let c be a strictly positive real number.

THEOREM I. There exists a unique pair (Π_c^+, Π_c^-), where Π_c^+ is an $E^- \times E^+$ matrix and Π_c^- is an $E^+ \times E^-$ matrix, such that, if

$$S = \begin{pmatrix} I^+ & \Pi_c^- \\ \Pi_c^+ & I^- \end{pmatrix}, \tag{1}$$

then S is invertible and

$$S^{-1}[V^{-1}(Q - cI)]S = \begin{pmatrix} \tilde{Q}_c^+ & 0 \\ 0 & -\tilde{Q}_c^- \end{pmatrix}, \tag{2}$$

where $\tilde{Q}_c^+ \in \mathcal{Q}(E^+)$ and $\tilde{Q}_c^- \in \mathcal{Q}(E^-)$. Moreover, Π_c^+ and Π_c^- are strictly substochastic: thus, for $i \in E^-$, $j \in E^+$,

$$\Pi_c^+(i,j) \geq 0, \quad \sum_{k \in E^+} \Pi_c^+(i,k) < 1 .$$

Theorem I will be said to yield the 'Wiener-Hopf factorization' of the matrix $V^{-1}(Q - cI)$.

Now let X be a Markov chain on $E \cup \{\partial\}$ (∂ is the cemetery state) with Q-matrix Q. Thus the transition matrix function of X is $P(t) = \exp(tQ)$. For $t \geq 0$, define :

$$\phi(t) = \int_0^t v(X_s)ds, \quad \tau^+(t) = \inf\{s : \phi(s) > t\}.$$

As usual, we shall (for example) write τ_t^+ for $\tau^+(t)$ when more convenient. Note that $X(\tau_t^+) \in E^+ \cup \{\partial\}$.

THEOREM II. <u>For</u> $i \in E^-$ <u>and</u> $j \in E^+$,

$$E^i[\exp(-c\tau_0^+);\ X(\tau_0^+) = j] = \Pi_c^+(i,j). \tag{3}$$

<u>For</u> $i \in E^+$, $j \in E^+$, <u>and</u> $t \geq 0$,

$$E^i[\exp(-c\tau_t^+);\ X(\tau_t^+) = j] = [\exp(t\tilde{Q}_c^+)](i,j). \tag{4}$$

The corresponding 'minus' results follow, on replacing ϕ by $(-\phi)$.

The problem of finding the joint distribution of τ_t^+ and $X(\tau_t^+)$ is of course solved by Theorems I and II. The way in which Π_c^+ and Π_c^- may be calculated will be clear from the proofs.

<u>Comment</u>. The reader may feel that the martingale techniques used in this paper are more sophisticated than those required for this Markov chain problem. The following two statements are therefore apposite. First, we do not know how to prove the purely algebraic Theorem I without appealing to probability theory (and ultimately to martingale theory). Second, the martingale technique generalises to other (more interesting) cases, though the problem of obtaining explicit answers proves to be very difficult. Work on the 'continuous' state-space case will be published later.

2. <u>Basic martingales</u>. It is important to regard the strictly positive number c as <u>fixed</u> throughout the remainder of the paper, except in section 7.

Let f be a function on $(E \cup \{\partial\}) \times \mathbb{R} \times [0,\infty)$ such that $f(\partial,\cdot,\cdot) = 0$. A natural extension of Dynkin's formula shows that (for <u>every</u> initial distribution)

$$f(X_t, \phi_t, t) - \int_0^t \mathcal{A} f(X_s, \phi_s, s)\,ds$$

is a local martingale, where

$$\mathcal{A} f(x,\phi,t) = Qf + V\frac{\partial f}{\partial \phi} + \frac{\partial f}{\partial t} .$$

Here, of course,

$$Qf(x,\phi,t) = \sum_{y \in E} Q(x,y) f(y,\phi,t) .$$

In particular, if g is any vector on E, and

$$(5) \qquad f(x,\phi,t) = \{\exp[-ctI - \phi V^{-1}(Q - cI)]g\}(x) \quad \text{on } E,$$

then $f(X_t, \phi_t, t)$ is a local martingale (in fact, a <u>martingale</u>, because it is bounded on every finite interval).

3. <u>Definition of $\mathcal{N}$.</u> Before recalling part of the theory of the Jordan form, we recall the proof of the well-known fact that $V^{-1}(Q - cI)$ cannot have an eigenvalue on the imaginary axis. For suppose that μ lies on the imaginary axis and that

$$(6) \qquad (Q - cI)g = \mu V g$$

for some non-zero vector g. Choose i in E with $|g(i)| \geq |g(j)|$ for all j in E. The i-th coordinate of (6) reads :

$$[Q(i,i) - c - \mu v(i)]g(i) = -\sum_{j \neq i} Q(i,j) g(j) .$$

But the left-hand side has modulus at least equal to $(|Q(i,i)| + c)|g(i)|$, while the right-hand side has modulus at most equal to $|Q(i,i)||g(i)|$. The contradiction establishes the 'well-known fact'.

One of the main steps in Jordan-form theory shows that <u>the space of complex vectors on</u> E <u>has a basis</u> $\mathcal{G}$ <u>such that every</u> g <u>in</u> $\mathcal{G}$ <u>solves an equation</u>

(7) $$[V^{-1}(Q - cI) - \mu I]^k g = 0 ,$$

where μ is an eigenvalue of $V^{-1}(Q - cI)$ and k is a positive integer. Fix $\mathcal{G}$, and let $\mathcal{N}$ [respectively, $\mathcal{P}$] denote the set of those vectors g in $\mathcal{G}$ for which the associated μ-value has (strictly) negative [respectively, positive] real part.

4. The structure of $\mathcal{N}$. Let $g \in \mathcal{N}$, so that g satisfies (7) for some k and some μ with negative real part. Then the function f at (5) may be written

$$f(\cdot,\phi,t) = \exp(-ct - \mu\phi)\exp\{-\phi[V^{-1}(Q - cI) - \mu I]\}g ,$$

and the second exponential may be expanded in a power series in which all terms after the (k-1)-th annihilate g. Hence, since μ has negative real part,

(8) for $g \in \mathcal{N}$ and f as at (5), $f(X_s,\phi_s,s)$ is bounded on $[0,\tau_t^+]$ for every t.

In particular, on applying the optional stopping theorem at time τ_0^+, we find that (for all $i \in E$)

$$\underset{\sim}{E}^i[\exp(-c\tau_0^+)g \circ X(\tau_0^+)] = g(i) .$$

Now define Π_c^+ via the probabilistic formula (3). We have just shown that

(9) if $g \in \mathcal{N}$, then $g = \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} g^+$ where g^+ denotes the restriction of g to E^+.

Hence $\mathcal{N}$ has at most $|E^+|$ elements, and, by a similar argument, $\mathcal{P}$ has at most $|E^-|$ elements. The only explanation is that

(10) $\mathcal{N}$ has precisely $|E^+|$ elements and the elements g^+ where $g \in \mathcal{N}$ form a basis for the space of vectors on E^+.

5. Proof of the uniqueness of Wiener-Hopf factorization. Suppose that for some $E^- \times E^+$ matrix K_c^+, some $E^+ \times E^-$ matrix K_c^-, some $\hat{Q}_c^+$ in $\mathcal{Q}(E^+)$,

and some $\hat{Q}_c^-$ in $\mathcal{Q}(E^-)$, we have that $\begin{pmatrix} I^+ & K_c^- \\ K_c^+ & I^- \end{pmatrix}$ is invertible, and

$$\begin{pmatrix} I^+ & K_c^- \\ K_c^+ & I^- \end{pmatrix}^{-1} V^{-1}(Q-cI) \begin{pmatrix} I^+ & K_c^- \\ K_c^+ & I^- \end{pmatrix} = \begin{pmatrix} \hat{Q}_c^+ & 0 \\ 0 & -\hat{Q}_c^- \end{pmatrix}.$$

Then the eigenvalues of $V^{-1}(Q-cI)$ with negative real part must coincide with the eigenvalues of $\hat{Q}_c^+$. Moreover, if

$$(\hat{Q}_c^+ - \mu I^+)^k u^+ = 0 \tag{11}$$

for some positive integer k, some μ(with negative real part) and some vector u^+ on E^+, then

$$\{V^{-1}(Q-cI) - \mu I\}^k \begin{pmatrix} I^+ \\ K_c^+ \end{pmatrix} u^+ = 0,$$

so that, from the argument leading to (9),

$$\begin{pmatrix} I^+ \\ K_c^+ \end{pmatrix} u^+ = \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} u^+ . \tag{12}$$

By the theory of the Jordan canonical form for $\hat{Q}_c^+$, equation (11) holds for a set of vectors u^+ spanning the vectors on E^+, and so therefore does equation (12). Hence

$$K_c^+ = \Pi_c^+$$

and the required uniqueness follows from this fact and its 'minus' analogue.

6. <u>Existence of the Wiener-Hopf factorization</u>. For the moment, regard $t \geq 0$ as fixed. Let $g \in \mathcal{N}$, define f as at (5), and recall that $f(X_s, \phi_s, s)$ is a martingale. Using (8) as justification, apply the optional stopping theorem at time τ_t^+ to obtain

$$\underset{\sim}{E}^{\cdot}[\exp(-c\tau_t^+) h \circ X(\tau_t^+)] = g = \exp[tV^{-1}(Q-cI)]h,$$

where $h = \exp[-tV^{-1}(Q-cI)]g$. Note that h will automatically satisfy the same version of (7) as does g, so that h, like g, has property (9). Hence, since $X(\tau_t^+) \in E^+$, we have

$$(13) \quad \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} \underset{\sim}{E}^{\cdot}[\exp(-c\tau_t^+) h^+ \circ X(\tau_t^+)] = \exp[tV^{-1}(Q - cI)] \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} h^+ .$$

We have obtained (13) for a class of vectors h^+ which clearly spans the space of all vectors on E^+, so that <u>(13) holds for all vectors</u> h^+ <u>on</u> E^+.

It is almost immediate from the strong Markov property of X that

$$\tilde{P}_c^+(t;i,j) = \underset{\sim}{E}^i[\exp(-c\tau_t^+);\ X(\tau_t^+) = j] \qquad (i,j \in E^+)$$

defines a subMarkovian transition function on E^+, so that

$$\tilde{P}_c^+(t) = \exp(t\tilde{Q}_c^+)$$

for some $\tilde{Q}_c^+$ in $\mathcal{Q}(E^+)$. (One can alternatively deduce this from (13).) On differentiating (13) with respect to t and setting $t = 0$, we obtain

$$(14) \quad \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} \tilde{Q}_c^+ = V^{-1}(Q - cI) \begin{pmatrix} I^+ \\ \Pi_c^+ \end{pmatrix} .$$

Theorems I and II now follow from (14) and its 'minus' analogue.

Department of Pure Mathematics
University of Liverpool
P.O. Box 147
Liverpool, L69 3BX

(Barlow)

Department of Pure Mathematics
University College of Swansea
Singleton Park
Swansea, SA2 8PP

(Rogers, Williams)