Stochastic Processes and their Applications 46 (1993) 283-299
North-Holland

On correlation calculus for multivariate martingales

Kacha Dzhaparidze
Centre for Mathematics and Computer Science, Amsterdam, Netherlands

Peter Spreij
Department of Econometrics, Free University, Amsterdam, Netherlands

Received 2 March 1992
Revised 30 April 1992

In this paper the correlation between two multivariate martingales is studied. This correlation can be expressed in a nondecreasing process, that remains zero in the case of linear dependence. A key result is an integral representation for this process.

martingale * quadratic variation * correlation * Moore-Penrose inverse

1. Introduction

Let (Ω, F, P) be a complete filtered probability space. Let $M:\Omega\times[0,\infty)\to\mathbb{R}^n$ and $m:\Omega\times[0,\infty)\to\mathbb{R}^k$ be locally square integrable martingales. We assume that both M_0 and m_0 are zero.

Denote by $\langle m, M\rangle$ the predictable covariation process of m and M. So $\langle m, M\rangle$: $\Omega\times[0,\infty)\to\mathbb{R}^{k\times n}$ and if m^i and M^j are the ith and jth components of m and M respectively, then the ij entry $\langle m, M\rangle^{ij}$ of $\langle m, M\rangle$ equals the real valued process $\langle m^i, M^j\rangle$. $\langle m\rangle=\langle m, m\rangle$ and $\langle M\rangle=\langle M, M\rangle$ are defined likewise.

Assume now that for some $t>0$ the matrices $\langle m\rangle_t$ and $\langle M\rangle_t$ are invertible. Then, parallel to what one can do when dealing with multivariate random variables, it is natural to express the correlation between m and M over the interval $[0, t]$ by

$$\rho(m, M)_t=\langle m\rangle_t^{-1/2}\langle m, M\rangle_t\langle M\rangle_t^{-1/2}.$$

Let $c(m, M)_t=\langle m\rangle_t-\langle m, M\rangle_t\langle M\rangle_t^{-1}\langle M, m\rangle_t$. Then we have the identity

$$I-\rho(m, M)_t\rho(M, m)_t=\langle m\rangle_t^{-1/2}c(m, M)_t\langle m\rangle_t^{-1/2}.$$

It follows that $c(m, M)_t$ carries the same amount of information about the correlation between m and M as $\rho(m, M)_t$. It turns out that it is more convenient to study

Correspondence to: Dr. Peter Spreij, Department of Econometrics, Free University, De Boelelaan 1105, 1081 HV Amsterdam, Netherlands.

0304-4149/93/$06.00

$c(m, M)_t$ than $\rho(m, M)_t$. The process $c(m, M)$ is of interest in its own right, because it appears at several places in probability and statistics. For example, this process — or rather a slightly different one — appears in Dzhaparidze and Spreij (1989), where we studied a strong law of large numbers for matingales. The results of the present paper offer an alternative approach to such a study.

In the situation where m and M are Gaussian martingales with deterministic brackets $c(m, M)_t$ has an interpretation as an (L_2-) projection error. Indeed $\langle m, M\rangle_t\langle M\rangle_t^{-1}M_t$ can be considered as an L_2-projection of m_t on M_t and it is also the conditional expectation of m_t given M_t and the conditional covariance matrix of m_t given M_t is precisely $c(m, M)_t$. See Lipster and Shiryaev (1979, Theorem 13.1), which also describes the case where the inverse doesn't exist. In the general framework that we consider in the present paper (we don't assume that the brackets are deterministic) it is not clear whether a similar interpretation holds.

In a statistical context $c(m, M)$ can be interpreted as a measure of deficiency when comparing an arbitrary estimator with an optimal one. Consider for instance the following simple regression example. Let $y_i = x_i^{\mathrm{T}}\beta + \varepsilon_i$ for $i = 1, \ldots, n$. Assume that the ε's are independent standard normal random variables, $\beta \in \mathbb{R}^k$ and that the regressors are deterministic. Write $Y^{\mathrm{T}} = [y_1, \ldots, y_n]$, $X = [x_1, \ldots, x_n]$. Let Q be some positive definite matrix and denote by $\hat{\beta}$ the minimizer of the quadratic form $(Y - X^{\mathrm{T}}\beta)^{\mathrm{T}}Q(Y - X^{\mathrm{T}}\beta)$. Then $\hat{\beta} = (XQX^{\mathrm{T}})^{-1}XQY$, assuming that X has rank k. For the special case that Q is equal to the identity matrix we have as the minimizer $\beta^* = (XX^{\mathrm{T}})^{-1}XY$. The Gauss-Markov theorem states that $\mathrm{Cov}(\hat{\beta}) \geq \mathrm{Cov}(\beta^*)$. Hence

$$0 \leq XQX^{\mathrm{T}}(\mathrm{Cov}(\hat{\beta}) - \mathrm{Cov}(\beta^*))XQX^{\mathrm{T}} = XQ^2X^{\mathrm{T}} - XQX^{\mathrm{T}}(XX^{\mathrm{T}})^{-1}XQX^{\mathrm{T}}.$$

Define the martingales m and M by

$$m_t = \sum_{j=1}^{t} x_j\varepsilon_j \quad \text{and} \quad M_t = \sum_{j=1}^{t}\sum_{i=1}^{k} x_iQ_{ij}\varepsilon_j.$$

Then the right hand side of the last inequality is just $c(m, M)_n$. For more general estimation problems we refer to Dzhaparidze and Spreij (1990) for details.

In the present paper we drop the restrictions that $\langle m\rangle_t$ and $\langle M\rangle_t$ are invertible. So we have to replace $\langle M\rangle_t^{-1}$ in the definition of $c(m, M)_t$ by a suitable generalized inverse. The Moore-Penrose inverse turns out to be a good choice. Working with a generalized inverse however complicates the analysis of $c(m, M)$ considerably.

The rest of the paper is organized as follows. In Section 2 we describe some properties of $\langle M\rangle$, its Moore-Penrose inverse process $\langle M\rangle^+$ and invariance properties of M under a to $\langle M\rangle_t$ related orthogonal projection. Section 3 contains an important integral representation of $c(m, M)$. In Section 4 linear dependence between m and M is defined by $c(m, M) = 0$ and characterized by the property that there is a constant (*random*!) matrix C such that $m = CM$. See Example 1 in the next section for a motivating example.

The familiar case of linear dependence where m and M are replaced with random variables is easily recognized.

2. Some technical results

In this section we describe some properties of the process $\langle M\rangle$. $\langle M\rangle$ takes its values in the space of positive semidefinite $n\times n$ matrices $\mathcal{P}_n$, and if $t>s$, then $\langle M\rangle_t-\langle M\rangle_s\in\mathcal{P}_n$.

For fixed t, $\omega\langle M\rangle_t=\langle M\rangle_t(\omega)$ may have non trivial kernel. This is typically the case if $M_t=\sum_{i=1}^t x_i\varepsilon_i$, where ε_i is a real valued martingale difference sequence and x and $\mathbb{R}^n$-valued predictable process. Then $\langle M\rangle_t$ for $t<n$ is always a singlular matrix. For $t>s$ we always have $\mathrm{Im}\langle M\rangle_t\supset\mathrm{Im}\langle M\rangle_s$, where $\mathrm{Im}\langle M\rangle_t$ is the image space of $\langle M\rangle_t$, a linear subspace of $\mathbb{R}^n$.

Define $r:\Omega\times[0,\infty)\to\{0,\dots,n\}$ by

$$r_t=\dim\mathrm{Im}\langle M\rangle_t=\mathrm{rank}\langle M\rangle_t.$$

Then r is a predictable process (see Proposition 2.1). Although $\langle M\rangle$ is a right continuous process, r may fail to be right (or left) continuous. See Example 2 below. Define the stopping times T_k $(k=0,\dots,n+1)$ by $T_0=0$ and $T_{k+1}=\inf\{t>T_k\colon r_t>r_{T_k}\}$ $(\inf\emptyset=\infty)$. Then each $T_k:\Omega\to[0,\infty]$, and $T_{n+1}=\infty$. The T_k are in general not predictable (see Example 1), which is one of the sources of the technical complexity in the analysis hereafter. For $(\omega,t)\in\rrbracket T_k,T_{k+1}\llbracket$ we have that $\mathrm{Im}\langle M\rangle_t$ does not depend on t, and hence r is constant on this stochastic interval. So we can find a (random) matrix $F(k)$ of size $n\times r_t$ such that the columns of $F(k)$ span $\mathrm{Im}\langle M\rangle_t$ and $F(k)^{\mathrm{T}}F(k)=I_{r_t}$, the $r_t\times r_t$ identity matrix. Similarly we can find matrices $G(k)$ of size $n\times r_{T_k}1_{\{T_k<\infty\}}$ such that the columns of $G(k)$ span $\mathrm{Im}\langle M\rangle_{T_k}1_{\{T_k<\infty\}}$ and such that $G(k)^{\mathrm{T}}G(k)=I_{r_{T_k}1_{\{T_k<\infty\}}}$. Moreover, since $\mathrm{Im}\langle M\rangle_t\supset\mathrm{Im}\langle M\rangle_s$ for $t>s$, we can always assume that $F(k)$ is of the form $[G(k),U_1(k)]$, where $U_1(k)$ is a $n\times(r_t-r_{T_k}1_{T_k<\infty})$ matrix for $(\omega,t)\in\rrbracket T_k,T_{k+1}\llbracket$, and likewise $G(k)$ is of the form $[F(k-1),U_2(k)]$. Then for $(\omega,t)\in\rrbracket T_k,T_{k+1}\llbracket$ there exists a $r_t\times r_t$ matrix $V_t(k)$ such that

$$\langle M\rangle_t=F(k)V_t(k)F(k)^{\mathrm{T}}$$

and there exists an $r_{T_k}1_{\{T_k<\infty\}}\times r_{T_k}1_{\{T_k<\infty\}}$ matrix $W(k)$ such that

$$\langle M\rangle_{T_k}1_{\{T_k<\infty\}}=G(k)W(k)G(k)^{\mathrm{T}}.$$

Notice that the $V_t(k)$ and the $W(k)$ are in general not diagonal. Hence

$$\langle M\rangle_.=\sum_{k=0}^n 1_{\rrbracket T_k,T_{k+1}\llbracket}F(k)V_.(k)F(k)^{\mathrm{T}}+\sum_{k=0}^n 1_{\llbracket T_k\rrbracket}G(k)W(k)G(k)^{\mathrm{T}}.\tag{2.1}$$

On the sets where the $V_t(k)$ and $W(k)$ are defined, these matrices are invertible. Therefore we can define the generalized inverse process $\langle M\rangle^+$ by

$$\langle M\rangle_.^+=\sum_{k=0}^n 1_{\rrbracket T_k,T_{k+1}\llbracket}F(k)V_.(k)^{-1}F(k)^{\mathrm{T}}+\sum_{k=0}^n 1_{\llbracket T_k\rrbracket}G(k)W(k)^{-1}G(k)^{\mathrm{T}}.\tag{2.2}$$

Proposition 2.1. *$\langle M\rangle_t^+$ defined by equation (2.2) is for each t the Moore–Penrose inverse of $\langle M\rangle_t$ and r and $\langle M\rangle^+$ are predictable processes.*

Proof. First we show that the map rank: $\mathbb{R}^{m\times n} \to \{0, \ldots, m \wedge n\}$ is upper semi-continuous, that is the sets $G_p = \{A \in \mathbb{R}^{m\times n}: \text{rank } A \geqslant p\}$ are open in the ordinary topology on $\mathbb{R}^{m\times n}$. Let $A \in G_p$, and rank $A = q \geqslant p$. Then A contains a submatrix $A_q \in \mathbb{R}^{q\times q}$ with rank $A_q = q$. Let $\{\varepsilon_k\} \subset \mathbb{R}^{m\times n}$ be a sequence of matrices converging to zero. Let ε_{qk} be the submatrix of ε_k that is obtained in the same was as A_q, that is by deleting the same rows and columns. Then $\lim_{k\to\infty} \det(A_q + \varepsilon_{qk}) \neq 0$. (by continuity of the determinant). Hence $\text{rank}(A_q + \varepsilon_{qk}) = q$ for all k large enough and consequently $\text{rank}(A + \varepsilon_k) \geqslant q$ for the same k. This shows that G_p is open. As a consequence rank is a (Borel) measurable map. Since r is the composition $r =$ rank $\langle M\rangle$, it is predictable. Since $\langle M\rangle_t$ and $\langle M\rangle_t^+$ are both symmetric and since they commute, it follows from Lancaster and Tismenetsky (1985, p. 432), that $\langle M\rangle_t^+$ is the Moore–Penrose inverse of $\langle M\rangle_t$. To show predictability of $\langle M\rangle^+$, we need the following characterization of the Moore–Penrose inverse for any real matrix R: $R^+ = \lim_{n\to\infty}(R^{\mathrm{T}}R + (1/n)I)^{-1}R^{\mathrm{T}}$. That this characterization holds is easily seen for positive semidenfinite matrices (see the appendix). For the general case with a considerably more difficult proof we refer to Rao and Mitra (1971, Theorem 3.5.3). Apply this characterization to $\langle M\rangle_S$ for any stopping time S to obtain that $\langle M\rangle^+$ as a limit of predictable processes is predictable too. □

Remark. Proposition 2.1 really needs a proof, since another generalized inverse of $\langle M\rangle_t$ may not yield a predictable process. Consider the following example.

$$\langle M\rangle_t = \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix}.$$

Let a_t be an artibrary stochastic process, possibly not adapted. Then for $t > 0$,

$$\begin{bmatrix} 1/t & a_t \\ a_t & ta_t^2 \end{bmatrix}$$

is a generalized inverse of $\langle M\rangle_t$, different from the Moore–Penrose inverse (which corresponds with $a_t = 0$), and viewed as a stochastic process it is in general not predictable.

Example 1. Let N be the standard Poisson process. Define $T = \inf\{t > 0: N_t = 1\}$. Then T is a totally inaccessible stopping time. Define now the martingale M by $M_t = N_t - t - (N_{t\wedge T} - t \wedge T)$. Then $\langle M\rangle_t = t - t \wedge T$. But now $T_1 = \inf\{t > 0: \langle M\rangle_t > 0\} = T$. So T_1 is not predictable. Notice that $r_t = 1_{\{t>T\}}$ is predictable.

Let now k be an arbitrary $\mathscr{F}_T$-measurable random variable and $\kappa_t = k1_{\{t>T\}}$. Then κ is predictable (cf. Bremaud, 1981, p. 304) and $m = \kappa.M$ (the dot means stochastic integration) is a martingale indistinguishable from the product κM. Although κ is not a constant, it is strightforward to show that $c(m, M) = 0$ (see the introduction for the definition of $c(m, M)$). We will return to this in Section 4.

We need some technical properties of M and $\langle M\rangle$, to be used in Section 3. These are formulated in the next three lemmas. In the notation introduced above we have the following:

Lemma 2.2. *On the set* $\{T_k < \infty\}$ *we have*:

(i) $V_{T_k-}(k-1) = \lim_{t\uparrow T_k} V_t(k-1)$ *exists and is invertible.*

(ii) *If* $F(k) = G(k)$, *then* $\lim_{t\downarrow T_k} V_t(k) = W(k)$. *If* $F(k) = [G(k), U_1(k)]$, *with* $U_1(k)$ *nontrivial, then we can write*

$$V_t(k) = R_t(k)R_t(k)^{\mathrm{T}}$$

with

$$R_t(k) = \begin{bmatrix} a_t(k) & b_t(k) \\ 0 & c_t(k) \end{bmatrix},$$

decomposed in blocks of appropriate sizes such that $\lim_{t\downarrow T_k} b_t(k) = 0$, $\lim_{t\downarrow T_k} c_t(k) = 0$ *and* $\lim_{t\downarrow T_k} a_t(k)a_r(k)^{\mathrm{T}} = W(k)$.

Proof. (i) is obvious.

(ii) If $F(k) = G(k)$, then right continuity of $\langle M\rangle$ yields the result. Assume therefore that $F(k) = [G(k), U_1(k)]$. Then

$$\langle M\rangle_{T_k} = F(k)\begin{bmatrix} W(k) & 0 \\ 0 & 0 \end{bmatrix}F(k)^{\mathrm{T}},$$

with the zero blocks of appropriate dimension.

Decompose $V_t(k)$ in blocks of the same dimension as

$$\begin{bmatrix} V_t(k)_{11} & V_t(k)_{12} \\ V_t(k)_{21} & V_t(k)_{22} \end{bmatrix}.$$

Since $V_t(k) > 0$, we also have $V_t(k)_{22} > 0$. Since on $]\!]T_k, T_{k+1}[\![$ also $\langle M\rangle_t - \langle M\rangle_{T_k} \geqslant 0$, we have that

$$\begin{bmatrix} V_t(k)_{11} - W(k) & V_t(k)_{12} \\ V_t(k)_{21} & V_t(k)_{22} \end{bmatrix} \geqslant 0.$$

Hence

$$V_t(k)_{11} - W(k) - V_t(k)_{12}V_t(k)_{22}^{-1}V_t(k)_{21} \geqslant 0.$$

Use the decomposition $V_t(k) = R_t(k)R_t(k)^{\mathrm{T}}$ to write this inequality as

$$a_t(k)a_t(k)^{\mathrm{T}} + b_t(k)b_t(k)^{\mathrm{T}} - W(k)$$
$$- b_t(k)c_t(k)^{\mathrm{T}}[c_t(k)c_t(k)^{\mathrm{T}}]^{-1}c_t(k)b_t(k)^{\mathrm{T}} \geqslant 0.$$

But $c_t(k)$ is invertible, so this inequality becomes

$$a_t(k)a_t(k)^{\mathrm{T}} - W(k) \geqslant 0. \tag{2.3}$$

Right continuity of $\langle M\rangle$ gives $\lim_{t\downarrow T_k} V_t(k)_{11} = W(k)$. So

$$0 = \lim_{t\downarrow T_k}[V_t(k)_{11} - W(k)] = \lim_{t\downarrow T_k}[(a_t(k)a_t(k)^{\mathrm{T}} - W(k)) + b_t(k)b_t(k)^{\mathrm{T}}].$$

The term in brackets is because of equation (2.3) the sum of two nonnegative matrices. Hence $\lim_{t\downarrow T_k} a_t(k)a_t(k)^{\mathrm{T}} = W(k)$ and $\lim_{t\downarrow T_k} b_t(k) = 0$. Because $\lim_{t\downarrow T_k} V_t(k)_{22} = 0$, we obtain $\lim_{t\downarrow T_k} c_t(k) = 0$. □

Introduce the following notation. $P_t = \langle M\rangle_t \langle M\rangle_t^+$. Observe that P_t for fixed (t, ω) is the orthogonal projection on $\mathrm{Im}\langle M\rangle_t$ along $\mathrm{Ker}\langle M\rangle_t$. P as a process doesn't depend on t on $]\!] T_k, T_{k+1}[\![$. It is, like r, nor right or left continuous at the T_k, although (trivially) left and right limits exist and are finite. Furthermore, for $t > s$, we have $P_t P_s = P_s P_t = P_s$, because $\mathrm{Im}\langle M\rangle_s \subset \mathrm{Im}\langle M\rangle_t$.

Lemma 2.3. *M is indistinguishable from the stochastic integral $P.M$ and from the product PM.*

Proof. P is predictable (from Proposition 2.1). Hence $P.M$ defines again a martingale. Then

$$\langle M - P.M\rangle = \langle (I-P).M\rangle = \int_0^{\cdot} (I-P)\,\mathrm{d}\langle M\rangle (I-P)^{\mathrm{T}}.$$

On $]\!] T_k, T_{k+1}[\![$ we have

$$P\,\mathrm{d}\langle M\rangle = \mathrm{d}(P\langle M\rangle) = \mathrm{d}\langle M\rangle$$

which makes the integral zero over $]\!] T_k, T_{k+1}[\![$. On $\{T_k < \infty\}$ we can apply the same argument if $P_{T_k} = P_{T_k-}$. Otherwise we get

$$\begin{aligned}(I - P_{T_k})\Delta\langle M\rangle_{T_k} &= (I-P_{T_k})[\langle M\rangle_{T_k} - \langle M\rangle_{T_k-}]\\ &= -(I-P_{T_k})\langle M\rangle_{T_k-} = -(I-P_{T_k})P_{T_k-}\langle M\rangle_{T_k-} = 0,\end{aligned}$$

since $P_{T_k}P_{T_k-} = P_{T_k-}$. Hence $\langle M - P.M\rangle$ is indistinguishable from the zero process. Consider now the product PM. On $]\!] T_k, T_{k+1}[\![$ we have $\mathrm{d}(PM) = P\,\mathrm{d}M$. Let $T_1 < \infty$. Then

$$P_{T_1}M_{T_1} = P_{T_1}\Delta M_{T_1} = \Delta(P.M)_{T_1} = \Delta M_{T_1} = M_{T_1}.$$

Now we use an induction argument. Let $T_k < \infty$ and assume that $P_{T_{k-1}}M_{T_{k-1}} = M_{T_{k-1}}$. Then

$$\begin{aligned}\Delta(P_{T_k}M_{T_k}) &= P_{T_k+}M_{T_k} - P_{T_k-}M_{T_k-}\\ &= P_{T_k+}\Delta M_{T_k} + (P_{T_k+} - P_{T_k-})M_{T_k-}\\ &= \Delta M_{T_k} + (P_{T_k+} - P_{T_k-})(M_{T_k-} - M_{T_{k-1}}) + (P_{T_k+} - P_{T_k-})P_{T_{k-1}}M_{T_{k-1}}\\ &= \Delta M_{T_k} + (P_{T_k+} - P_{T_k-})\int_{(T_{k-1}, T_k)} P\,\mathrm{d}M + 0\\ &= \Delta M_{T_k} + (P_{T_k+} - P_{T_k-})P_{T_k-}\int_{(T_{k-1}, T_k)} \mathrm{d}M = \Delta M_{T_k}.\end{aligned}$$

Hence PM and M are indistinguishable. □

The covariation process $\langle m, M\rangle$ enjoys the following property.

Lemma 2.4. $\langle m, M\rangle = \langle m, M\rangle P$.

Proof.

$$\begin{aligned}\langle m, M\rangle_t P_t 1_{]\!]T_k,T_{k+1}[\![} &= \int_{[0,t]} 1_{]\!]T_k,T_{k+1}[\![}\,\mathrm{d}\langle m, M\rangle_s P_s \\ &= \int_{[0,t]} 1_{]\!]T_k,T_{k+1}[\![}\,\mathrm{d}\langle m, P.M\rangle_s \\ &= \int_{[0,t]} 1_{]\!]T_k,T_{k+1}[\![}\,\mathrm{d}\langle m, M\rangle_s \quad \text{(by Lemma 2.3)} \\ &= \langle m, M\rangle_t 1_{]\!]T_k,T_{k+1}[\![}.\end{aligned}$$

On $\{T_k < \infty\}$ we have

$$\begin{aligned}\langle m, M\rangle_{T_k} P_{T_k} &= \Delta\langle m, M\rangle_{T_k} P_{T_k} + \langle m, M\rangle_{T_k-}(P_{T_k} - P_{T_k-}) + \langle m, M\rangle_{T_k-} P_{T_k-} \\ &= \Delta\langle m, P.M\rangle_{T_k} + \langle m, M\rangle_{T_k-}(P_{T_k} - P_{T_k-}) + \langle m, M\rangle_{T_k-} P_{T_k-} \\ &= \Delta\langle m, M\rangle_{T_k} + \langle m, M\rangle_{T_k-} P_{T_k-},\end{aligned}$$

because the second term equals zero, as can be seen by the first part of the proof and by using an induction argument like in the proof of Lemma 2.3. By the same argument it follows that

$$\langle m, M\rangle_{T_k-} P_{T_k-} = \lim_{t\uparrow T_k} \langle m, M\rangle_t P_t = \lim_{t\uparrow T_k} \langle m, M\rangle_t = \langle m, M\rangle_{T_k-}.$$

So $\langle m, M\rangle_{T_k} P_{T_k} = \langle m, M\rangle_{T_k}$. Combining this with the first part of the proof we get $\langle m, M\rangle = \langle m, M\rangle P$. □

Remark. Lemmas 2.3 and 2.4 as well the results in subsequent sections can be generalized by taking other generalized inverses of $\langle M\rangle$. Consider for instance once more the example in the remark after Proposition 2.1 Then

$$P_t = \begin{bmatrix} 1 & ta_t \\ 0 & 0 \end{bmatrix}.$$

Write

$$M_t = \begin{bmatrix} W_t \\ 0 \end{bmatrix},$$

with W a standard Brownian motion, then $P_t M_t = (P.M)_t = M_t$. However one is in general faced with the problem that (although P_t is still a (non-orthogonal) projection) P fails to be a predictable process. So, proving a statement like in Lemma 2.3 will certainly be more difficult. We will leave this possible generalization aside, since for our purposes the specific choice of the Moore–Penrose inverse suffices.

3. The process $c(m, M)$

Let m and M be as in Section 1. Define the predictable process (related to the correlation between m and M) $c(m, M): \Omega \times [0, \infty) \to \mathbb{R}^{k\times k}$ by

$$c(m, M) = \langle m \rangle - \langle m, M \rangle \langle M \rangle^{+} \langle M, m \rangle.$$

The main result of this section is an integral representation for $c(m, M)$. The difficulty that we encounter is that $\langle M \rangle^{+}$ and even $\langle m, M \rangle \langle M \rangle^{+}$ may not be right continuous. See example 2. Typically right limits of $\langle M \rangle^{+}$ at the T_k are not finite. Take for example the trivial case where $\langle M \rangle_t = t - t \wedge 1$, then $\langle M \rangle_t^{+} = 1/(t-1)$, for $t > 1$. Therefore we need some agreements concerning the notation that we will follow. The considerations above forbid us to define $\Delta \langle M \rangle_t^{+}$ as $\langle M \rangle_{t+}^{+} - \langle M \rangle_{t-}^{+}$. Therefore we adopt the *convention*

$$\Delta \langle M \rangle_t^{+} = \langle M \rangle_t^{+} - \langle M \rangle_{t-}^{+}.$$

All integrals of the type $J_t = \int_{[0,t]} \alpha \, \mathrm{d}\langle M \rangle^{+}$ are then to be understood such that $\Delta J_t = \alpha_t \Delta \langle M \rangle_t^{+} = \alpha_t (\langle M \rangle_t^{+} - \langle M \rangle_{t-}^{+})$, provided of sourse that α is such that this convention makes sense, for instance it is such that J is right continuous.

We need the following representation result (cf. Lipster and Shiryaev, 1990, pp. 112, 113 for the univariate case; the proof of the multivariate case proceeds along the same lines).

Lemma 3.1. *There exists a (in general not unique) predictable process $\kappa : \Omega \times [0, \infty) \to \mathbb{R}^{k\times n}$, such that $m - \kappa.M$ is an $\mathbb{R}^k$-valued locally square integrable martingale, orthogonal to M in the sense that $\langle m - \kappa.M, M \rangle = 0$. However the martingale $m - \kappa.M$ is uniquely defined (up to indistinguishability).* □

With a process κ as in Lemma 3.1 we can write

$$c(m, M) = \langle m - \kappa.M \rangle + \langle \kappa.M \rangle - \langle m, M \rangle \langle M \rangle^{+} \langle M, m \rangle$$

$$= \langle m - \kappa.M \rangle + c(\kappa.M, M).$$

The proof of Theorem 3.3 below involves some calculus rules. As for $\langle M \rangle^{+}$, we also use for P the notation $\Delta P_t = P_t - P_{t-}$.

Lemma 3.2. (i) $\mathrm{d}\langle M\rangle_t \langle M\rangle^+_{t-} = -\langle M\rangle_t \, \mathrm{d}\langle M\rangle^+_t + \mathrm{d}P_t$.

(ii) $\mathrm{d}\langle M\rangle_t = -\langle M\rangle_{t-} \, \mathrm{d}\langle M\rangle^+_t \langle M\rangle_t + \mathrm{d}P_t \langle M\rangle_t$.

Proof. On $]\!]T_k, T_{k+1}[\![$ the ordinary calculus rules apply to $V_t(k)$ and P doesn't vary with t on this stochastic interval. Hence the result follows in this case. Consider now what happens if $t = T_k < \infty$. If $\langle M\rangle$ happens to be left continuous at this point we are back in the previous case. So assume that $\Delta\langle M\rangle_{T_k} \neq 0$. Then

$$\Delta\langle M\rangle_{T_k}\langle M\rangle^+_{T_k-} + \langle M\rangle_{T_k}\Delta\langle M\rangle^+_{T_k} = \langle M\rangle_{T_k}\langle M\rangle^+_{T_k} - \langle M\rangle_{T_k-}\langle M\rangle^+_{T_k-} = \Delta P_{T_k}.$$

This proves (i). Similarly we have

$$\begin{aligned}
&\Delta\langle M\rangle_{T_k} + \langle M\rangle_{T_k-}\Delta\langle M\rangle^+_{T_k}\langle M\rangle_{T_k} \\
&\quad = \langle M\rangle_{T_k} - \langle M\rangle_{T_k-} + \langle M\rangle_{T_k-}P_{T_k} - P_{T_k-}\langle M\rangle_{T_k} \\
&\quad = (I - P_{T_k-})\langle M\rangle_{T_k-}\langle M\rangle_{T_k-}(I - P_{T_k}) = \Delta P_{T_k}\langle M\rangle_{T_k}
\end{aligned}$$

which proves the second assertion. □

In the notation that we introduced above we are now able to present the principal result of this section.

Theorem 3.3. (i) *$c(m, M)$ is a right continuous process.*

(ii) *With κ as in Lemma 3.1 we have for $m = \kappa.M$ the following integral representation*:

$$\begin{aligned}
c(m, M) &= -\int (\kappa\langle M\rangle - \langle m, M\rangle)\, \mathrm{d}\langle M\rangle^+ (\kappa\langle M\rangle - \langle m, M\rangle)^{\mathrm{T}} \\
&= -\int (\kappa - \langle m, M\rangle\langle M\rangle^+)\langle M\rangle \, \mathrm{d}\langle M\rangle^+ \langle M\rangle(\kappa - \langle m, M\rangle\langle M\rangle^+)^{\mathrm{T}} \\
&= -\int (\kappa - \langle m, M\rangle_-\langle M\rangle^+_-)\langle M\rangle_- \, \mathrm{d}\langle M\rangle^+ \langle M\rangle_-(\kappa - \langle m, M\rangle_-\langle M\rangle^+_-)^{\mathrm{T}} \\
&= +\int (\kappa - \langle m, M\rangle_-\langle M\rangle^+_-)(I - \Delta\langle M\rangle\langle M\rangle^+)\, \mathrm{d}\langle M\rangle \\
&\qquad \times (\kappa - \langle m, M\rangle_-\langle M\rangle^+_-)^{\mathrm{T}}.
\end{aligned}$$

Proof. (i) This is a simple consequence of right continuity of all involved processes if we restrict our attention to the open intervals $]\!]T_k, T_{k+1}[\![$. Therefore we consider what happens at the T_k (on $\{T_k < \infty\}$). Define the process q on $]\!]T_k, T_{k+1}[\![$ by $q_t = \langle m, M\rangle_t F(k) R_t(k)^{-1}$, by $q_t = \langle m, M\rangle_t F(k) R_t(k)^{-1}$, where $R_t(k)$ is as in Lemma 2.2. We will show that $\lim_{t\downarrow T_k} q_t q_k^{\mathrm{T}}$ exists. Write

$$q_t = q^1_t + q^2_t,$$

with

$$q_t^1 = \langle m, M\rangle_{T_k} F(k) R_t(k)^{-\mathrm{T}}$$

and

$$q_t^2 = (\langle m, M\rangle_t - \langle m, M\rangle_{T_k}) F(k) R_t(k)^{-\mathrm{T}}.$$

First we will show that $\lim_{t\downarrow T_k} q_t^2 = 0$. It is sufficient to prove that $\mathrm{tr}[q_t^2(q_t^2)^{\mathrm{T}}]$ tends to zero for $t\downarrow T_k$. Write

$$q_t^2(q_t^2)^{\mathrm{T}} = \int_{(T_k,t]} \kappa\, \mathrm{d}\langle M\rangle \langle M\rangle_t^+ \int_{(T_k,t]} \mathrm{d}\langle M\rangle \kappa^{\mathrm{T}} \geqslant 0.$$

Let κ_i be the ith row of κ and write $\langle M\rangle_t^+ = \sum_{j=1}^n Q_{jt}Q_{jt}^{\mathrm{T}}$, where the Q_{jt} are $\mathbb{R}^n$ valued random variables and $Q_{it}^T Q_{jt} = 0$ if $i \neq j$. Then

$$\mathrm{tr}(q_t^2 q_t^{2\mathrm{T}}) = \sum_{i,j}\left[\int_{(T_k,t]} \kappa_i\, \mathrm{d}\langle M\rangle Q_{jt}\right]^2,$$

which is by Schwarz' inequality less than

$$\begin{aligned}
&\sum_{i,j}\int_{(T_k,t]} \kappa_i\, \mathrm{d}\langle M\rangle \kappa_i^{\mathrm{T}} \int_{(T_k,t]} Q_{jt}^{\mathrm{T}}\, \mathrm{d}\langle M\rangle Q_{jt}\\
&\quad = \sum_i \int_{(T_k,t]} \kappa_i\, \mathrm{d}\langle M\rangle \kappa_i^{\mathrm{T}} \sum_j Q_{jt}^{\mathrm{T}}(\langle M\rangle_t - \langle M\rangle_{T_k}) Q_{jt}\\
&\quad = \mathrm{tr}\int_{(T_k,t]} \kappa\, \mathrm{d}\langle M\rangle \kappa^{\mathrm{T}}\, \mathrm{tr}[(\langle M\rangle_t - \langle M\rangle_{T_k})\langle M\rangle_t^+]. \qquad (3.1)
\end{aligned}$$

The first factor of this product tends to zero as $t\downarrow T_k$. Consider now the seond factor. First we notice that

$$\mathrm{tr}[\langle M\rangle_t \langle M\rangle_t^+] = \mathrm{tr}[F(k)F(k)^{\mathrm{T}}] = \mathrm{tr}[F(k)^{\mathrm{T}}F(k)] = r_t.$$

(Remember that $r_t = \mathrm{rank}\langle M\rangle_t$.) Next we compute

$$\begin{aligned}
\mathrm{tr}[\langle M\rangle_{T_k}\langle M\rangle_t^+] &= \mathrm{tr}[G(k)W(k)G(k)^{\mathrm{T}}F(k)V_t(k)^{-1}F(k)^{\mathrm{T}}]\\
&= \mathrm{tr}[V_t(k)^{-1}F(k)^{\mathrm{T}}G(k)W(k)G(k)^{\mathrm{T}}F(k)]\\
&= \mathrm{tr}\left[V_t(k)^{-1}\begin{bmatrix} W(k) & 0\\ 0 & 0\end{bmatrix}\right]\\
&= \mathrm{tr}\left[R_t(k)^{-1}\begin{bmatrix} W(k) & 0\\ 0 & 0\end{bmatrix}R_t(k)^{-\mathrm{T}}\right]\\
&= \mathrm{tr}\left\{\begin{bmatrix} a_t(k)^{-1} & *\\ 0 & *\end{bmatrix}\begin{bmatrix} W(k) & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix} a_t(k)^{-\mathrm{T}} & 0\\ * & *\end{bmatrix}\right\}\\
&= \mathrm{tr}\begin{bmatrix} (a_t(k)a_t(k)^{\mathrm{T}})-{}^{1}W(k) & 0\\ 0 & 0\end{bmatrix}\\
&= \mathrm{tr}[(a_t(k)a_t(k)^{\mathrm{T}})^{-1}W(k)]
\end{aligned}$$

which tends to $\operatorname{tr}[W(k)^{-1}W(k)] = r_{T_k}$. Hence $\lim_{t\downarrow T_k}[(\langle M\rangle_t - \langle M\rangle_{T_k})\langle M\rangle_t^+] = r_{T_k+} - r_{T_k} < \infty$. So from equation (3.1) we obtain that indeed $q_t^2 \to 0$ as $t \downarrow T_k$. Secondly we look at q_t^1. From Lemma 2.4 we see that there exists a random matrix $A(k)$ such that $\langle m, M\rangle_{T_k} = A(k)G(k)^{\mathrm{T}}$. Hence

$$\begin{aligned} q_t^1 &= A(k)G(k)^{\mathrm{T}}F(k)R_t(k)^{-\mathrm{T}} \\ &= A(k)[I\ \ 0]\begin{bmatrix} a_t(k)^{-\mathrm{T}} & 0 \\ * & * \end{bmatrix} \\ &= A(k)[a_t(k)^{-\mathrm{T}}\ \ 0]. \end{aligned}$$

So

$$q_t^1(q_t^1)^{\mathrm{T}} = A(k)(a_t(k)a_t(k)^{\mathrm{T}})^{-1}A(k)^{\mathrm{T}} \to A(k)W(k)^{-1}A(k)^{\mathrm{T}},$$

since $W(k)$ is invertible and $a_t(k)a_t(k)^{\mathrm{T}} \to W(k)$ by Lemma 2.2. Because of the fact that $\lim_{t\downarrow T_k} q_t^2 = 0$, and that $a_t(k)$ is bounded for $t\downarrow T_k$, we get

$$\lim_{t\downarrow T_k} q_t q_t^{\mathrm{T}} = \lim_{t\downarrow T_k} q_t^1(q_t^1)^{\mathrm{T}} = A(k)W(k)^{-1}A(k)^{\mathrm{T}}.$$

But

$$\begin{aligned} \langle m, M\rangle_{T_k}\langle M\rangle_{T_k}^+\langle M, m\rangle_{T_k} &= A(k)G(k)^{\mathrm{T}}G(k)W(k)^{-1}G(k)^{\mathrm{T}}G(k)A(k)^{\mathrm{T}} \\ &= A(k)W(k)^{-1}A(k)^{\mathrm{T}}, \end{aligned}$$

which gives right continuity of $\langle m, M\rangle\langle M\rangle^+\langle M, m\rangle$ at the T_k (on $\{T_k < \infty\}$), thus proving the first assertion of the theorem. In order to prove the second one we proceed as follows. Because $c(m, M)$ is right continuous we can use the results of Lemma 3.2 in the computations below.

$$\begin{aligned} \mathrm{d}c(m, M) &= \kappa\,\mathrm{d}\langle M\rangle\kappa^{\mathrm{T}} - \langle m, M\rangle_-\langle M\rangle_-^+\,\mathrm{d}\langle M\rangle\kappa^{\mathrm{T}} \\ &\quad - \langle m, M\rangle_-\,\mathrm{d}\langle M\rangle^+\langle M, m\rangle - \kappa\,\mathrm{d}\langle M\rangle\langle M\rangle^+\langle M, m\rangle \end{aligned} \tag{3.2}$$

from which we obtain by Lemma 3.2,

$$\begin{aligned} \mathrm{d}c(m, M) &= -(\kappa\langle M\rangle_- - \langle m, M\rangle_-)\,\mathrm{d}\langle M\rangle^+(\kappa\langle M\rangle - \langle m, M\rangle)^{\mathrm{T}} \\ &\quad + \kappa\,\mathrm{d}P\langle M\rangle\kappa^{\mathrm{T}} - \langle m, M\rangle_-\,\mathrm{d}P\kappa^{\mathrm{T}} - \kappa\,\mathrm{d}P\langle M, m\rangle \qquad (3.3) \\ &= -(\kappa\langle M\rangle - \langle m, M\rangle)\,\mathrm{d}\langle M\rangle^+(\kappa\langle M\rangle - \langle m, M\rangle)^{\mathrm{T}} \\ &\quad + \kappa\,\mathrm{d}\langle M\rangle_-\kappa^{\mathrm{T}} - \langle m, M\rangle_-\,\mathrm{d}P\kappa^{\mathrm{T}} - \kappa\,\mathrm{d}P\langle M, m\rangle_-. \qquad (3.4) \end{aligned}$$

It is immediately seen that on $]\!]T_k, T_{k+1}[\![$ the last three terms vanish, whereas on $\{T_k < \infty\}$ we have

$$\Delta P_{T_k}\langle M\rangle_{T_k-} = \Delta P_{T_k}P_{T_k-}\langle M\rangle_{T_k-} = 0$$

and

$$\langle m, M\rangle_{T_k}\Delta P_{T_k} = \langle m, M\rangle_{T_k}P_{T_k-}\Delta P_{T_k} = 0,$$

since $P_{T_k-}\Delta P_{T_k} = 0$. This proves the first formula of the second assertion. The other ones follow similarly. □

Remark. At $t=T_k$ it is not true that $\Delta\langle M\rangle_t^+\leq 0$ and that $\langle M\rangle_t\Delta\langle M\rangle_t^+\langle M\rangle_t\leq 0$. However for all t one has $\langle M\rangle_{t-}\Delta\langle M\rangle_t^+\langle M\rangle_{t-}\leq 0$. This is trivially true on the open intervals $]\!]T_k, T_{k+1}[\![$. Consider what happens at T_k on $\{T_k<\infty\}$ if $\Delta\langle M\rangle_{T_k}\neq 0$. We know that $G(k)W(k)G(k)^{\mathrm{T}}-F(k-1)V_{T_k-}(k-1)F(k-1)^{\mathrm{T}}\geq 0$ or, with an obvious decomposition of $W(k)$:

$$\begin{bmatrix} W(k)_{11}-V_{T_k-}(k-1) & W(k)_{21} \\ W(k)_{12} & W(k)_{22}\end{bmatrix}\geq 0.$$

Hence, since $W(k)_{22}$ is invertible, we get

$$W(k)_{11}-W(k)_{12}W(k)_{22}^{-1}W(k)_{21}-V_{T_k-}(k-1)\geq 0. \tag{3.5}$$

Now look at

$$\begin{aligned}&\langle M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+\langle M\rangle_{T_k-}\\ &\quad=\langle M\rangle_{T_k-}\langle M\rangle_{T_k}^+\langle M\rangle_{T_k-}-\langle M\rangle_{T_k-}\\ &\quad=F(k-1)V_{T_k-}(k-1)[F(k-1)^{\mathrm{T}}G(k)W(k)^{-1}G(k)^{\mathrm{T}}F(k-1)\\ &\qquad\qquad -V_{T_k-}(k-1)^{-1}]V_{T_k-}(k-1)F(k-1)^{\mathrm{T}}.\end{aligned}$$

Consider the term in brackets. Again in obvious notation, it becomes

$$\begin{aligned}&[W(k)^{-1}]_{11}-V_{T_k-}(k-1)^{-1}\\ &\quad=[W(k)_{11}-W(k)_{12}W(k)_{22}^{-1}W(k)_{21}]^{-1}-V_{T_k-}(k-1)^{-1}\leq 0,\end{aligned}$$

from equation (3.5). Thus we have proved the following:

Corollary 3.4. *The process $c(m, M)$ is nondecreasing.* □

4. Linear dependence

In this section we will study a suitably defined notion of linear dependence between two square integrable martingales m and M. By analogy with the situation in which one deals with multidimensional random variables we have the following:

Definition 4.1. (i) m is said to be linearly dependent on M if the process $c(m, M)\in\mathbb{R}^{k\times k}$ is indistinguishable from zero.

(ii) m and M are said to be mutually linearly dependent if both $c(m, M)$ and $c(M, m)$ are indistinguishable from zero.

Here is the main result of this section.

Theorem 4.2. *m is linearly dependent on M iff there exists a (possibly random) matrix $C\in\mathbb{R}^{k\times n}$ with $C\langle M\rangle$ a predictable process such that $m=CM$. Moreover in this case $C\langle M\rangle=\langle m, M\rangle$. Furthermore m and M are mutually linearly dependent iff there exist matrices C_1 and C_2 such that $m=C_1M$ and $M=C_2m$. In the latter case we also have that C_1 and C_2 are each others Moore-Penrose inverses.*

Remark. The matrix C in Theorem 4.2 is not necessarily $\mathscr{F}_0$-measurable. See Example 1 in Section 2 and Example 3 below.

Proof of Theorem 4.2. Define $\gamma_t = \langle m, M\rangle_t \langle M\rangle_t^+$. Then $\gamma_t \langle M\rangle_t = \langle m, M\rangle_t$ from Lemma 2.4. On $]\!]T_k, T_{k+1}[\![$ we have

$$\begin{aligned} \mathrm{d}\gamma_t &= \langle m, M\rangle_{t-}\, \mathrm{d}\langle M\rangle_t^+ + \mathrm{d}\langle m, M\rangle_t \langle M\rangle_t^+ \\ &= \gamma_{t-}\langle M\rangle_{t-}\, \mathrm{d}\langle M\rangle_t^+ + \kappa_t\, \mathrm{d}\langle M\rangle_t \langle M\rangle_t^+ \\ &= (\gamma_{t-} - \kappa_t)\langle M\rangle_{t-}\, \mathrm{d}\langle M\rangle_t^+ . \end{aligned}$$

So if $c(m, M) = 0$, then from Theorem 3.3 we obtain that γ is constant on $]\!]T_k, T_{k+1}[\![$. This also implies that γ admits right limits at T_k if $T_k < \infty$. We need some more properties of γ. On $\{T_k < \infty\}$ we have

$$(\gamma_{T_k+} - \gamma_{T_k})G(k) = 0, \tag{4.1}$$

$$\gamma_{T_k} - \gamma_{T_k-} = \kappa_{T_k}[G(k)G(k)^{\mathrm{T}} - F(k-1)F(k-1)^{\mathrm{T}}] = \kappa_{T_k}\Delta P_{T_k}. \tag{4.2}$$

Indeed right continuity of $\langle m, M\rangle$ gives

$$\gamma_{T_k}\langle M\rangle_{T_k} = \langle m, M\rangle_{T_k} = \lim_{t\downarrow T_k}\langle m, M\rangle_t = \lim_{t\downarrow T_k} \gamma_t\langle M\rangle_t = \gamma_{T_k+}\langle M\rangle_{T_k}.$$

Hence $(\gamma_{T_k+} - \gamma_{T_k})\langle M\rangle_{T_k} = 0$, which is equivalent to equation (4.1). Next we use Lemma 3.2 to write

$$\begin{aligned} \gamma_{T_k} - \gamma_{T_k-} &= \langle m, M\rangle_{T_k}\langle M\rangle_{T_k}^+ - \langle m, M\rangle_{T_k-}\langle M\rangle_{T_k-}^+ \\ &= \langle m, M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+ + \kappa_{T_k}\Delta\langle M\rangle_{T_k}\langle M\rangle_{T_k}^+ \\ &= \gamma_{T_k-}\langle M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+ + \kappa_{T_k}\Delta\langle M\rangle_{T_k}\langle M\rangle_{T_k}^+ \\ &= \gamma_{T_k-}\langle M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+ - \kappa_{T_k}\langle M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+ + \kappa_{T_k}\Delta P_{T_k} \\ &= (\gamma_{T_k-} - \kappa_{T_k})\langle M\rangle_{T_k-}\Delta\langle M\rangle_{T_k}^+ + \kappa_{T_k}\Delta P_{T_k}. \end{aligned}$$

The assumption that $c(m, M) = 0$ yields the first term zero from Theorem 3.3, which gives equation (4.2). Notice that equation (4.1) and equation (4.2) imply

$$(\gamma_{T_k} - \gamma_{T_k-})\langle M\rangle_{T_k-} = 0, \qquad (\gamma_{T_{k+}} - \gamma_{T_k})\langle M\rangle_{T_k} = 0. \tag{4.3}$$

Hence $\gamma_{T_k}\langle M\rangle_{T_k-} = 0$ and $\Delta(\gamma_{T_k}\langle M\rangle_{T_k}) = \gamma_{T_k}\Delta\langle M\rangle_{T_k}$, or $\Delta\langle m, M\rangle_{T_k} = \gamma_{T_k}\Delta\langle M\rangle_{T_k}$. Define now $C = \lim_{t\to\infty}\gamma_t$. We claim that this is the matrix in the assertion of the theorem. Notice that on the set $\Omega_k = \{T_k < \infty, T_{k+1} = \infty\}$ C equals $\gamma_{T_{k+}}$. Furthermore $\bigcup_{k=0}^n \Omega_k = \Omega$ and $\Omega_k \cap \Omega_l = \emptyset$ if $k \neq l$. First we prove the following facts. CM is a martingale and $CM_t = \gamma_t M_t = (\gamma.M)_t$.

From Lemma 2.3: $CM_t = C\langle M\rangle_t\langle M\rangle_t^+ M_t$. On Ω_k we have for $j \leq k$,

$$\begin{aligned} C\langle M\rangle_{T_j} = \gamma_{T_k+}\langle M\rangle_{T_j} &= \sum_{i=1}^{k}(\gamma_{T_i+} - \gamma_{T_{i-1}+})\langle M\rangle_{T_j} \\ &= \sum_{i=1}^{k}(\gamma_{T_i+} - \gamma_{T_i-})\langle M\rangle_{T_j} = \gamma_{T_j+}\langle M\rangle_{T_j}, \end{aligned}$$

since $(\gamma_{T_i+} - \gamma_{T_i-})\langle M\rangle_{T_j} = 0$ if $i<j$. But

$$\gamma_{T_j+}\langle M\rangle_{T_j} = (\gamma_{T_j+} - \gamma_{T_j})\langle M\rangle_{T_j} + \gamma_{T_j}\langle M\rangle_{T_j} = \gamma_{T_j}\langle M\rangle_{T_j}$$

by equation (4.3).

Furthermore on $\Omega_k \times [0,\infty) \cap]\!]T_j, T_{j+1}[\![$ we have in the same way $C\langle M\rangle_t = \gamma_{T_j+}\langle M\rangle_t$, because $(\gamma_{T_i} + - \gamma_{T_i-})F(j) = 0$ if $j<i$ and so $C\langle M\rangle$ is equal to $\gamma\langle M\rangle$. Hence

$$CM_t = \gamma_t M_t = (\gamma.M)_t + \int_{[0,t]} \mathrm{d}\gamma_s M_{s-}.$$

Now on Ω_k for $j \leqslant k$ we have

$$\Delta\gamma_{T_j} M_{T_j-} = \Delta\gamma_{T_j}\langle M\rangle_{T_j}\langle M\rangle^+_{T_j} M_{T_j-} = 0.$$

Hence

$$\int_{[0,t]} \mathrm{d}\gamma_s M_{s-} = \sum_{T_j \leqslant t} \Delta\gamma_{T_j} M_{T_j-} = 0.$$

Predictability of γ (Lemma 2.3) gives that $CM = \gamma.M$ is indeed a martingale.

Finally we have to show that m and CM are indistinguishable. Compute

$$\langle m - CM\rangle = \langle m - \gamma.M\rangle = \langle(\kappa-\gamma).M\rangle = \int_{[0,t]} (\kappa-\gamma)\,\mathrm{d}\langle M\rangle(\kappa-\gamma)^{\mathrm{T}}.$$

Consider

$$(\kappa-\gamma)_t\,\mathrm{d}\langle M\rangle_t = \mathrm{d}\langle m, M\rangle_t - \gamma_t\,\mathrm{d}\langle M\rangle_t = \mathrm{d}(\gamma_t\langle M\rangle_t) - \gamma_t\,\mathrm{d}\langle M\rangle_t = \mathrm{d}\gamma_t\langle M\rangle_{t-},$$

which is zero on all $]\!]T_k, T_{k+1}[\![$, because here $\mathrm{d}\gamma_t = 0$. At $t = T_k < \infty$ we also get zero from equation (4.2). This proves the only if part.

Next we prove the converse statement. Assume that $C\langle M\rangle$ is predictable, equivalently CP is predictable. Then the product $m = CM$ is a martingale. Indeed $CM = CPM$ is adapted. Let now $\gamma = CP$. Then

$$m = \gamma.M + \int_0^{\cdot} \mathrm{d}\gamma M_- = \gamma.M + \int_0^{\cdot} \mathrm{d}\gamma P_- M_-.$$

The last integral is easily seen to be zero. So m is equal to $\gamma.M$ and thus a martingale. Moreover we also obtain

$$\langle m, M\rangle = \gamma.\langle M\rangle = \gamma\langle M\rangle - \int_0^{\cdot} \mathrm{d}\gamma\langle M\rangle_-,$$

where again the last integral vanishes. But $\gamma\langle M\rangle = C\langle M\rangle$. Similarly $\langle m\rangle = C\langle M\rangle C^{\mathrm{T}}$. Hence $c(m, M) = 0$. Assume finally that m and M are mutually linearly dependent.

Then there exists matrices C_1 and C_2 as in the first part of the theorem. They are of the form as in the first part of the proof. Therefore we can compute

$$\begin{aligned} C_1C_2C_1 &= \lim_{t\to\infty} \langle m, M\rangle_t \langle M\rangle_t^+ \langle M, m\rangle_t \langle m\rangle^+ \langle m, M\rangle_t \langle M\rangle_t^+ \\ &= \lim_{t\to\infty} \langle m, M\rangle_t \langle M\rangle_t^+ \langle M\rangle_t \langle M\rangle_t^+ \\ &= \lim_{t\to\infty} \langle m, M\rangle_t^+ \langle M\rangle_t^+ = C_1 . \end{aligned}$$

Here we used in the second equality the fact that $c(M, m) = 0$. Similarly one can prove that $C_2C_1C_2 = C_2$ and $C_1C_2 = (C_1C_2)^{\mathrm{T}}$ which shows that C_1 and C_2 are each others Moore–Penrose inverses (cf. Lancaster and Tismenetsky, 1985). This completes the proof. □

Remark. Consider the other extreme case. One always has $c(m, M)_t \leq \langle m\rangle_t$. Here equality holds iff $\langle m, M\rangle_t = 0$. Indeed, assume that equality holds, then $\langle m, M\rangle_t \langle M\rangle_t = 0$, and hence $\langle m, M\rangle_t P_t = 0$ and by Lemma 2.4 this implies $\langle m, M\rangle_t = 0$. The converse statement is trivial.

By localization it is possible to formulate a whole string of corollaries, which are roughly all of the following type.

Corollary 4.3. *Let S be a stopping time and assume that*

$$c(m, M)_S 1_{\{S<\infty\}} + c(m, M)_{\infty-} 1_{\{S=\infty\}} = 0.$$

Then the stopped martingale m^S depends linearly on the stopped martingale M^S. Equivalently there exists C such that $1_{[\![0,S]\!]}(m - CM) = 0$.

Proof. It holds that $c(m, M)^S = c(m^S, M^S)$. Hence the assumption in the corollary implies $\lim_{t\to\infty} c(m^S, M^S)_t = 0$. So $c(m^S, M^S)_t = 0$ $\forall t \geq 0$, since $c(m^S, M^S)$ is non decreasing (Corollary 3.4). The result now follows from Theorem 4.2. □

Example 2. Let W be Brownian motion and ε an $N(0, 1)$ distributed random variable. Assume that W and ε are independent. Let $\mu_t = W_t + 1_{\{t\geq 1\}}\varepsilon$. Define $\xi : [0, \infty) \to \mathbb{R}^2$ by

$$\xi(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} 1_{\{1\}}(t) + \begin{bmatrix} 1 \\ t-1 \end{bmatrix} 1_{(1,\infty)}(t)$$

and $M = \xi.\mu$. Let

$$\mathcal{F}_t = \sigma\{W_s, s \leq t; 1_{\{t\geq 1\}}\varepsilon\}.$$

Then M is a martingale with respect to the filtration $F = \{\mathcal{F}_t\}_{t\geq 0}$ and

$$\langle M\rangle_t = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} 1_{\{1\}}(t) + \begin{bmatrix} t & \frac{1}{2}(t-1)^2 \\ \frac{1}{2}(t-1)^2 & \frac{1}{3}(t-1)^3 \end{bmatrix} 1_{(1,\infty)}(t)$$

for $\langle \mu\rangle = t + 1_{(1,\infty)}(t)$. Hence $r_t = \operatorname{rank}\langle M\rangle_t = 1_{\{1\}}(t) + 2.1_{(1,\infty)}(t)$.

Let $K:[0,\infty)\to\mathbb{R}^{2\times 2}$ be given by $K(t)=K^1 1_{\{1\}}(t)+K^2 1_{(1,\infty)}(t)$, and $m=K.M$. Then $\langle m,M\rangle_t=K.\langle M\rangle_t$ equals

$$K^1\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}1_{\{1\}}(t)+\left\{K^1\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}+K^2\begin{bmatrix}t-1 & \frac{1}{2}(t-1)^2\\ \frac{1}{2}(t-1)^2 & \frac{1}{3}(t-1)^3\end{bmatrix}\right\}1_{(1,\infty)}(t).$$

A computation shows

$$\langle M\rangle_t^+=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}1_{\{1\}}(t)+\begin{bmatrix}\dfrac{4}{t+3} & \dfrac{-6}{(t-1)(t+3)}\\ \dfrac{-6}{(t-1)(t+3)} & \dfrac{12t}{(t-1)^3(t+3)}\end{bmatrix}1_{(1,\infty)}(t).$$

Let $K^1=[K^1_{ij}]$ and $K^2=[K^2_{ij}]$. Then

$$\begin{aligned}\gamma_t&=\langle m,M\rangle_t\langle M\rangle_t^+\\ &=K^1\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}1_{\{1\}}(t)\\ &\quad+\begin{bmatrix}\dfrac{(4K^1_{11}+(t-1)K^2_{11})}{t+3} & K^2_{12}+\dfrac{6(K^2_{11}-K^1_{11})}{(t-1)(t+3)}\\ \dfrac{(4K^1_{21}+(t-1)K^2_{21})}{t+3} & K^2_{22}+\dfrac{6(K^2_{21}-K^1_{21})}{(t-1)(t+3)}\end{bmatrix}1_{(1,\infty)}(t).\end{aligned}$$

Hence $\lim_{t\downarrow 1}\gamma_t$ doesn't exist for arbitrary K.

Assume now that $c(m,M)=0$, then from Theorem 4.2 we know that γ is constant on $(1,\infty)$. So the following equalities have to hold: $K^1_{11}=K^2_{11}$ and $K^1_{21}=K^2_{21}$. Now γ becomes

$$\gamma_t=K^1\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}1_{\{1\}}(t)+K^2 1_{(1,\infty)}(t).$$

And in agreement with Theorem 4.2 (cf. its proof) we see that $m=\gamma_{1+}M$.

Example 3. Let ε_i be i.i.d. N(0, 1) random variables. Let $\mathcal{F}_n=\sigma\{\varepsilon_1,\dots,\varepsilon_n\}$. Let $x_1\cdots x_n$ be an orthonormal basis for $\mathbb{R}^n$ and $x_i=0$ for $i\geq n+1$. Let furthermore $K_i:\Omega\to\mathbb{R}^{k\times n}$ be $\mathcal{F}_{i-1}$ measurable. Define

$$M_t=\sum_{i\leq t}x_i\varepsilon_i,\qquad m_t=\sum_{i\leq t}K_i\Delta M_i.$$

Then

$$\langle M\rangle_t=\sum_{i\leq t}x_ix_i^{\mathrm{T}},\qquad \langle M\rangle_t^+=\sum_{i\leq t\wedge n}x_ix_i^{\mathrm{T}}.$$

A simple calculation shows that $c(m,M)=0$ and that the matrix C in Theorem 4.2 becomes $C=\sum_{i\leq n}K_ix_ix_i^{\mathrm{T}}$, which is $\mathcal{F}_{n-1}$ measurable.

Appendix

We provide a simple proof of the characterization of the Moore–Penrose inverse of a positive semidefinite matrix R. Write $R = QQ^{\mathrm{T}}$, where Q has full column rank, so that $Q^{\mathrm{T}}Q$ is invertible. Use the matrix inversion lemma, which is a simple extension of the Sherman–Morrison formula (cf. Lancaster and Tismenetsky, 1985, p. 64) to write $(R^2+(1/n)I)^{-1}R$ as

$$
\begin{aligned}
&= n[I - Q((1/n)(Q^{\mathrm{T}}Q)^{-1} + Q^{\mathrm{T}}Q)^{-1}Q^{\mathrm{T}}]QQ^{\mathrm{T}} \\
&= n[I - Q((1/n)(Q^{\mathrm{T}}Q)^{-2} + I)^{-1}(Q^{\mathrm{T}}Q)^{-1}Q^{\mathrm{T}}]QQ^{\mathrm{T}} \\
&= n[I - Q(I - (1/n)(Q^{\mathrm{T}}Q)^{-2} + \mathrm{O}(1/n^2))(Q^{\mathrm{T}}Q)^{-1}Q^{\mathrm{T}}]QQ^{\mathrm{T}} \\
&= Q(Q^{\mathrm{T}}Q)^{-2}Q^{\mathrm{T}} + \mathrm{O}(1/n).
\end{aligned}
$$

Clearly $Q(Q^{\mathrm{T}}Q)^{-2}Q^{\mathrm{T}}$ is the Moore–Penrose inverse of R.

References

P. Brémaud, Point Processes and Queues (Springer, New York, 1981).

K. Dzhaparidze and P.J.C. Spreij, On SLLN for multivariate martingales, Res. Mem. 1989-79, Vrije Univ. (Amsterdam, 1989).

K. Dzhaparidze and P.J.C. Spreij, On second order optimality of regular projective estimators, part I, Res. Mem. 1990-50, Vrije Univ. (Amsterdam, 1990).

P. Lancaster and M. Tismenetsky, The Theory of Matrices (Academic Press, Orlando, 1985).

R.S. Lipster and A.N. Shiryayev, Statistics of Random Processes II (Springer, New York, 1979).

R.S. Lipster and A.N. Shiryayev, Theory of Martingales (Kluwer, Dordrecht, 1990).

C.R. Rao and S.K. Mitra, Generalized Inverse of Matrices and its Applications (Wiley, New York, 1971).