A Minimality Property of the Minimal Martingale Measure

*Martin Schweizer**
Technische Universität Berlin
Fachbereich Mathematik, MA 7–4
Straße des 17. Juni 136
D – 10623 Berlin
Germany

Abstract: Let X be a continuous adapted process for which there exists an equivalent local martingale measure (ELMM). The minimal martingale measure $\widehat{P}$ is the unique ELMM for X with the property that local P-martingales strongly orthogonal to the P-martingale part of X are also local $\widehat{P}$-martingales. We prove that if $\widehat{P}$ exists, it minimizes the reverse relative entropy $H(P|Q)$ over all ELMMs Q for X. A counterexample shows that the assumption of continuity cannot be dropped.

Key words: minimal martingale measure, relative entropy, equivalent martingale measures

1991 Mathematics Subject Classification: 60G48, 90A09

JEL Classification Numbers: G10

(Statistics and Probability Letters 42 (1999), 27–31)

This version: 10.11.98

---

* Research for this paper was partially carried out within Sonderforschungsbereich 373. This paper was printed using funds made available by the Deutsche Forschungsgemeinschaft.

1. The result

In this section, we introduce the framework for our problem and present our main result. Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $I\!\!F = (\mathcal{F}_t)_{0 \le t \le T}$ satisfying the usual conditions of right-continuity and completeness, where $T \in (0, \infty]$ is a fixed time horizon. For all unexplained terminology from stochastic analysis, we refer to Protter (1990). We consider an $I\!\!R^d$-valued $I\!\!F$-adapted process $X = (X_t)_{0 \le t \le T}$ and assume that X has P-a.s. *continuous* trajectories. Intuitively, X represents the discounted price evolution of d risky assets in a financial market, and we want to exclude the possibility of having arbitrage ("money-pumps") in this market. We therefore assume that X admits an *equivalent local martingale measure (ELMM)*, i.e., there exists a probability measure $Q \approx P$ with $Q = P$ on $\mathcal{F}_0$ such that X is a local Q-martingale; see for instance Delbaen/Schachermayer (1994) for a more detailed discussion of the economic significance of such a condition. Together with the continuity of X, it implies by Theorem 2.2 of Choulli/Stricker (1996) that X is a special semimartingale satisfying the *structure condition (SC)*: In the canonical decomposition $X = X_0 + M + A$, the process M is an $I\!\!R^d$-valued locally square-integrable local P-martingale, and the $I\!\!R^d$-valued process A of finite variation has the form

$$A_t = \int_0^t d\langle M \rangle_s \, \lambda_s \qquad , \qquad 0 \le t \le T$$

for an $I\!\!R^d$-valued predictable process λ such that

$$K_t := \int_0^t \lambda_s^{\mathrm{tr}} \, d\langle M \rangle_s \, \lambda_s = \sum_{i,j=1}^d \int_0^t \lambda_s^i \lambda_s^j \, d\langle M^i, M^j \rangle_s < \infty \qquad P\text{-a.s. for all } t \in [0, T].$$

The process K is called the *mean-variance tradeoff process* of X.

Since X admits an least one ELMM, one can ask about ELMMs having some special properties. One possibility is the *minimal martingale measure* $\widehat{P}$ introduced by Föllmer/Schweizer (1991) and generalized by Ansel/Stricker (1992, 1993). This is defined by

$$\text{(1.1)} \qquad \frac{d\widehat{P}}{dP} := \widehat{Z}_T \qquad \text{with } \widehat{Z} := \mathcal{E}\left(-\int \lambda \, dM\right),$$

where we assume that the exponential local P-martingale $\widehat{Z}$ is strictly positive and a true P-martingale so that $E[\widehat{Z}_T] = 1$. If in addition $\widehat{Z}_T \in L^2(P)$, then Theorem (3.5) of Föllmer/Schweizer (1991) shows that every square-integrable P-martingale L strongly P-orthogonal to M is also a $\widehat{P}$-martingale (and strongly $\widehat{P}$-orthogonal to X). Thus $\widehat{P}$ is minimal in the sense that it preserves the martingale structure as far as possible under the constraint of turning X into a martingale. Moreover, $\widehat{P}$ is also the natural candidate for an ELMM for X by Girsanov's theorem.

Because the preceding description of minimality is somewhat awkward, there have been several attempts to characterize $\widehat{P}$ in a different way. An economic characterization in a multidimensional diffusion framework has been given in Hofmann/Platen/Schweizer (1992). Föllmer/Schweizer (1991) and Schweizer (1995a) have shown that for X continuous, $\widehat{P}$ minimizes the "free energy" $H(Q|P) - \frac{1}{2} E_Q[K_T]$ over all ELMMs Q for X satisfying $E_Q[K_T] < \infty$. Here we recall that for two probability measures P, Q and a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, the relative entropy of Q with respect to P on $\mathcal{G}$ is

$$H_{\mathcal{G}}(Q|P) := \begin{cases} E_Q\left[\log \left.\dfrac{dQ}{dP}\right|_{\mathcal{G}}\right] & \text{, if } Q \ll P \text{ on } \mathcal{G} \\ +\infty & \text{, otherwise.} \end{cases}$$

We also recall that $H_{\mathcal{G}}(Q|P)$ is always nonnegative, increasing in $\mathcal{G}$, and that $H(Q|P) := H_{\mathcal{F}}(Q|P)$ is 0 if and only if $Q = P$. In particular, the above characterization of $\widehat{P}$ implies that $\widehat{P}$ minimizes the relative entropy $H(Q|P)$ over all ELMMs Q for X if X is continuous and the final value K_T of the mean-variance tradeoff process is *deterministic*. Under the same conditions, $\widehat{P}$ also minimizes $\mathrm{Var}\left[\frac{dQ}{dP}\right]$ or $\left\|\frac{dQ}{dP}\right\|_{L^2(P)}$ over all ELMMs Q for X; see Theorem 7 of Schweizer (1995a). Miyahara (1996) has shown that $\widehat{P}$ also minimizes $H(Q|P)$ over all ELMMs Q if X is a Markovian diffusion given by the multidimensional stochastic differential equation

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dW_t.$$

But all these results either use a very specific structure for X or impose the very restrictive condition that K_T should be deterministic. In contrast, the main result of this paper is completely general.

Theorem 1. *Suppose that X is a continuous adapted process admitting at least one equivalent local martingale measure Q. If $\widehat{P}$ defined by (1.1) is a probability measure equivalent to P, then $\widehat{P}$ minimizes the reverse relative entropy $H(P|Q)$ over all ELMMs Q for X.*

We remark that the idea of considering $H(P|Q)$ instead of $H(Q|P)$ first appeared in Platen/Rebolledo (1996). The assumption about $\widehat{P}$ of course just states that the minimal martingale measure $\widehat{P}$ should exist; it is thus a minimal requirement for the theorem's assertion. Theorem 1 is only true for a *continuous* process X; we shall show by a counterexample in the next section that the conclusion fails in general if X has jumps.

The next result is a preparation for the proof of Theorem 1. It does not really need any martingale structure; we could replace N_τ by any positive random variable with expectation 1. The present formulation just makes clear how we apply the lemma later on.

Lemma 2. *Suppose that N is a strictly positive local P-martingale with $N_0 = 1$. For any stopping time τ such that the stopped process N^τ is a P-martingale, we then have* $E[\log N_\tau] \in [-\infty, 0]$.

Proof. We cannot use Jensen's inequality because we do not know whether $\log N_\tau$ is integrable. But since N^τ is a strictly positive P-martingale starting from 1, N_τ is strictly positive and has expectation 1. Thus we can define a probability measure $R \approx P$ by $\frac{dR}{dP} := N_\tau$, and so we obtain

$$E_P[-\log N_\tau] = E_P\left[\log \frac{dP}{dR}\right] = H(P|R) \in [0, \infty].$$

q.e.d.

Proof of Theorem 1: Let Q be any ELMM for X and denote by Z its density process with respect to P. We may also assume that $H(P|Q) < \infty$ since there is nothing to prove otherwise. Because X is continuous, we can write Z as $Z = \widehat{Z}\mathcal{E}(L)$ for a local P-martingale L with $L_0 = 0$; see Theorem 1 of Schweizer (1995a) or Corollary 2.3 of Choulli/Stricker (1996). Let $(\tau_n)_{n \in I\!N}$ be a localizing sequence for $\mathcal{E}(L)$ and $\int \lambda \, dM$ and fix $n \in I\!N$. Then

$$\left.\frac{dP}{dQ}\right|_{\mathcal{F}_{\tau_n}} = \frac{1}{Z_{\tau_n}} = \frac{1}{\widehat{Z}_{\tau_n}} \frac{1}{\mathcal{E}(L)_{\tau_n}} = \left.\frac{dP}{d\widehat{P}}\right|_{\mathcal{F}_{\tau_n}} \frac{1}{\mathcal{E}(L)_{\tau_n}},$$

and so Lemma 2 with $N := \mathcal{E}(L)$ implies that

$$H_{\mathcal{F}_{\tau_n}}(P|Q) = H_{\mathcal{F}_{\tau_n}}(P|\widehat{P}) - E_P\left[\log \mathcal{E}(L)_{\tau_n}\right] \geq H_{\mathcal{F}_{\tau_n}}(P|\widehat{P})$$

and therefore

$$\sup_{n \in I\!N} H_{\mathcal{F}_{\tau_n}}(P|\widehat{P}) \leq \sup_{n \in I\!N} H_{\mathcal{F}_{\tau_n}}(P|Q) \leq H(P|Q) < \infty, \tag{1.2}$$

since $H_{\mathcal{G}}(P|Q)$ is increasing in $\mathcal{G}$. From Lemma 2 of Barron (1985), we thus obtain

$$\sup_{n \in I\!N} \left|\log \frac{1}{\widehat{Z}_{\tau_n}}\right| = \sup_{n \in I\!N} \left|\log \left.\frac{dP}{d\widehat{P}}\right|_{\mathcal{F}_{\tau_n}}\right| \in L^1(P),$$

and since $\widehat{Z}_{\tau_n} \to \widehat{Z}_T$ P-a.s. because τ_n increases stationarily to T, the dominated convergence theorem yields

$$H(P|\widehat{P}) = E_P\left[\log \frac{1}{\widehat{Z}_T}\right] = \lim_{n \to \infty} E_P\left[\log \frac{1}{\widehat{Z}_{\tau_n}}\right] = \lim_{n \to \infty} H_{\mathcal{F}_{\tau_n}}(P|\widehat{P}) \leq H(P|Q)$$

by (1.2). As Q was arbitrary, the proof is complete.

q.e.d.

Remark. A closer look at the above proof shows that we only need continuity of X to write the density process Z of an arbitrary ELMM as $Z = \widehat{Z}\mathcal{E}(L)$ for some local P-martingale L null at 0. One can ask if this is also possible for a general semimartingale X satisfying the structure condition (SC), but the answer is negative. An explicit counterexample can be obtained by taking for X the sum of a Brownian motion with drift and a compensated Poisson process. Alternatively, this is a consequence of the counterexample in the next section.

2. The counterexample

If the process X is not continuous, the assertion of Theorem 1 is no longer true: We present here a counterexample with an ELMM Q^* such that $H(P|Q^*) < H(P|\widehat{P})$. It uses a bounded process in finite discrete time and basically consists of a number of elementary computations.

Fix some $U > 1$ and consider for X a trinomial tree with time horizon 2 and parameters $U, 1, \frac{1}{U}$. Formally, let Y_1, Y_2 be i.i.d. under P taking the values $U, 1, \frac{1}{U}$ with probability $\frac{1}{3}$ each. The process $X = (X_k)_{k=0,1,2}$ is then given by $X_0 := 1$, $X_1 := Y_1$ and $X_2 := Y_1 Y_2$, and $I\!\!F$ is the filtration generated by X. We use the notation $\Delta X_k := X_k - X_{k-1}$ for the increments of X.

Any equivalent martingale measure (EMM) Q for X can be identified with a vector $q \in (0,1)^4$ via its transition probabilities

$$q_1 := Q[X_1 = U] \quad , \quad q_2 := Q[X_2 = U | X_1 = U]$$
$$q_3 := Q[X_2 = U | X_1 = 1] \quad , \quad q_4 := Q\left[X_2 = U \middle| X_1 = \frac{1}{U}\right].$$

The other transition probabilities are then determined by the martingale property of X under Q and the fact that they add to 1 at each node in the tree. An elementary computation yields

$$\begin{aligned}(2.1)\quad H(P|Q) &= E_P\left[-\log \frac{dQ}{dP}\right] \\ &= -\frac{2}{3}\log q_1 - \frac{1}{3}\log\big(1-(U+1)q_1\big) - \frac{1}{9}\sum_{i=2}^{4}\Big(2\log q_i + \log\big(1-(U+1)q_i\big)\Big) \\ &\quad + \log 9 - \frac{2}{3}\log U,\end{aligned}$$

and setting the gradient with respect to q equal to 0 gives an EMM Q^* with

$$q_i^* = \frac{2}{3(U+1)} \qquad \text{for } i = 1, \dots, 4$$

as a candidate for the entropy-optimal EMM. Under Q^*, the random variables Y_1, Y_2 are still i.i.d. and take the values $U, 1, \frac{1}{U}$ with probability $\frac{2}{3(U+1)}$, $\frac{1}{3}$ and $\frac{2U}{3(U+1)}$, respectively, so that

Q^* is clearly equivalent to P. Inserting into (2.1) yields after some simplification

$$H(P|Q^*) = \log \frac{81}{\sqrt[3]{16}} + \frac{2}{3} \log \frac{(U+1)^2}{U}.$$

To compute the minimal EMM $\widehat{P}$ for X, we use the results of Schweizer (1995b). According to equations (2.21) and (1.2) in that paper, $\widehat{P}$ is given by the density

$$\frac{d\widehat{P}}{dP} = \widehat{Z}_2 = \prod_{k=1}^{2} \frac{1-\alpha_k \Delta X_k}{1-\alpha_k \Delta A_k} = \prod_{k=1}^{2} \frac{E\left[\Delta X_k^2 \middle| \mathcal{F}_{k-1}\right] - \Delta X_k E[\Delta X_k|\mathcal{F}_{k-1}]}{E\left[\Delta X_k^2 \middle| \mathcal{F}_{k-1}\right] - (E[\Delta X_k|\mathcal{F}_{k-1}])^2}.$$

Computing this explicitly shows that $\widehat{P}$ can be identified with the vector $\widehat{q}$ given by

$$\widehat{q}_i = \frac{U+1}{2(U^2+U+1)} \qquad \text{for } i = 1, \dots, 4.$$

This means that under $\widehat{P}$, Y_1 and Y_2 are again i.i.d. and take the values $U, 1, \frac{1}{U}$ with probability $\frac{U+1}{2(U^2+U+1)}$, $\frac{U^2+1}{2(U^2+U+1)}$ and $\frac{U^2+U}{2(U^2+U+1)}$, respectively. Inserting into (2.1) now yields

$$H(P|\widehat{P}) = \log 36 - \frac{2}{3} \log \frac{U(U^2+1)(U+1)^2}{(U^2+U+1)^3}.$$

If we take for instance $U = 2$, we obtain

$$q_i^* = \frac{2}{9} \quad , \quad \widehat{q}_i = \frac{3}{14} \qquad \text{for } i = 1, \dots, 4$$

and

$$H(P|Q^*) = 4.473 < 4.475 = H(P|\widehat{P}).$$

This shows that $\widehat{P}$ need not minimize the reverse relative entropy if X is not continuous so that we have indeed a counterexample. Numerical evidence suggests that $H(P|Q^*) < H(P|\widehat{P})$ for every $U > 1$, but we have not bothered to check this theoretically.

References

J. P. Ansel and C. Stricker (1992), "Lois de Martingale, Densités et Décomposition de Föllmer-Schweizer", *Annales de l'Institut Henri Poincaré 28 (1992)*, 375–392

J. P. Ansel and C. Stricker (1993), "Unicité et Existence de la Loi Minimale", *Séminaire de Probabilités XXVII, Lecture Notes in Mathematics 1557, Springer*, 22–29

A. R. Barron (1985), “The Strong Ergodic Theorem for Densities: Generalized Shannon-McMillan-Breiman Theorem”, *Annals of Probability 13*, 1292–1303

T. Choulli and C. Stricker (1996), “Deux Applications de la Décomposition de Galtchouk-Kunita-Watanabe”, *Séminaire de Probabilités XXX, Lecture Notes in Mathematics 1626, Springer*, 12–23

F. Delbaen and W. Schachermayer (1994), “A General Version of the Fundamental Theorem of Asset Pricing”, *Mathematische Annalen 300*, 463–520

H. Föllmer and M. Schweizer (1991), “Hedging of Contingent Claims under Incomplete Information”, *in: M. H. A. Davis and R. J. Elliott (eds.), “Applied Stochastic Analysis”, Stochastics Monographs, Vol. 5, Gordon and Breach, London/New York*, 389–414

N. Hofmann, E. Platen and M. Schweizer (1992), “Option Pricing under Incompleteness and Stochastic Volatility”, *Mathematical Finance 2*, 153–187

Y. Miyahara (1996), “Canonical Martingale Measures of Incomplete Assets Markets”, *in: Probability Theory and Mathematical Statistics, Proceedings of the Seventh Japan-Russia Symposium, Tokyo 1995*, 343–352

E. Platen and R. Rebolledo (1996), “Principles for Modelling Financial Markets”, *Journal of Applied Probability 33*, 601–613

P. Protter (1990), “Stochastic Integration and Differential Equations. A New Approach”, *Springer*

M. Schweizer (1995a), “On the Minimal Martingale Measure and the Föllmer-Schweizer Decomposition”, *Stochastic Analysis and Applications 13*, 573–599

M. Schweizer (1995b), “Variance-Optimal Hedging in Discrete Time”, *Mathematics of Operations Research 20*, 1–32