ON MINIMIZATION AND MAXIMIZATION OF ENTROPY IN VARIOUS DISCIPLINES

A.S. Cherny*, V.P. Maslov**

* *Moscow State University,*
Faculty of Mechanics and Mathematics,
Department of Probability Theory,
119992 Moscow, Russia.
E-mail: cherny@mech.math.msu.su

** *Moscow State University,*
Faculty of Physics,
Department of Quantum Statistics and Field Theory,
119992 Moscow, Russia.
E-mail: viktor_maslov@hotmail.com

Abstract. This paper deals with some problems related to the relative entropy minimization under linear constraints. We discuss the relation between this problem and statistical physics, information theory and financial mathematics. Furthermore, in financial mathematics we provide the explicit form of the minimal entropy martingale measure in the general discrete-time asset price model. We also give the explicit solution of the problem of the exponential utility maximization in the general discrete-time asset price model.

Key words and phrases. Amount of information, average cost of coding, data compression, density, entropy, Esscher transform, exponential utility, free energy, Gibbs state, interior energy, Kullback-Leibler information, mass, metastable state, minimal entropy martingale measure, Nernst theorem, pressure, relative entropy, stable state, temperature, volume.

1 Introduction

1. Probability theory: the minimum and the critical points of the relative entropy. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a measurable space with a σ-finite measure P. Let $X : \Omega \to \mathbb{R}^d$ be a measurable function. We will consider the problem of minimizing the *relative entropy* (also known as the *Kullback-Leibler information*)

$$H(\mathsf{Q}, \mathsf{P}) = \begin{cases} \displaystyle\int_\Omega \ln \frac{d\mathsf{Q}}{d\mathsf{P}}\, d\mathsf{Q} & \text{if } \displaystyle\int_\Omega \Big|\ln \frac{d\mathsf{Q}}{d\mathsf{P}}\Big| d\mathsf{Q} < \infty, \\ +\infty & \text{if } \displaystyle\int_\Omega \Big|\ln \frac{d\mathsf{Q}}{d\mathsf{P}}\Big| d\mathsf{Q} = \infty \end{cases} \tag{1.1}$$

over the set

$$\mathcal{E} = \Big\{\mathsf{Q} : \mathsf{Q} \text{ is a probability measure on } (\Omega, \mathcal{F}), \quad \mathsf{Q} \ll \mathsf{P},\ \int_\Omega \|X\| d\mathsf{Q} < \infty,\ \int_\Omega X d\mathsf{Q} = 0\Big\}. \tag{1.2}$$

This problem (and even more general problem, where we have infinitely many constraints of the form $\int_\Omega X^i d\mathsf{Q} = 0$) is well known in probability theory; see Bucklew [1; p. 30], Cover and Thomas [2; Ch. 11], Csiszar [3], Föllmer and Schied [6; §3.2], Frittelli [7]. Theorem 2.2 that we prove below can be derived from [3; Theorem 3.1] in the case, where P is a finite measure. However, we prefer to give a direct proof for two reasons: first, the case, where $\mathsf{P}(\Omega) = \infty$, does not follow directly from [3] (we need to consider this case for the applications in statistical physics and information theory); second, the proof in [3] does not employ the method of Lagrange multipliers, while we prefer to use this method in order to stress the relationship between the problem under consideration and finding the stable state of a system in statistical physics.

The measure, at which the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ attains its minimum, is its critical point. It is also interesting to find other critical points. In Theorem 2.6 we describe all the critical points of this functional on the set $\mathcal{E}$.

The problem of the relative entropy minimization under linear constraints arises in various disciplines:

- theory of large deviations,
- statistics,
- statistical physics,
- information theory,
- financial mathematics.

In the theory of large deviations this problem arises in the Sanov theorem and the conditional limit theorem (see Cover and Thomas [2; Ch. 12]). In statistics it arises in the proof of Stein's lemma and in the proof of Chernoff's theorem (see Cover and Thomas [2; Ch. 12]). The relationship with the other three disciplines is described below.

2. Statistical physics: stable and metastable states. In statistical physics the problem of the relative entropy minimization under linear constraints corresponds to maximizing the (Boltzmann) *entropy* of a system under the fixed *interior energy* (the entropy in physics is taken with the opposite sign as compared to the relative entropy in probability theory), i.e. determining the *stable state* of a system; see Landau and Lifshits [11] (also see Jaynes [10]). In Section 3, we translate the results of Section 2 into the physical language and thus obtain formulas for the *Gibbs state*, the *free energy*, the *interior energy*, and the *entropy* of a system in a stable state as well as basic thermodynamic relations between these quantities (see Theorem 3.1).

The problem of finding the critical points of the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ over $\mathcal{E}$ means determining the *metastable states* of a system; see Maslov [12].

3. Information theory: data compression. In information theory the problem of the relative entropy minimization under linear constraints corresponds to maximizing the *amount of information* contained in a coding under the fixed *average cost of the coding*. This problem was considered by Stratonovich [17] under the name "first variational problem". Its solution shows, in particular, that the maximal amount of information, which can be contained in the words of the D-ary alphabet with the average length L, is approximately $L \ln D$ for the large L. This result is often derived from completely different considerations (McMillan inequality); see Cover and Thomas [2; Ch. 5]. The interpretation in information theory of the problem being considered is described in Section 4.

4. Financial mathematics: minimal entropy martingale measure. In financial mathematics the problem of the relative entropy minimization under linear constraints corresponds to finding the *minimal entropy martingale measure* in the one-period asset price model.

In the case, where the random variables describing the asset prices have exponential moments, the solution is well known (see, for instance, Föllmer and Schied [6; Corollary 3.27]). Namely, the minimal entropy martingale measure is the *Esscher transform* of the original probability measure P.

In the multiperiod discrete-time models and the continuous-time models the form of the minimal entropy martingale measure is known in some particular cases; see Fujiwara and Miyahara [8], Miyahara and Novikov [13]. In Section 5, we give the explicit form of the minimal entropy martingale measure in the general discrete-time model satisfying natural integrability conditions (see Theorem 5.6).

The problem of finding the minimal entropy martingale measure is dual to the problem of the *exponential utility* maximization as pointed out in many papers; see Delbaen et al. [4], Frittelli [7], Goll and Rüschendorf [9], Schachermayer [15]. We also provide in Section 5 an explicit solution of the exponential utility maximization problem in the general discrete-time model satisfying natural integrability conditions (see Theorem 5.10). Then the duality of the two problems becomes obvious from their solutions.

2 Probability Theory: The Minimum and the Critical Points of the Relative Entropy

1. Relative entropy minimization under linear constraints. Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a measurable space with a σ-finite measure P. Let $X : \Omega \to \mathbb{R}^d$ be a measurable function. Let us find the minimal value of the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ defined in (1.1) over the set $\mathcal{E}$ defined in (1.2). We will apply the method of Lagrange multipliers and first find the minimum of the expression $\int_\Omega \langle \tau, X \rangle d\mathsf{Q} + H(\mathsf{Q}, \mathsf{P})$ over all the probability measures $\mathsf{Q} \ll \mathsf{P}$ (by $\langle \cdot, \cdot \rangle$ we denote the scalar product in $\mathbb{R}^d$).

Lemma 2.1. *Let $Z : \Omega \to \mathbb{R}$ be a random variable. Suppose that $\int_\Omega e^Z d\mathsf{P} < \infty$ and $\int_\Omega |Z| e^Z d\mathsf{P} < \infty$. Then the minimum of the expression $-\int_\Omega Z d\mathsf{Q} + H(\mathsf{Q}, \mathsf{P})$ over the probability measures $\mathsf{Q} \ll \mathsf{P}$ with $\int_\Omega |Z| d\mathsf{Q} < \infty$ is attained at the unique measure*

$$\mathsf{Q}_* = const\, e^Z \mathsf{P}.$$

`Proof`. For any measure Q with the described properties, we have

$$\begin{aligned} H(\mathsf{Q}, \mathsf{Q}_*) &= \int_\Omega \ln \frac{d\mathsf{Q}}{d\mathsf{Q}_*}\, d\mathsf{Q} = \int_\Omega q \ln\Big(q e^{-Z} \int_\Omega e^Z d\mathsf{P}\Big) d\mathsf{P} \\ &= -\int_\Omega Z q d\mathsf{P} + \int_\Omega q \ln q\, d\mathsf{P} + \ln \int_\Omega e^Z d\mathsf{P} \\ &= -\int_\Omega Z d\mathsf{Q} + H(\mathsf{Q}, \mathsf{P}) + \ln \int_\Omega e^Z d\mathsf{P}, \end{aligned}$$

where $q = \frac{d\mathsf{Q}}{d\mathsf{P}}$. This yields the desired statement. □

Set $S = \operatorname{supp}(\mathsf{P} \circ X^{-1})$. Let L denote the smallest affine subspace of $\mathbb{R}^d$ containing S. By $\overset{\circ}{S}$ we denote the relative interior of S, i.e. the interior of S in the relative topology of L. If $0 \in \overset{\circ}{S}$, then $\mathcal{E} \neq \emptyset$ (see Shiryaev [16; Ch. V, §2e]).

Consider the function

$$\varphi(\tau) = \ln \int_\Omega e^{\langle \tau, X \rangle} d\mathsf{P}, \quad \tau \in \mathbb{R}^d,$$

which takes on values in $(-\infty, \infty]$. It follows from the Jensen inequality that φ is convex.

We will also consider the function

$$\psi(\tau) = \frac{\int_\Omega X e^{\langle \tau, X\rangle} d\mathsf{P}}{\int_\Omega e^{\langle \tau, X\rangle} d\mathsf{P}} \tag{2.1}$$

defined on the set

$$\Big\{\tau \in \mathbb{R}^d : \int_\Omega \|X\| e^{\langle \tau, X\rangle} d\mathsf{P} < \infty,\ \int_\Omega e^{\langle \tau, X\rangle} d\mathsf{P} < \infty\Big\}.$$

Note that, for τ from the interior of the set $\{\tau \in \mathbb{R}^d : \varphi(\tau) < \infty\}$, we have $\psi(\tau) = \operatorname{grad}\varphi(\tau)$.

Theorem 2.2. *Suppose that* $0 \in \overset{\circ}{S}$.
(i) *We have*

$$\inf\{H(\mathsf{Q}, \mathsf{P}) : \mathsf{Q} \in \mathcal{E}\} = -\inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}.$$

(ii) *If there exists* $\tau_* \in \mathbb{R}^d$ *such that* $\varphi(\tau_*) = \inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}$ *and* $\psi(\tau_*) = 0$, *then the minimum of* $H(\mathsf{Q}, \mathsf{P})$ *over* $\mathcal{E}$ *is attained at the unique measure*

$$\mathsf{Q}_* = const\, e^{\langle \tau_*, X\rangle}\, \mathsf{P}.$$

Otherwise, the minimum of $H(\mathsf{Q}, \mathsf{P})$ *over* $\mathcal{E}$ *is not attained.*

Remark. We have $\varphi(\tau) = \varphi(\operatorname{pr}_L \tau)$, $\psi(\tau) = \psi(\operatorname{pr}_L \tau)$, where pr_L denotes the orthogonal projection on L (the condition $0 \in \overset{\circ}{S}$ guarantees that L is not only an affine subspace, but is also a linear subspace). The function φ is strictly convex on L. Hence, the set of points τ_* satisfying the conditions $\varphi(\tau_*) = \inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}$ and $\psi(\tau_*) = 0$ is either empty or has the form $\tau_0^* + L^\perp$, where $L^\perp$ is the orthogonal complement to L. Thus, τ_* is not determined uniquely, while Q_* is nevertheless unique. □

Proof. **(i)** Let us first prove that

$$\inf\{H(\mathsf{Q}, \mathsf{P}) : \mathsf{Q} \in \mathcal{E}\} \le -\inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}. \tag{2.2}$$

Consider sets $A_n \in \mathcal{F}$ such that $A_n \subseteq A_{n+1}$, $0 < \mathsf{P}(A_n) < \infty$, and $\bigcup A_n = \Omega$. Define $B_n = A_n \cap \{\|X\| \le n\}$, $\mathsf{P}_n = \mathsf{P}(\,\cdot \cap B_n)$, and

$$\varphi_n(\tau) = \ln \int_\Omega e^{\langle \tau, X\rangle} d\mathsf{P}_n, \quad \tau \in \mathbb{R}^d. \tag{2.3}$$

We have $\varphi_n \le \varphi_{n+1}$, and the functions φ_n tend to φ pointwise. Furthermore, $\varphi(\tau) = \varphi(\operatorname{pr}_L \tau)$, $\varphi_n(\tau) = \varphi_n(\operatorname{pr}_L \tau)$, the function φ is strictly convex on L, and $\varphi(\tau) \to \infty$ on L as $\|\tau\| \to \infty$. Consequently, for any sufficiently large n, there exists a point $\tau_n^* \in \mathbb{R}^d$, at which φ_n attains its minimum. For the measure $\mathsf{Q}_n^* = const\, e^{\langle \tau_n^*, X\rangle} \mathsf{P}_n$, we have

$$\int_\Omega X d\mathsf{Q}_n^* = \frac{\int_\Omega X e^{\langle \tau_n^*, X\rangle} d\mathsf{P}_n}{\int_\Omega e^{\langle \tau_n^*, X\rangle} d\mathsf{P}_n} = \operatorname{grad}\varphi_n(\tau_n^*) = 0.$$

Hence, $\mathsf{Q}_n^* \in \mathcal{E}$. Moreover,

$$\begin{aligned} H(\mathsf{Q}_n^*, \mathsf{P}) &= \int_\Omega \ln \frac{d\mathsf{Q}_n^*}{d\mathsf{P}}\, d\mathsf{Q}_n^* = \int_\Omega \ln \frac{d\mathsf{Q}_n^*}{d\mathsf{P}_n}\, d\mathsf{Q}_n^* \\ &= \int_\Omega \langle \tau_n^*, X\rangle d\mathsf{Q}_n^* - \ln \int_\Omega e^{\langle \tau_n^*, X\rangle} d\mathsf{P}_n \\ &= -\varphi_n(\tau_n^*) \xrightarrow[n\to\infty]{} -\inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}. \end{aligned}$$

This proves (2.2).

Let us now prove that

$$\inf\{H(\mathsf{Q},\mathsf{P}) : \mathsf{Q} \in \mathcal{E}\} \geq -\inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}. \tag{2.4}$$

Assume that this is not true, i.e. there exists a measure $\mathsf{Q}_0 \in \mathcal{E}$ and a point $\tau_0 \in \mathbb{R}^d$ such that $H(\mathsf{Q}_0,\mathsf{P}) < -\varphi(\tau_0)$. Consider sets $A_n \in \mathcal{F}$ such that $A_n \subseteq A_{n+1}$, $0 < \mathsf{P}(A_n) < \infty$, and $\bigcup A_n = \Omega$. Take $B_n = A_n \cap \{\|X\| \leq n\}$, $\mathsf{P}_n = \mathsf{P}(\,\cdot \cap B_n)$, $\mathsf{Q}_n = \mathsf{Q}_0(\,\cdot \mid B_n)$ and define φ_n by (2.3). Since

$$\int_\Omega X d\mathsf{Q}_n \xrightarrow[n\to\infty]{} \int_\Omega X d\mathsf{Q}_0 = 0,$$

$$H(\mathsf{Q}_n,\mathsf{P}_n) \xrightarrow[n\to\infty]{} H(\mathsf{Q}_0,\mathsf{P}),$$

$$\varphi_n(\tau_0) \xrightarrow[n\to\infty]{} \varphi(\tau_0),$$

there exists n such that

$$-\int_\Omega \langle \tau_0, X\rangle d\mathsf{Q}_n + H(\mathsf{Q}_n,\mathsf{P}_n) < -\varphi_n(\tau_0).$$

On the other hand, it follows from Lemma 2.1 that

$$\begin{aligned}
&-\int_\Omega \langle \tau_0, X\rangle d\mathsf{Q}_n + H(\mathsf{Q}_n,\mathsf{P}_n)\\
&\geq -\frac{\int_\Omega \langle \tau_0, X\rangle e^{\langle \tau_0, X\rangle} d\mathsf{P}_n}{\int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}_n} + \int_\Omega \frac{e^{\langle \tau_0, X\rangle}}{\int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}_n} \ln \frac{e^{\langle \tau_0, X\rangle}}{\int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}_n} d\mathsf{P}_n\\
&= -\ln \int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}_n = -\varphi_n(\tau_0).
\end{aligned}$$

The obtained contradiction shows that (2.4) is true.

(ii) The first statement is a straightforward consequence of Lemma 2.1.

Suppose now that there exists no point $\tau_* \in \mathbb{R}^d$ such that $\varphi(\tau_*) = \inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}$ and $\psi(\tau_*) = 0$. Assume that there exists a measure $\mathsf{Q}_0 \in \mathcal{E}$, at which the minimum of $H(\mathsf{Q},\mathsf{P})$ over $\mathcal{E}$ is attained. Denote $\frac{d\mathsf{Q}_0}{d\mathsf{P}}$ by q_0. Let us first prove that there exist $\tau_0 \in \mathbb{R}^d$ and $c_0 \in \mathbb{R}$ such that $\ln q_0 = \langle \tau_0, X\rangle + c_0$ P-a.e. on $\{q_0 > 0\}$. If this is not true, then there exists a set $B \in \mathcal{F}$ such that $0 < \mathsf{P}(B) < \infty$, X is bounded on B, $\ln q_0$ is bounded on B, and the restriction $\ln q_0|_B$ does not belong to the space $\mathcal{L} = \{\langle \tau, X\rangle|_B + c : \tau \in \mathbb{R}^d, c \in \mathbb{R}\}$. Let p denote the $L^2(B, \mathcal{F}|_B, \mathsf{P}|_B)$-projection of $\ln q_0|_B$ onto $\mathcal{L}$. Then the function $r = \ln q_0|_B - p$ is measurable, bounded, and it satisfies

$$\int_B |r| d\mathsf{P} < \infty, \quad \int_B r d\mathsf{P} = 0, \quad \int_B |r| X d\mathsf{P} < \infty, \quad \int_B r X d\mathsf{P} = 0. \tag{2.5}$$

Since $\ln q_0$ is bounded on B, the measure $\mathsf{Q}_0 + \varepsilon r \mathsf{P}$ is a probability measure for ε from a sufficiently small neighborhood of zero. It follows from (2.5) that $\mathsf{Q}_0 + \varepsilon r \mathsf{P} \in \mathcal{E}$. Furthermore,

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} H(\mathsf{Q}_0 + \varepsilon r \mathsf{P}, \mathsf{P}) = \int_\Omega r(1 + \ln q_0) d\mathsf{P} = \int_B r \ln q_0 d\mathsf{P} \neq 0.$$

This contradicts the choice of Q_0.

Thus, we have proved that Q_0 has the form $\mathsf{Q}_0 = \mathit{const}\, e^{\langle \tau_0, X\rangle} I_A \mathsf{P}$ with some $\tau_0 \in \mathbb{R}^d$, $A \in \mathcal{F}$. Let us now prove that $\mathsf{P}(\Omega \setminus A) = 0$. If this is not true, then we can find a set $B \in \mathcal{F}$

such that $0 < \mathsf{P}(B) < \infty$, X is bounded on B, and $\mathsf{P}(B \setminus A) > 0$. It follows from the first part of (ii) that there exists $\tau_1 \in \mathbb{R}^d$ such that the measure $\mathsf{Q}_1 = const\, e^{\langle \tau_1, X\rangle} I_B \mathsf{P}$ belongs to $\mathcal{E}$. Then $(1-\varepsilon)\mathsf{Q}_0 + \varepsilon \mathsf{Q}_1 \in \mathcal{E}$ for any $\varepsilon \in [0,1]$. We have

$$\begin{aligned} &H((1-\varepsilon)\mathsf{Q}_0 + \varepsilon\mathsf{Q}_1, \mathsf{P}) \\ &= \int_A ((1-\varepsilon)q_0 + \varepsilon q_1)\ln((1-\varepsilon)q_0 + \varepsilon q_1) d\mathsf{P} + \int_{B\setminus A} \varepsilon q_1 \ln \varepsilon q_1 d\mathsf{P}, \end{aligned}$$

where $q_1 = \frac{d\mathsf{Q}_1}{d\mathsf{P}}$. It is clear that there exists a sufficiently small $\varepsilon > 0$ such that $H((1-\varepsilon)\mathsf{Q}_0 + \varepsilon\mathsf{Q}_1, \mathsf{P}) < H(\mathsf{Q}_0, \mathsf{P})$. This contradicts the choice of Q_0.

Thus, we have proved that Q_0 has the form $\mathsf{Q}_0 = const\, e^{\langle \tau_0, X\rangle}\mathsf{P}$ with some $\tau_0 \in \mathbb{R}^d$. Then

$$\varphi(\tau_0) = -\ln \int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P} = -H(\mathsf{Q}_0, \mathsf{P}), \tag{2.6}$$

$$\psi(\tau_0) = \frac{\int_\Omega X e^{\langle \tau_0, X\rangle} d\mathsf{P}}{\int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}} = \int_\Omega X d\mathsf{Q}_0 = 0. \tag{2.7}$$

If there exists a point $\tau_1 \in \mathbb{R}^d$ such that $\varphi(\tau_1) < \varphi(\tau_0)$, then

$$\frac{\int_\Omega \langle \tau_1 - \tau_0, X\rangle e^{\langle \tau_0, X\rangle} d\mathsf{P}}{\int_\Omega e^{\langle \tau_0, X\rangle} d\mathsf{P}} = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \varphi((1-\varepsilon)\tau_0 + \varepsilon\tau_1) < 0,$$

which contradicts (2.7). Consequently, $\varphi(\tau_0) = \inf\{\varphi(\tau) : \tau \in \mathbb{R}^d\}$. The contradiction obtained shows that the measure Q_0 does not exist. □

Corollary 2.3. *Let $m \in \mathbb{R}$, $\sigma \neq 0$. The minimum of $\int_\mathbb{R} q(x)\ln q(x) dx$ over the set*

$$\left\{ q \geq 0 : \int_\mathbb{R} q(x)dx = 1,\ \int_\mathbb{R} xq(x)dx = m,\ \int_\mathbb{R} (x-m)^2 q(x)dx = \sigma^2 \right\}$$

is attained (*only*) *at the density of the normal distribution with the mean m and the variance σ^2.*

Corollary 2.4. *Let $m > 0$. The minimum of $\int_{\mathbb{R}_+} q(x)\ln q(x)dx$ over the set*

$$\left\{ q \geq 0 : \int_{\mathbb{R}_+} q(x)dx = 1,\ \int_{\mathbb{R}_+} xq(x)dx = m \right\}$$

is attained (*only*) *at the density of the exponential distribution with the mean m.*

2. Critical points of the relative entropy. We have found the measure, at which the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ attains its minimum over $\mathcal{E}$. Let us now find the critical points of this functional. For $\mathsf{Q} \in \mathcal{E}$, let $\mathcal{D}_\mathsf{Q}$ denote the set that consists of measurable bounded functions $r : \Omega \to \mathbb{R}$ with the following property: there exists $\delta > 0$ such that $\mathsf{Q} + r\varepsilon\mathsf{P} \in \mathcal{E}$ for any $\varepsilon \in [0, \delta)$.

Definition 2.5. An element $\mathsf{Q} \in \mathcal{E}$ is a *critical point* of the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ if for any $r \in \mathcal{D}_\mathsf{Q}$, the derivative $\frac{d}{d\varepsilon}\big|_{\varepsilon=0} H(\mathsf{Q} + \varepsilon r\mathsf{P}, \mathsf{P})$ does not exist or equals zero.

Theorem 2.6. *An element $\mathsf{Q} \in \mathcal{E}$ is a critical point of the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ if and only if Q has the form*

$$\mathsf{Q} = const\, e^{\langle \tau, X\rangle} I_A \mathsf{P}$$

with some $\tau \in \mathbb{R}^d$ and $A \in \mathcal{F}$.

Proof. Suppose that Q has the described form. Then, for any $r \in \mathcal{D}_{\mathsf{Q}}$,

$$H(\mathsf{Q} + \varepsilon r\mathsf{P}, \mathsf{P}) = \int_\Omega (q + \varepsilon r) \ln(q + \varepsilon r) d\mathsf{P},$$

where $q = \frac{d\mathsf{Q}}{d\mathsf{P}}$. If $\mathsf{P}(\{r > 0, q = 0\}) = 0$, then

$$\begin{aligned}\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} H(\mathsf{Q} + \varepsilon r\mathsf{P}, \mathsf{P}) &= \int_\Omega r(1 + \ln q) d\mathsf{P} = \int_A r \ln q \, d\mathsf{P} \\ &= \int_A r\langle \tau, X\rangle d\mathsf{P} - \int_A r d\mathsf{P} \, \ln \int_A e^{\langle \tau, X\rangle} d\mathsf{P} = 0.\end{aligned}$$

If $\mathsf{P}(\{r > 0, q = 0\}) > 0$, then this derivative does not exist.

The reverse implication has been verified in the proof of Theorem 2.2 (ii). □

3 Statistical Physics: Stable and Metastable States

1. Stable states. Let us consider a system that consists of many particles. Each of the particles can be in one of the possible *elementary states* $\omega \in \Omega$. The *state of the system* is a positive finite measure Q such that $\mathsf{Q} \ll \mathsf{P}$, where P is a fixed positive σ-finite measure on Ω. The value $\mathsf{Q}(\Omega)$ is interpreted as the *mass* of the system and $\mathsf{P}(\Omega)$ is interpreted as the *volume* of the system. If the system is in a state Q, then, for any measurable set $A \subseteq \Omega$, the total mass of the particles, whose states belong to A, is $\mathsf{Q}(A)$. Let $X : \Omega \to \mathbb{R}_+$ be a function, whose value at a point ω is interpreted as the *energy level* corresponding to ω. The *interior energy* of the system in a state Q is $\int_\Omega X d\mathsf{Q}$ (i.e. we consider a system with no interaction). The (Boltzmann) *entropy* of the system in a state Q is $-\int_\Omega \ln\bigl(\frac{1}{\mathsf{Q}(\Omega)} \frac{d\mathsf{Q}}{d\mathsf{P}}\bigr) d\mathsf{Q}$ (note that if Q is a probability measure, then this quantity coincides with $H(\mathsf{Q}, \mathsf{P})$ up to a sign).

In this interpretation, the problem of the relative entropy minimization under linear constraints means maximizing the entropy under the fixed interior energy (here the constraint $\int_\Omega X d\mathsf{Q} = 0$ should be replaced by the constraint $\int_\Omega X d\mathsf{Q} = E$, where E is a positive number). The problem of finding the measure Q_*, at which the minimum of $H(\mathsf{Q}, \mathsf{P})$ (over the measures $\mathsf{Q} \ll \mathsf{P}$ satisfying $\mathsf{Q}(\Omega) = M$ and $\int_\Omega X d\mathsf{Q} = E$) is attained, means determining the *stable state* of a system corresponding to the mass M and the interior energy E.

The value

$$\int_\Omega X d\mathsf{Q} + T \int_\Omega \ln\Bigl(\frac{1}{\mathsf{Q}(\Omega)} \frac{d\mathsf{Q}}{d\mathsf{P}}\Bigr) d\mathsf{Q} \tag{3.1}$$

is interpreted as the *free energy* of the system in a state Q, while the Lagrange multiplier $T > 0$ is interpreted as the *temperature.*

In this interpretation, the problem of the minimization of (3.1) (see Lemma 2.1) means minimizing the free energy under the fixed temperature. The problem of finding the measure Q_*, at which the minimum of (3.1) (over all the measures $\mathsf{Q} \ll \mathsf{P}$ with $\mathsf{Q}(\Omega) = M$) is attained, means determining the stable state corresponding to the mass M and the temperature T.

The statement below follows from Lemma 2.1. It yields the stable state, the free energy, the interior energy, and the entropy of the stable state.

Theorem 3.1. *Let $M > 0$, $T > 0$. Suppose that $\int_\Omega e^{-X/T} d\mathsf{P} < \infty$ and $\int_\Omega X e^{-X/T} d\mathsf{P} < \infty$.*

(i) *The minimum of* (3.1) *over the positive measures* $\mathsf{Q} \ll \mathsf{P}$ *with* $\mathsf{Q}(\Omega) = M$ *is attained at the unique measure*

$$\mathsf{Q}_* = M \frac{e^{-X/T}}{\int_\Omega e^{-X/T} d\mathsf{P}} \mathsf{P}.$$

(ii) *Let* F, E, *and* S *denote the free energy, the interior energy, and the entropy of* Q_*. *Then*

$$F = -MT \ln\Big(\int_\Omega e^{-X/T} d\mathsf{P}\Big),$$

$$E = M \frac{\int_\Omega X e^{-X/T} d\mathsf{P}}{\int_\Omega e^{-X/T} d\mathsf{P}},$$

$$S = \frac{M}{T} \frac{\int_\Omega X e^{-X/T} d\mathsf{P}}{\int_\Omega e^{-X/T} d\mathsf{P}} + M \ln\Big(\int_\Omega e^{-X/T} d\mathsf{P}\Big).$$

(iii) *Suppose that* P *is finite, i.e.* $\mathsf{P} = V\mathsf{P}_0$, *where* P_0 *is a probability measure. Then*

$$\frac{\partial F}{\partial V} = -p, \quad \frac{\partial E}{\partial V} = 0, \quad \frac{\partial S}{\partial V} = \rho, \quad \frac{\partial F}{\partial T} = -S, \quad \frac{\partial E}{\partial T} = T\frac{\partial S}{\partial T},$$

where $p := \frac{MT}{V}$ *is interpreted as the pressure and* $\rho := \frac{M}{V}$ *is interpreted as the density.*

(iv) *If the measure* $\mathsf{P} \circ X^{-1}$ *is not degenerate, then* E *and* S *are strictly increasing in* T.

(v) *Set* $a = \sup\{x \in \mathbb{R} : X \geq x \ \mathsf{P}\text{-a.e.}\}$. *Then*

$$E \xrightarrow[T\downarrow 0]{} a, \quad S \xrightarrow[T\downarrow 0]{} M \ln \mathsf{P}(X = a).$$

Remark. The measure Q_* defined in (i) corresponds to the *Gibbs state* (see Landau and Lifshits [11]). The equalities in (iii) are known thermodynamic relations. The statement of (v) for the case, where P is the counting measure, corresponds to the *Nernst theorem* (see Landau and Lifshits [11]). □

2. Metastable states. The problem of finding the critical points of the functional $\mathsf{Q} \mapsto H(\mathsf{Q}, \mathsf{P})$ (see Theorem 2.6) means determining the *metastable states* of a system, i.e. the states, in which the system may stay for a very long time before it reaches the stable state. Theorem 2.6 shows that the metastable states have the form

$$\mathsf{Q}_*^A = M \frac{e^{-X/T}}{\int_A e^{-X/T} d\mathsf{P}} I_A \mathsf{P},$$

where $A \in \mathcal{F}$. The statements of Theorem 3.1 (ii)–(v) remain true for the free energy, the interior energy, and the entropy of Q_*^A if one replaces P by $\mathsf{P}|_A$.

4 Information Theory: Data Compression

Let Ω be a collection of *codewords* (Ω may be finite or countable). Each codeword ω has its *cost* $X(\omega) \geq 0$. Let $(\Omega_0, \mathcal{F}_0, \mathsf{P}_0)$ be a probability space interpreted as the *source* of information. The source *coding* is a map $F : \Omega_0 \to \Omega$. This map induces the measure $\mathsf{Q} = \mathsf{Q}_0 \circ F^{-1}$ on Ω. We will call the quantity $\sum_{\omega\in\Omega} X(\omega) q(\omega)$ the *average cost* of the coding F. The quantity $-\sum_{\omega\in\Omega} q(\omega) \ln q(\omega)$ will be called the *amount of information* contained in the coding F. The problem of the *data compression* means maximizing the amount of information contained in a coding under the fixed average cost of the coding.

Consider the function

$$\varphi(\tau) = \ln \sum_{\omega\in\Omega} e^{\tau X(\omega)}, \quad \tau \in \mathbb{R}$$

(it may take on the value $+\infty$). Note that φ is differentiable inside the interval $\{\tau \in \mathbb{R} : \varphi(\tau) < \infty\}$.

In what follows, we will use the notation

$$H(C) = \sup\Big\{-\sum_{\omega\in\Omega} q(\omega)\ln q(\omega)\Big\},$$

where the supremum is taken over the probability measures Q on Ω with $\sum_{\omega\in\Omega} X(\omega)q(\omega) \le C$.

Theorem 4.1. *Let $C > 0$. Suppose that there exists $\tau_* \in \mathbb{R}$ such that $\varphi'(\tau_*) = C$. Then*

$$H(C) = \varphi(\tau_*) - C\tau_*.$$

`Proof`. This statement follows from Theorem 2.2 (we only need to consider $X - C$ instead of X and to take P as a counting measure). □

Example 4.2. *Let $\Omega = \{(a_1, \dots, a_n) : n \in \mathbb{N},\ a_i \in A\}$, where A is a D-ary alphabet. Let the cost of a codeword $(a_1, \dots, a_n)$ be n. Then*

$$\lim_{C\to\infty} \frac{H(C)}{C} = \ln D.$$

`Proof`. We have

$$\varphi(\tau) = \ln \sum_{n=1}^{\infty} e^{\tau n} D^n = \ln \frac{De^\tau}{1 - De^\tau} = \ln\Big(\frac{1}{1 - De^\tau} - 1\Big), \quad \tau \in \mathbb{R}.$$

The function φ is finite on $(-\infty, -\ln D)$ and

$$\varphi'(\tau) = \frac{De^\tau(1 - De^\tau)}{(1 - De^\tau)^2 De^\tau} = \frac{1}{1 - De^\tau}, \quad \tau \in (-\infty, -\ln D).$$

For the point $\tau_* = \tau_*(C)$ defined as the solution of the equation $\varphi'(\tau_*) = C$, we have

$$\lim_{C\to\infty} \frac{\varphi(\tau_*) - C\tau_*}{C} = \lim_{C\to\infty} \frac{\ln(C-1)}{C} - \lim_{C\to\infty} \tau_* = -\lim_{C\to\infty} \tau_* = \ln D$$

(see Figure 1). □

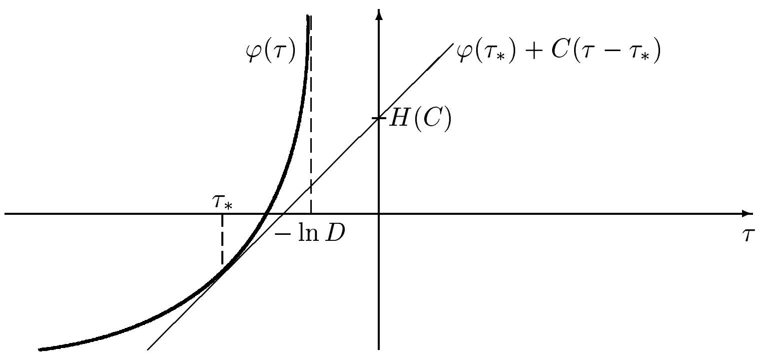

Figure 1

Example 4.3. *Let* $\Omega = \{(a_1, \dots, a_n) : n \in \mathbb{N},\ a_i \in A\}$, *where* $A = \{A_1, \dots, A_D\}$ *is a* D*-ary alphabet. Let the cost of a letter* A_i *be* $\lambda(A_i)$, *so that the cost of a codeword* $(a_1, \dots a_n)$ *is* $\lambda(a_1) + \dots + \lambda(a_n)$. *Then*

$$\lim_{C\to\infty} \frac{H(C)}{C} = -\tau_0,$$

where τ_0 *is defined as the solution of the equation*

$$\sum_{i=1}^{D} e^{\tau_0 \lambda(A_i)} = 1.$$

`Proof`. We have

$$\begin{aligned}\varphi(\tau) &= \ln \sum_{n=1}^{\infty} \sum_{i_1,\dots,i_n} \exp\{\tau\lambda(A_{i_1}) + \dots + \tau\lambda(A_{i_n})\} \\ &= \ln \sum_{n=1}^{\infty} \Bigl(\sum_{i=1}^{D} e^{\tau\lambda(A_i)}\Bigr)^n = \ln \frac{\eta(\tau)}{1-\eta(\tau)} = \ln\Bigl(\frac{1}{1-\eta(\tau)} - 1\Bigr), \qquad \tau \in \mathbb{R},\end{aligned}$$

where

$$\eta(\tau) = \sum_{i=1}^{D} e^{\tau\lambda(A_i)}, \qquad \tau \in \mathbb{R}.$$

The function φ is finite on the interval $(-\infty, \tau_0)$, and

$$\varphi'(\tau) = \frac{\eta'(\tau)}{\eta(\tau)(1-\eta(\tau))}, \qquad \tau \in (-\infty, \tau_0).$$

For the point $\tau_* = \tau_*(C)$ defined as the solution of the equation $\varphi'(\tau_*) = C$, we have

$$\lim_{C\to\infty} \frac{\varphi(\tau_*) - C\tau_*}{C} = \lim_{C\to\infty} \frac{1}{C} \ln\Bigl(\frac{C\eta(\tau_*)}{\eta'(\tau_*)} - 1\Bigr) - \lim_{C\to\infty} \tau_* = -\tau_0. \qquad \square$$

5 Financial Mathematics: Minimal Entropy Martingale Measure

1. One-period model. Let

$$\bigl(\Omega, \mathcal{F}, (\mathcal{F}_n)_{0\le n\le N}, \mathsf{P}, (X_n)_{0\le n\le N}\bigr) \tag{5.1}$$

be an asset price *model*. Here $(\mathcal{F}_n)$ is a filtration and X is a d-dimensional $(\mathcal{F}_n)$-adapted process. From the financial point of view, X_n^i is the (discounted) price of the ith asset at the time n. Let us recall some basic definitions and facts of financial mathematics.

Definition 5.1. A *self-financing strategy* is a pair $\pi = (x, H)$, where $x \in \mathbb{R}$ and $H = (H_n^i;\ n = 1, \dots, N, i = 1, \dots, d)$ is an $(\mathcal{F}_n)$-predictable process, i.e. H_n is $\mathcal{F}_{n-1}$-measurable for any $n = 1, \dots, N$. The value x is called the *initial capital* of the strategy π. The (discounted) *capital* of the strategy π is the process

$$X_n^\pi = x + \sum_{i=1}^{n} \langle H_i, \Delta X_i \rangle, \qquad n = 0, \dots, N,$$

where $\Delta X_n = X_n - X_{n-1}$.

Definition 5.2. A strategy $\pi = (x, H)$ realizes *arbitrage* if
(i) $x = 0$,
(ii) $X_N^\pi \geq 0$ P-a.s.
(iii) $\mathsf{P}(X_N^\pi > 0) > 0$.
A model is said to be *arbitrage-free* if such a strategy does not exist.

Let $\mathsf{P}_n(\omega)$ be the conditional distribution $\mathsf{Law}(\Delta X_n \,|\, \mathcal{F}_{n-1})(\omega)$. By $S_n(\omega)$ we denote the support of $\mathsf{P}_n(\omega)$. Let $L_n(\omega)$ be the smallest affine subspace of $\mathbb{R}^d$ containing $S_n(\omega)$. By $\overset{\circ}{S}_n(\omega)$ we denote the relative interior of $S_n(\omega)$, i.e. the interior of $S_n(\omega)$ in the relative topology of $L_n(\omega)$.

Proposition 5.3. (First Fundamental Theorem of Asset Pricing). *The following conditions are equivalent:*
(i) *the model is arbitrage-free;*
(ii) *there exists a probability measure* $\mathsf{Q} \sim \mathsf{P}$ *such that* X *is an* $(\mathcal{F}_n, \mathsf{Q})$*-martingale;*
(iii) *for any* $n = 1, \dots, N$ *and* P*-a.e.* ω, $0 \in \overset{\circ}{S}_n(\omega)$.

For the proof, see, for example, Shiryaev [16; Ch.V, §2e].

In the statements below we use the notation

$$\mathcal{M}^a = \{\mathsf{Q} : \mathsf{Q} \text{ is a probability measure on } (\Omega, \mathcal{F}),\ \mathsf{Q} \ll \mathsf{P}, \\ \text{and } X \text{ is an } (\mathcal{F}_n, \mathsf{Q})\text{-local martingale}\}.$$

An important problem of financial mathematics is to find "the most natural" element of $\mathcal{M}^a$. A possible way to solve it is to find the element of $\mathcal{M}^a$, which minimizes the relative entropy $H(\mathsf{Q}, \mathsf{P})$. The corresponding measure is called the *minimal entropy martingale measure.* The following result is well known. It states that, for the one-period model, the minimal entropy martingale measure is obtained as the *Esscher transform* of P.

Theorem 5.4. *Consider an arbitrage-free model of the form* (5.1) *with* $N = 1$. *Suppose that* $\mathcal{F}_0$ *is* P*-trivial and* $\mathsf{E}_\mathsf{P} e^{\langle \tau, \Delta X_1 \rangle} < \infty$ *for any* $\tau \in \mathbb{R}^d$. *Then there exists a point* τ_*, *at which the function* $\varphi(\tau) := \mathsf{E}_\mathsf{P} e^{\langle \tau, \Delta X_1 \rangle}$ *attains its minimum. The minimum of* $H(\mathsf{Q}, \mathsf{P})$ *over* $\mathcal{M}^a$ *is attained at the unique measure*

$$\mathsf{Q}_* = const\, e^{\langle \tau_*, \Delta X_1 \rangle} \mathsf{P}.$$

`Proof`. Note that in this case

$$\mathcal{M}^a = \{\mathsf{Q} : \mathsf{Q} \text{ is a probability measure on } (\Omega, \mathcal{F}),\ \mathsf{Q} \ll \mathsf{P},\ \mathsf{E}_\mathsf{Q} \Delta X_1 = 0\}.$$

The desired statement now follows from Theorem 2.2. □

2. Multiperiod model. Our next goal is to describe the structure of the minimal entropy martingale measure in the general multiperiod model. We need an auxiliary lemma.

Lemma 5.5. *Let* $(\Omega, \mathcal{F}, \mathsf{P})$ *be a probability space,* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-field,* $Y : \Omega \to \mathbb{R}$ *be a random variable, and* $X : \Omega \to \mathbb{R}^d$ *be a random vector such that* $\mathsf{E}_\mathsf{P}\big(e^{\langle \tau, X \rangle + Y} \,\big|\, \mathcal{G}\big) < \infty$ P*-a.s. for any* $\tau \in \mathbb{R}^d$. *Suppose that* $0 \in \overset{\circ}{S}(\omega)$ *for* P*-a.e.* ω, *where* $\overset{\circ}{S}(\omega)$ *denotes the relative interior of the support of the conditional distribution* $\mathsf{Law}(X \,|\, \mathcal{G})(\omega)$. *Then there exists a* $\mathcal{G} \times \mathcal{B}(\mathbb{R}^d)$*-measurable version of the function*

$$\varphi(\omega, \tau) = \mathsf{E}_\mathsf{P}\big(e^{\langle \tau, X \rangle + Y} \,\big|\, \mathcal{G}\big)(\omega), \quad \omega \in \Omega, \tau \in \mathbb{R}^d$$

such that, for any ω, *the map* $\tau \mapsto \varphi(\tau,\omega)$ *is continuous and* $\varphi(\tau,\omega) \xrightarrow[\|\tau\|\to\infty]{} \infty$. *For this version, there exists a* $\mathcal{G}$*-measurable map* $\tau_* : \Omega \to \mathbb{R}^d$ *such that, for* P*-a.e.* ω,

$$\tau_*(\omega) = \operatorname*{argmin}_{\tau\in\mathbb{R}^d} \varphi(\omega,\tau).$$

Proof. Let P_ω denote the conditional distribution $\mathsf{Law}(X,Y \,|\, \mathcal{G})(\omega)$. Let $\widetilde{X} : \mathbb{R}^{d+1} \to \mathbb{R}^d$ and $\widetilde{Y} : \mathbb{R}^{d+1} \to \mathbb{R}$ be the canonical mappings, i.e. $\widetilde{X}(x^1,\dots,x^{d+1}) = (x^1,\dots,x^d)$, $\widetilde{Y}(x^1,\dots,x^{d+1}) = x^{d+1}$. Then the function

$$\varphi(\tau,\omega) = \begin{cases} \mathsf{E}_{\mathsf{P}_\omega} e^{\langle \tau,\widetilde{X}\rangle + \widetilde{Y}} & \text{if } 0 \in \overset{\circ}{S}(\omega) \text{ and } \mathsf{E}_{\mathsf{P}_\omega} e^{\langle \lambda,\widetilde{X}\rangle + \widetilde{Y}} < \infty \ \forall \lambda \in \mathbb{R}^d, \\ 0 & \text{otherwise} \end{cases}$$

satisfies the desired properties.

In order to prove the existence of τ_*, we will consider the function

$$\eta(\omega) = \inf\{\varphi(\omega,\tau) : \tau \in \mathbb{R}^d\} = \inf\{\varphi(\omega,\tau) : \tau \in \mathbb{Q}^d\}.$$

This function is $\mathcal{G}$-measurable, and therefore, the set $\{(\omega,\tau) : \varphi(\omega,\tau) = \eta(\omega)\}$ belongs to $\mathcal{G} \times \mathcal{B}(\mathbb{R}^d)$. Now, the existence of τ_* follows from the measurable selection theorem (see Dellacherie and Meyer [5; Theorem 8.2., p. 252]). □

Let us now consider an arbitrage-free model of the form (5.1). Suppose that $\mathsf{E}_{\mathsf{P}}\big(e^{\langle \tau,\Delta X_n\rangle} \,|\, \mathcal{F}_{n-1}\big) < \infty$ P-a.s. for any $n = 1,\dots,N$, $\tau \in \mathbb{R}^d$ (for example, the conditionally Gaussian models satisfy this assumption). Construct the random variables $\tau_N^*,\dots,\tau_1^*$ (going downwards from N to 1) by the equality

$$\tau_n^* = \operatorname*{argmin}_{\tau\in\mathbb{R}^d} \mathsf{E}\bigg(\exp\bigg\{\langle \tau,\Delta X_n\rangle + \sum_{i=n+1}^N \langle \tau_i^*,\Delta X_i\rangle\bigg\}\bigg|\, \mathcal{F}_{n-1}\bigg) \tag{5.2}$$

as described in Lemma 5.5. The applicability of this lemma follows from Proposition 5.3 and the inequalities

$$\begin{aligned}
&\mathsf{E}_{\mathsf{P}}\bigg(\exp\bigg\{\langle \tau,\Delta X_n\rangle + \sum_{i=n+1}^N \langle \tau_i^*,\Delta X_i\rangle\bigg\}\bigg|\, \mathcal{F}_{n-1}\bigg)\\
&= \mathsf{E}_{\mathsf{P}}\bigg(e^{\langle \tau,\Delta X_n\rangle}\mathsf{E}_{\mathsf{P}}\bigg(\exp\bigg\{\langle \tau_{n+1}^*,\Delta X_{n+1}\rangle + \sum_{i=n+2}^N \langle \tau_i^*,\Delta X_i\rangle\bigg\}\,\bigg|\, \mathcal{F}_n\bigg)\bigg|\, \mathcal{F}_{n-1}\bigg)\\
&\le \mathsf{E}_{\mathsf{P}}\bigg(e^{\langle \tau,\Delta X_n\rangle}\mathsf{E}_{\mathsf{P}}\bigg(\exp\bigg\{\sum_{i=n+2}^N \langle \tau_i^*,\Delta X_i\rangle\bigg\}\,\bigg|\, \mathcal{F}_n\bigg)\bigg|\, \mathcal{F}_{n-1}\bigg)\\
&= \mathsf{E}_{\mathsf{P}}\bigg(\exp\bigg\{\langle \tau,\Delta X_n\rangle + \sum_{i=n+2}^N \langle \tau_i^*,\Delta X_i\rangle\bigg\}\bigg|\, \mathcal{F}_{n-1}\bigg)\\
&\le \dots \le \mathsf{E}_{\mathsf{P}}\big(e^{\langle \tau,\Delta X_n\rangle} \,\big|\, \mathcal{F}_{n-1}\big), \quad \tau \in \mathbb{R}^d.
\end{aligned}$$

(These inequalities are verified for $n = N,\dots,1$ going downwards from N to 1.)

Theorem 5.6. *Consider an arbitrage-free model of the form* (5.1). *Suppose that* $\mathsf{E}_\mathsf{P}\big(e^{\langle \tau, \Delta X_n\rangle} \,|\, \mathcal{F}_{n-1}\big) < \infty$ P*-a.s. for any* $n = 1,\dots,N$, $\tau \in \mathbb{R}^d$. *Then the minimum of* $H(\mathsf{Q},\mathsf{P})$ *over* $\mathcal{M}^a$ *is attained at the unique measure*

$$\mathsf{Q}_* = const\ \exp\bigg\{\sum_{n=1}^{N} \langle \tau_n^*, \Delta X_n\rangle\bigg\}\mathsf{P}, \tag{5.3}$$

where the random variables τ_n^* *are defined by* (5.2). *Furthermore,*

$$H(\mathsf{Q}_*,\mathsf{P}) = -\ln \mathsf{E}_\mathsf{P} \exp\bigg\{\sum_{n=1}^{N} \langle \tau_n^*, \Delta X_n\rangle\bigg\}.$$

`Proof`. Suppose that there exists a measure $\mathsf{Q} \in \mathcal{M}^a$ such that $H(\mathsf{Q},\mathsf{P}) < H(\mathsf{Q}_*,\mathsf{P})$. Set

$$U_n = \frac{d\mathsf{Q}|\mathcal{F}_n}{d\mathsf{P}|\mathcal{F}_n}, \quad n = 0,\dots,N, \tag{5.4}$$

$$V_n = \frac{e^{\langle \tau_n^*, \Delta X_n\rangle}\mathsf{E}_\mathsf{P}\big(\exp\{\sum_{i=n+1}^{N}\langle \tau_i^*, \Delta X_i\rangle\} \,|\, \mathcal{F}_n\big)}{\mathsf{E}_\mathsf{P}\big(\exp\{\sum_{i=n}^{N}\langle \tau_i^*, \Delta X_i\rangle\} \,|\, \mathcal{F}_{n-1}\big)}, \quad n = 0,\dots,N, \tag{5.5}$$

where $\mathcal{F}_{-1} := \{\emptyset, \Omega\}$, $\Delta X_0 := 0$. Consider the measures

$$\mathsf{Q}_n = U_n V_{n+1} \dots V_N \mathsf{P}, \quad n = -1,\dots,N,$$

where $U_{-1} := 1$. Let us prove that, for any $k = 0,\dots,N$,

$$H(\mathsf{Q}_{k-1},\mathsf{P}) \le H(\mathsf{Q}_k,\mathsf{P}) \tag{5.6}$$

and

$$\mathsf{E}_\mathsf{P}(V_k \dots V_N \ln V_k \dots V_N \,|\, \mathcal{F}_{k-1}) = Y_{k-1}, \tag{5.7}$$

where

$$Y_n = -\ln \mathsf{E}_\mathsf{P}\bigg(\exp\bigg\{\sum_{i=n+1}^{N}\langle \tau_i^*, \Delta X_i\rangle\bigg\}\bigg|\, \mathcal{F}_n\bigg), \quad n = -1,\dots,N.$$

Suppose that we have proved these statements for $k = n+1,\dots,N$. Let us prove them for $k = n$. We have

$$\begin{aligned} H(\mathsf{Q}_{n-1},\mathsf{P}) &= \mathsf{E}_\mathsf{P} U_{n-1}V_n\dots V_N \ln U_{n-1}V_n\dots V_N \\ &= \mathsf{E}_\mathsf{P} U_{n-1}V_n\dots V_N \ln V_{n+1}\dots V_N \\ &\quad + \mathsf{E}_\mathsf{P} U_{n-1}V_n\dots V_N \ln V_n \\ &\quad + \mathsf{E}_\mathsf{P} U_{n-1}V_n\dots V_N \ln U_{n-1} \\ &= \mathsf{E}_\mathsf{P}\mathsf{E}_\mathsf{P}(U_{n-1}V_n\dots V_N \ln V_{n+1}\dots V_N \,|\, \mathcal{F}_n) \\ &\quad + \mathsf{E}_\mathsf{P}\mathsf{E}_\mathsf{P}(U_{n-1}V_n\dots V_N \ln V_n \,|\, \mathcal{F}_n) \\ &\quad + \mathsf{E}_\mathsf{P}\mathsf{E}_\mathsf{P}(U_{n-1}V_n\dots V_N \ln U_{n-1} \,|\, \mathcal{F}_{n-1}) \\ &= \mathsf{E}_\mathsf{P} U_{n-1}V_n Y_n + \mathsf{E}_\mathsf{P} U_{n-1}V_n \ln V_n + \mathsf{E}_\mathsf{P} U_{n-1} \ln U_{n-1}. \end{aligned} \tag{5.8}$$

Here we used (5.7) for $k = n+1$ and the obvious equality $\mathsf{E}_\mathsf{P}(V_i \,|\, \mathcal{F}_j) = 1$ for $i > j$. Similarly,

$$H(\mathsf{Q}_n,\mathsf{P}) = \mathsf{E}_\mathsf{P} U_{n-1}W_n Y_n + \mathsf{E}_\mathsf{P} U_{n-1}W_n \ln W_n + \mathsf{E}_\mathsf{P} U_{n-1} \ln U_{n-1}, \tag{5.9}$$

where

$$W_n = I(U_{n-1} > 0)\frac{U_n}{U_{n-1}}.$$

Let P_ω denote the conditional distribution $\mathsf{Law}(\Delta X_n, Y_n, \tau_n^*, U_n, V_n, W_n \mid \mathcal{F}_{n-1})(\omega)$. Let $\widetilde{\Delta X}_n$, $\widetilde{Y}_n$, $\widetilde{\tau}_n^*$, $\widetilde{U}_n$, $\widetilde{V}_n$, $\widetilde{W}_n$ denote the coordinate maps on $\mathbb{R}^{2d+4}$ ($\widetilde{\Delta X}_n$ and $\widetilde{\tau}_n^*$ take on values in $\mathbb{R}^d$, while the other four maps take on values in $\mathbb{R}$). It follows from (5.7) with $k = n+1$ that $Y_n \ge 0$. Using the equality

$$\begin{aligned}
&\mathsf{E}_\mathsf{P}\big(e^{\langle \tau, \Delta X_n\rangle - Y_n} \,\big|\, \mathcal{F}_{n-1}\big)\\
&= \mathsf{E}_\mathsf{P}\Bigg(e^{\langle \tau, \Delta X_n\rangle}\mathsf{E}_\mathsf{P}\Bigg(\exp\Bigg\{\sum_{i=n+1}^{N}\langle \tau_i^*, \Delta X_i\rangle\Bigg\}\Bigg|\,\mathcal{F}_n\Bigg)\Bigg|\,\mathcal{F}_{n-1}\Bigg)\\
&= \mathsf{E}_\mathsf{P}\Bigg(\exp\Bigg\{\langle \tau, \Delta X_n\rangle + \sum_{i=n+1}^{N}\langle \tau_i^*, \Delta X_i\rangle\Bigg\}\Bigg|\,\mathcal{F}_{n-1}\Bigg), \quad \tau \in \mathbb{R}^d,
\end{aligned}$$

we get

$$\widetilde{\tau}_n^* = \operatorname*{argmin}_{\tau \in \mathbb{R}^d} \mathsf{E}_{\mathsf{P}_\omega} e^{\langle \tau, \widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} \tag{5.10}$$

for P-a.e. ω. Equality (5.5) can be rewritten as

$$\widetilde{V}_n = \frac{e^{\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle - \widetilde{Y}_n}}{\mathsf{E}_{\mathsf{P}_\omega} e^{\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle - \widetilde{Y}_n}} \tag{5.11}$$

for P-a.e. ω. The condition $\mathsf{E}_{\mathsf{P}_\omega}\big(e^{\langle \tau, \Delta X_n\rangle - Y_n} \,\big|\, \mathcal{F}_{n-1}\big) < \infty$ P-a.s. implies that

$$\mathsf{E}_{\mathsf{P}_\omega}\|\widetilde{\Delta X}_n\| e^{\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} < \infty \tag{5.12}$$

for P-a.e. ω. It follows from the equality $Y_n \ge 0$ that

$$\mathsf{E}_{\mathsf{P}_\omega}\widetilde{Y}_n e^{\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} < \infty \tag{5.13}$$

for P-a.e. ω. By the definition of W_n, we have

$$\mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n = 1 \tag{5.14}$$

for P-a.e. ω from the set $\{U_{n-1} > 0\}$. The condition $\mathsf{Q} \in \mathcal{M}^a$ implies that

$$\mathsf{E}_\mathsf{Q}(\|\Delta X_n\| \,|\, \mathcal{F}_{n-1}) = \mathsf{E}_\mathsf{P}(W_n\|\Delta X_n\| \,|\, \mathcal{F}_{n-1}) < \infty \quad \mathsf{P}\text{-a.s.}$$

(see [16; Ch. II, §1c]). Consequently,

$$\mathsf{E}_{\mathsf{P}_\omega}\|\widetilde{\Delta X}_n\|\widetilde{W}_n < \infty \tag{5.15}$$

for P-a.e. ω from the set $\{U_{n-1} > 0\}$. Thanks to conditions (5.11)–(5.15) and Lemma 2.1, we can write

$$\begin{aligned}
&- \mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n\widetilde{Y}_n + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n \ln \widetilde{V}_n\\
&\le -\mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n\langle \widetilde{\tau}_n^*, \widetilde{\Delta X}_n\rangle + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n\widetilde{Y}_n + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n \ln \widetilde{W}_n
\end{aligned} \tag{5.16}$$

for P-a.e. ω from the set $\{U_{n-1} > 0\}$ Indeed, if $\mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n\widetilde{Y}_n = \infty$, then this inequality is trivial; otherwise, it follows from Lemma 2.1. By (5.10) and (5.11),

$$\mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n\widetilde{\Delta X}_n = 0 \tag{5.17}$$

for P-a.e. ω. Due to the equality

$$\mathsf{E}_\mathsf{P}(W_n\Delta X_n \,|\, \mathcal{F}_{n-1}) = I(U_{n-1} > 0)\mathsf{E}_\mathsf{Q}(\Delta X_n \,|\, \mathcal{F}_{n-1}) = 0,$$

we have

$$\mathsf{E}_{\mathsf{P}_\omega}\widetilde{\Delta X}_n\widetilde{W}_n = 0 \tag{5.18}$$

for P-a.e. ω from the set $\{U_{n-1} > 0\}$. Combining (5.16)–(5.18) together, we get

$$\mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n\widetilde{Y}_n + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n \ln \widetilde{V}_n \le \mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n\widetilde{Y}_n + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{W}_n \ln \widetilde{W}_n$$

for P-a.e. ω from the set $\{U_{n-1} > 0\}$. Using (5.8) and (5.9), we obtain the inequality $H(\mathsf{Q}_{n-1},\mathsf{P}) \le H(\mathsf{Q}_n,\mathsf{P})$. Furthermore,

$$\begin{aligned}
\mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n\widetilde{Y}_n + \mathsf{E}_{\mathsf{P}_\omega}\widetilde{V}_n \ln \widetilde{V}_n &= \int_\Omega \widetilde{Y}_n \frac{e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n}}{\int_\Omega e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} d\mathsf{P}_\omega} d\mathsf{P}_\omega \\
&+ \int_\Omega \frac{e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n}}{\int_\Omega e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} d\mathsf{P}_\omega} \ln \frac{e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n}}{\int_\Omega e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} d\mathsf{P}_\omega} d\mathsf{P}_\omega \\
&= -\ln \int_\Omega e^{\langle\widetilde{\tau}_n^*,\widetilde{\Delta X}_n\rangle - \widetilde{Y}_n} d\mathsf{P}_\omega
\end{aligned}$$

for P-a.e. ω (here we used (5.17)). Hence,

$$\begin{aligned}
&\mathsf{E}_\mathsf{P}(V_n \dots V_N \ln V_n \dots V_N \,|\, \mathcal{F}_{n-1}) \\
&= \mathsf{E}_\mathsf{P}(V_n \dots V_N \ln V_{n+1} \dots V_N \,|\, \mathcal{F}_{n-1}) + \mathsf{E}_\mathsf{P}(V_n \dots V_N \ln V_n \,|\, \mathcal{F}_{n-1}) \\
&= \mathsf{E}_\mathsf{P}(V_n Y_n \,|\, \mathcal{F}_{n-1}) + \mathsf{E}_\mathsf{P}(V_n \ln V_n \,|\, \mathcal{F}_{n-1}) \\
&= -\ln \mathsf{E}_\mathsf{P}\big(e^{\langle\tau_n^*,\Delta X_n\rangle - Y_n} \,\big|\, \mathcal{F}_{n-1}\big) = Y_{n-1}.
\end{aligned}$$

Thus, we have proved (5.6) and (5.7) for $k = 0,\dots,N$.

Note that $\mathsf{Q}_* = \mathsf{Q}_{-1}$. Using (5.17), we get

$$\begin{aligned}
\mathsf{E}_{\mathsf{Q}_*}(\Delta X_n \,|\, \mathcal{F}_{n-1}) &= \frac{\mathsf{E}_\mathsf{P}(V_0 \dots V_N \Delta X_n \,|\, \mathcal{F}_{n-1})}{\mathsf{E}_\mathsf{P}(V_0 \dots V_N \,|\, \mathcal{F}_{n-1})} \\
&= \frac{\mathsf{E}_\mathsf{P}(\mathsf{E}_\mathsf{P}(V_n \dots V_N \Delta X_n \,|\, \mathcal{F}_n) \,|\, \mathcal{F}_{n-1})}{\mathsf{E}_\mathsf{P}(V_n \dots V_N \,|\, \mathcal{F}_{n-1})} \\
&= \frac{\mathsf{E}_\mathsf{P}(V_n \Delta X_n \,|\, \mathcal{F}_{n-1})}{\mathsf{E}_\mathsf{P}(V_n \dots V_N \,|\, \mathcal{F}_{n-1})} = 0.
\end{aligned}$$

Consequently, $\mathsf{Q}_* \in \mathcal{M}^a$. Due to the equality $\mathsf{Q} = \mathsf{Q}_N$, we have $H(\mathsf{Q}_*,\mathsf{P}) \le H(\mathsf{Q},\mathsf{P})$ and

$$H(\mathsf{Q}_*,\mathsf{P}) = \mathsf{E}_\mathsf{P}(V_0 \dots V_N \ln V_0 \dots V_N) = Y_{-1} = -\ln \mathsf{E}_\mathsf{P} \exp\Big\{\sum_{n=1}^N \langle\tau_n^*, \Delta X_n\rangle\Big\}.$$

Thus, the minimum of $H(\mathsf{Q},\mathsf{P})$ over $\mathcal{M}^a$ is attained at the measure Q_*. The uniqueness of the minimizing measure follows from the convexity of $\mathcal{M}^a$ and the strict convexity of the function $x \mapsto x \ln x$. □

`Remark`. The method of constructing the minimal entropy martingale measure described above is somewhat similar to the construction of some martingale measure through the *conditional Esscher transform*; see Rogers [14]. □

Corollary 5.7. *Consider an arbitrage-free model of the form* (5.1). *Suppose that* ΔX_n *is independent of* $\mathcal{F}_{n-1}$ *for any* $n = 1,\dots,N$, *and* $\mathsf{E}_\mathsf{P} e^{\langle\tau,\Delta X_n\rangle} < \infty$ *for any* $n = 1,\dots,N$, $\tau \in \mathbb{R}^d$. *Then, for each* $n = 1,\dots,N$, *there exists a point* $\tau_n^* \in \mathbb{R}^d$, *at which the function*

$\varphi_n(\tau) = \mathsf{E}_\mathsf{P} e^{\langle \tau, \Delta X_n \rangle}$ *attains its minimum. The minimum of* $H(\mathsf{Q},\mathsf{P})$ *over* $\mathcal{M}^a$ *is attained at the unique measure*

$$\mathsf{Q}_* = const\ \exp\left\{\sum_{n=1}^N \langle \tau_n^*, \Delta X_n \rangle\right\}\mathsf{P}.$$

`Proof`. It is sufficient to note that the conditional expectations in (5.2) coincide with the usual ones. The result now follows from Theorem 5.6. □

Corollary 5.8. *Consider a one-dimensional arbitrage-free model of the form* (5.1)*. Suppose that* $X_n = e^{Y_n}$*, where* $\Delta Y_n = Y_n - Y_{n-1}$ *is independent of* $\mathcal{F}_{n-1}$ *for any* $n = 1,\dots,N$*, and* $\mathsf{E}_\mathsf{P} e^{\tau(e^{\Delta Y_n}-1)} < \infty$ *for any* $n = 1,\dots,N$*,* $\tau \in \mathbb{R}$*. Then, for each* $n = 1,\dots,N$*, there exists a point* $\lambda_n^* \in \mathbb{R}$*, at which the function* $\varphi_n(\lambda) = \mathsf{E}_\mathsf{P} e^{\lambda(e^{\Delta Y_n}-1)}$ *attains its minimum. The minimum of* $H(\mathsf{Q},\mathsf{P})$ *over* $\mathcal{M}^a$ *is attained at the unique measure*

$$\mathsf{Q}_* = const\ \exp\left\{\sum_{n=1}^N \frac{\lambda_n^*}{X_{n-1}}\,\Delta X_n\right\}\mathsf{P}.$$

`Proof`. Let us prove that, for any $k = N,\dots,1$, the random variable τ_k^* defined by (5.2) equals $\frac{\lambda_k^*}{X_{k-1}}$. Suppose that we have proved this statement for $k = n+1,\dots,N$. Let us prove it for $k = n$. We have

$$\begin{aligned}
&\mathsf{E}_\mathsf{P}\left(\exp\left\{\tau\Delta X_n + \sum_{i=n+1}^N \tau_i^*\Delta X_i\right\}\Big|\,\mathcal{F}_{n-1}\right)\\
&= \mathsf{E}_\mathsf{P}\left(e^{\tau\Delta X_n}\,\mathsf{E}_\mathsf{P}\left(\exp\left\{\sum_{i=n+1}^N \lambda_i^*\left(e^{\Delta Y_i}-1\right)\right\}\Big|\,\mathcal{F}_n\right)\Big|\,\mathcal{F}_{n-1}\right)\\
&= \mathsf{E}_\mathsf{P}\exp\left\{\sum_{i=n+1}^N \lambda_i^*\left(e^{\Delta Y_i}-1\right)\right\}\mathsf{E}_\mathsf{P}\left(e^{\tau\Delta X_n}\,|\,\mathcal{F}_{n-1}\right)\\
&= \mathsf{E}_\mathsf{P}\exp\left\{\sum_{i=n+1}^N \lambda_i^*\left(e^{\Delta Y_i}-1\right)\right\}\varphi_n(\tau X_{n-1}).
\end{aligned}$$

Thus, we have verified that $\tau_k^* = \frac{\lambda_k^*}{X_{k-1}}$ for $k = N,\dots,1$. The desired result now follows from Theorem 5.6. □

3. Exponential utility maximization. As already mentioned, the problem of finding the minimal entropy martingale measure is dual to the problem of the exponential utility maximization.

Definition 5.9. The *exponential utility* of a strategy π is defined as

$$U(\pi) = -\mathsf{E}_\mathsf{P} e^{-X_N^\pi}.$$

Theorem 5.10. *Consider an arbitrage-free model of the form* (5.1)*. Suppose that* $\mathsf{E}_\mathsf{P}\left(e^{\langle\tau,\Delta X_n\rangle}\,\middle|\,\mathcal{F}_{n-1}\right) < \infty$ P*-a.s. for any* $n = 1,\dots,N$*,* $\tau \in \mathbb{R}^d$*. Let* $x \in \mathbb{R}$*. Then the maximum of* $U(\pi)$ *over all the strategies* π *with the initial capital* x *is attained at the unique strategy* $\pi_* = (x, H^*)$*, where* $H_n^* = -\tau_n^*$ *and the random variables* τ_n^* *are given by* (5.2)*.*

Proof. Suppose that there exists a strategy $\pi = (x, H)$ such that $U(\pi) > U(\pi^*)$. This means that

$$\mathsf{E_P} \exp\Big\{-\sum_{i=1}^{N} \langle H_i, \Delta X_i \rangle\Big\} < \mathsf{E_P} \exp\Big\{-\sum_{i=1}^{N} \langle H_i^*, \Delta X_i \rangle\Big\}.$$

Consider the strategies $\pi^{(n)} = (x, H^{(n)})$, $n = 0, \dots, N$ defined as follows

$$H_i^{(n)} = \begin{cases} H_i & \text{if } i \le n, \\ H_i^* & \text{if } i > n. \end{cases}$$

Then, for each $n = 1, \dots, N$, we have

$$\begin{aligned}
&\mathsf{E_P}\Big(\exp\Big\{-\sum_{i=1}^{N} \langle H_i^{(n-1)}, \Delta X_i \rangle\Big\} \,\Big|\, \mathcal{F}_{n-1}\Big) \\
&= \exp\Big\{-\sum_{i=1}^{n-1} \langle H_i, \Delta X_i \rangle\Big\} \mathsf{E_P}\Big(\exp\Big\{-\langle H_n^*, \Delta X_n \rangle - \sum_{i=n+1}^{N} \langle H_i^*, \Delta X_i \rangle\Big\} \Big| \mathcal{F}_{n-1}\Big) \\
&\le \exp\Big\{-\sum_{i=1}^{n-1} \langle H_i, \Delta X_i \rangle\Big\} \mathsf{E_P}\Big(\exp\Big\{-\langle H_n, \Delta X_n \rangle - \sum_{i=n+1}^{N} \langle H_i^*, \Delta X_i \rangle\Big\} \Big| \mathcal{F}_{n-1}\Big) \\
&= \mathsf{E_P}\Big(\exp\Big\{-\sum_{i=1}^{N} \langle H_i^{(n)}, \Delta X_i \rangle\Big\} \,\Big|\, \mathcal{F}_{n-1}\Big),
\end{aligned}$$

which implies that $U(\pi^{(n-1)}) \le U(\pi^{(n)})$. Since $\pi_* = \pi^{(0)}$ and $\pi = \pi^{(N)}$, we obtain a contradiction. Thus, the maximum of $U(\pi)$ is attained at the strategy π_*.

The uniqueness of the maximizing strategy follows from the strict convexity of the function $x \mapsto -e^{-x}$. □

Remark. We have

$$-H(\mathsf{Q}_*, \mathsf{P}) = x + \ln(-U(\pi_*)),$$

where Q_* is the measure given by (5.3). □

References

[1] *J. Bucklew.* Large deviation techniques in decision, simulation, and estimation. Wiley, 1990.

[2] *T.M. Cover, J.A. Thomas.* Elements of information theory. Wiley, 1991.

[3] *I. Csiszar.* I-divergence geometry of probability distributions and minimization problems.// The Annals of Probability, **3** (1975), No. 1, p. 146–158.

[4] *F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer, C. Stricker.* Exponential hedging and entropic penalties.// Preprint.

[5] *C. Dellacherie, P.-A. Meyer.* Probabilités et potentiel. Vol. I, Hermann, 1976.

[6] *H. Föllmer, A. Schied.* Stochastic finance. An introduction in discrete time. Walter de Gruyter, 2002.

[7] *M. Frittelli.* The minimal entropy martingale measure and the valuation problem in incomplete markets.// Mathematical Finance, **10** (2000), No. 1, p. 39–52.

[8] *T. Fujiwara, Y. Miyahara.* On the minimal entropy martingale measure for geometric Lévy processes.// Discussion papers in economics, Nagoya City University, **299** (2001), p. 1–21.

[9] *T. Goll, L. Rüschendorf.* Minimax and minimal distance martingale measures and their relationship to portfolio optimization.// Preprint.

[10] *E.T. Jaynes.* Papers on probability, statistics, and statistical physics. Riedel, Dordrecht, 1983.

[11] *L.D. Landau, E.M. Lifshits.* Theoretical Physics. Vol. V: Statistical physics. Moscow, Nauka, 1995.

[12] *V.P. Maslov.* The notions of enropy, Hamiltonian, temperature, and thermodynamical limit in probability theory used for solving model problems in econophysics.// Russian Journal of Mathematical Physics, **9** (2002), No. 4, p. 437–445.

[13] *T. Miyahara, A. Novikov.* Geometric Lévy process pricing model.// Proceedings of the Steklov Mathematical Institute, **237** (2002), p. 185–200.

[14] *L.C.G. Rogers.* Equivalent martingale measures and no-arbitrage.// Stochastics and Stochastics Reports, **51** (1995), p. 41–50.

[15] *W. Schachermayer.* Optimal investment in incomplete markets when wealth may become negative.// Annals of Applied Probability, **11** (2001), No. 3, p. 694–734.

[16] *A.N. Shiryaev.* Essentials of stochastic finance. World Scientific, 1998.

[17] *R.L. Stratonovich.* Information theory. Moscow, Sovetskoe Radio, 1975 (in Russian).