# RESEARCH

# Open Access

# Non-smooth analysis method in optimal investment-BSDE approach



Helin Wu<sup>1</sup>, Yong Ren<sup>2</sup> and Feng Hu<sup>3\*</sup>

\*Correspondence: hufengqf@163.com <sup>3</sup>School of Statistics, Qufu Normal University, Qufu, China Full list of author information is available at the end of the article

# Abstract

In this paper, the investment process is modeled by backward stochastic differential equation. We investigate a necessary condition for optimal investment problem by the method of non-smooth analysis. Furthermore, some applications of our result are given.

MSC: 60H10; 93E20; 49J52

**Keywords:** Backward stochastic differential equation (BSDE); Non-smooth analysis; Optimal investment

# **1** Introduction

Single-period mean-variance portfolio selection was first introduced and studied by Markowitz [14]. Markowitz's pioneer work [14] has laid down the foundation for modern financial portfolio theory. In 2000, Li and Ng [11] extended the Markowitz model to the dynamic setting. From then on, various continuous time Markowitz models were studied in much of the literature (see, e.g., [1, 12, 13, 21]).

Generally, there are two approaches, i.e. the forward (primal) approach and the backward (dual) approach, employed to solve the above problem in the continuous time case. The first approach is inspired by indefinite stochastic linear quadratic control (see, e.g., [3, 20]) and builds the relationship between the mean-variance problem and a family of indefinite stochastic linear quadratic optimal control problems (see, e.g., [12, 13, 21]). The second approach is first studied by Bielecki et al. in [1], which is the generalization of the well-known risk neutral computational approach in the discrete time case (see, e.g., [9, 17, 18]). The backward approach includes two steps: the first step is to compute the optimal terminal wealth; the second step is to compute the replicating portfolio strategy corresponding to the obtained optimal terminal wealth. As shown in [1], the optimal terminal wealth is first obtained using the Lagrange multiplier method and then the optimal replicating portfolio strategy is obtained by solving a backward stochastic differential equation (BSDE for short). Along this line, in 2008, Ji and Peng [10] used Ekeland's variational principle to obtain a necessary condition for portfolio optimization problem by the backward approach. In this paper, we study optimal investment problem by the method of nonsmooth analysis, which makes more general portfolio optimization problem be inside the scopes of our consideration.



© The Author(s) 2018. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Now the framework of optimal investment problem that we study are introduced. Let  $(\Omega, \mathcal{F}, P)$  be a probability space and  $(W_t)_{t\geq 0}$  be a *d*-dimensional standard Brownian motion with respect to filtration  $(\mathcal{F}_t)_{t\geq 0}$  generated by Brownian motion and all *P*-null subsets, i.e.,

$$\mathcal{F}_t = \sigma\{W_s; s \leq t\} \vee \mathcal{N},$$

where  $\mathcal{N}$  is the set of all *P*-null subsets. Fix a real number T > 0. Denote  $F := (\mathcal{F}_t)_{t \in [0,T]}$ Here, let  $L^2(\mathcal{F}_T)$  denote the space of  $\mathcal{F}_T$ -measurable random variables  $\xi : \Omega \mapsto R$  satisfying  $E[\xi^2] < \infty$ ;  $H_T^2(\mathbb{R}^n)$  denote the space of F-predictable processes  $\varphi : \Omega \times [0,T] \mapsto \mathbb{R}^n$ satisfying  $E[\int_0^T |\varphi_t|^2 dt] < \infty$  and  $S_T^2(\mathbb{R})$  denote the space of F-progressively measurable RCLL processes  $\varphi : \Omega \times [0,T] \mapsto \mathbb{R}$  satisfying  $E[\sup_{0 \le t \le T} |\varphi_t|^2] < \infty$ .

Given an initial capital x, the investment process can be modeled by the following onedimensional BSDE on [0, T]:

$$y_t = \xi + \int_t^T g(s, y_s, z_s) \,\mathrm{d}s - \int_t^T z_s \cdot \,\mathrm{d}W_s \tag{1}$$

with  $y_0 \leq x$ , where  $\xi \in L^2(\mathcal{F}_T)$  and  $g(\omega, t, y, z) : \Omega \times [0, T] \times R \times R^d \mapsto R$  is a function uniformly satisfying the Lipschitz condition, i.e., there exists a positive constant M such that for all  $(y_1, z_1), (y_2, z_2) \in R \times R^d$ 

$$\left| g(\omega, t, y_1, z_1) - g(\omega, t, y_2, z_2) \right| \le M \left( |y_1 - y_2| + |z_1 - z_2| \right) \tag{A.1}$$

and

$$g(\cdot, 0, 0) \in H^2_T(R).$$
 (A.2)

By Pardoux and Peng [15], we know that: Suppose that *g* satisfies (A.1) and (A.2). Then, for any  $\xi \in L^2(\mathcal{F}_T)$ , BSDE (1) has a unique pair of adapted processes  $(y_t, z_t) \in S_T^2(\mathbb{R}) \times H_T^2(\mathbb{R}^d)$ . Usually, we call  $\{y_t\}$  a wealth process and  $\{z_t\}$  a portfolio process.

Assume that the investor has initial wealth x, he invests it in the financial market according to BSDE (1). Denote  $\mathcal{E}_{0,T}^g(\xi) := y_0$ . Then his investment strategy is determined by all available terminal values for him, i.e.,

 $\mathcal{A}(x) \coloneqq \left\{ \xi \in L^2(\mathcal{F}_T) | \mathcal{E}_{0,T}^g(\xi) \le x \right\}.$ 

In this paper, we study the following problem: if there exists  $\xi^* \in A(x)$ , which is the optimal objective of problem

$$\min_{\xi \in \mathcal{A}(x)} \rho(\xi),\tag{2}$$

what is the necessary condition for  $\xi^*$ ? Here,  $\rho$  is a function defined on  $L^2(\mathcal{F}_T)$  and represents a risk measure. We also assume that  $\rho$  satisfies the Lipschitz condition.

This paper is organized as follows. In Sect. 2, we give some results about non-smooth analysis that are used in this paper. In Sect. 3, we study the optimal investment problem on wealth and portfolio processes. With the help of the method of non-smooth analysis, a necessary condition for an optimal objective is obtained, which generalizes the result of Ji and Peng [10]. In Sect. 4, we give some examples as the applications of our result.

#### 2 Some results about non-smooth analysis

In this section, we give some results about non-smooth analysis that are used in this paper. The following definitions, lemmas and propositions can be found in [4, 5].

Suppose that *X* is a Banach space, and *X*<sup>\*</sup> is its dual space. Obviously,  $L^2(\mathcal{F}_T)$  is a Banach space with the norm  $\|\xi\| := (E[\xi^2])^{\frac{1}{2}}$  for any  $\xi \in L^2(\mathcal{F}_T)$ , and  $L^2(\mathcal{F}_T)$  is also its dual space.

**Definition 1** Suppose that *f* is a function defined on *X*, Given an  $x \in X$ , if  $\limsup_{y \in X, y \to x, t \downarrow 0} \frac{f(y+tv)-f(y)}{t} \in R$  for any  $v \in X$ , denote

$$f^{o}(x;\nu) := \limsup_{y \in X, y \to x, t \downarrow 0} \frac{f(y+t\nu) - f(y)}{t}.$$
(3)

We call  $f^{o}(x; v)$  the generalized directional derivative of f at x.

*Remark* 1 Suppose that a function  $f : X \mapsto R$  satisfies Lipschitz condition, i.e., for any  $x_1$ ,  $x_2 \in X$ ,

$$|f(x_1) - f(x_2)| \le M ||x_1 - x_2||$$

holds for some M > 0, depending on f. By Definition 1, we have:

(i) Obviously,  $f^o(x; v)$  is sub-linear on X, then, by the Hahn–Banach theorem, the generalized directional derivative set of f at x

$$\partial^{o} f(x) \coloneqq \left\{ \zeta \in X^{*} | \zeta(\nu) \le f^{o}(x;\nu), \forall \nu \in X \right\}$$

$$\tag{4}$$

is nonempty and weak star compact in  $X^*$ .

(ii) If f(x) attains minimum value at some point  $x_0 \in X$ , i.e.,

$$f(x) \ge f(x_0), \quad \forall x \in X,$$

the Fermat condition

 $0 \in \partial^o f(x_0)$ 

holds.

Given a set  $C \subset X$ , the distance function  $d_C : X \mapsto R$  is defined as

$$d_C(x) := \inf_{y \in C} \|y - x\|,$$

for a given  $x \in X$ .

**Lemma 1** (Exact penalization) Suppose that f is a Lipschitz function with coefficient M defined on  $X, x^* \in C \subset X$  and f takes its minimum value at  $x^*$  on C. Then, for any  $\hat{M} \ge M$ ,  $g(x) := f(x) + \hat{M}d_C(x)$  attains minimum value at  $x_0$  on X. On the contrary, if  $\hat{M} > M$  and C is closed, then the minimum point of  $g, x_0$  must belong to C.

*Remark* 2 From Lemma 1, we know that, supposing that f is a Lipschitz function with coefficient M defined on X and C is a closed subset of X and letting  $g(y) := f(y) + \hat{M}d_C(y)$ , for any  $y \in X$ , where  $\hat{M} > M$ , then, if there exists  $x^* \in C$  such that

$$f(x^*) = \min_{x \in C} f(x),\tag{5}$$

then we have

$$g(x^*) = \min_{x \in X} g(x) = f(x^*).$$
 (6)

**Definition 2** Assume  $x \in C$ . Given  $v \in X$ , if  $d_C^o(x; v) = 0$ , then v is said to be a contingent derivative at x. Denote

$$T_C(x) := \{ v \in X | d_C^o(x; v) = 0 \}.$$

 $T_C(x)$  is the set of contingent derivatives at x. By polarity, the normal derivative set is defined as

$$N_C(x) := \big\{ \zeta \in X^* | \zeta(\nu) \le 0, \forall \nu \in T_C(x) \big\}.$$

**Proposition 1** *Suppose that*  $x \in C$ *, then we have* 

$$N_C(x) = cl \left\{ \bigcup_{\lambda \ge 0} \lambda \partial^o d_C(x) \right\},\,$$

where cl means the weak star closure.

**Proposition 2** Given an  $x \in X$ , suppose that h is Lipschitz in a neighborhood of x and  $0 \notin \partial^{\circ} h(x)$ . Let

$$C := \{ y \in X | h(y) \le h(x) \}.$$
(7)

Then we have

$$N_C(x) \subset \bigcup_{\lambda \ge 0} \lambda \partial^o h(x).$$
(8)

*Remark* 3 Given an  $x \in X$ , suppose that h is Lipschitz in a neighborhood of x and  $0 \notin \partial^{\circ}h(x)$ . Let  $C := \{y \in X | h(y) \le h(x)\}$ . We also assume that f is a Lipschitz function with coefficient M defined on X, and f takes its minimum value at x on C. By Definition 1, Remark 2, Definition 2 and Proposition 2, then the Fermat condition

$$0 \in \partial^o f(x) + N_C(x) \tag{9}$$

holds, i.e., there exist a non-negative constant  $\lambda$ ,  $\zeta \in \partial^o f(x)$  and  $\eta \in \partial^o h(x)$  such that

 $\zeta + \lambda \eta = 0$ 

holds.

**Definition 3** A function *f* defined on *X* is called strictly differentiable at  $x \in X$ , if there exists  $x^* \in X^*$  such that

$$\lim_{y \in X, y \to x, t \to 0^+} \frac{f(y + t\nu) - f(y)}{t} = \langle x^*, \nu \rangle,$$
(10)

for any  $\nu \in X$ .

**Lemma 2** A function f defined on X is strictly differentiable at  $x \in X$ , then  $\partial^o f(x) = \{x^*\}$ , where  $x^*$  can be obtained by (10).

**Lemma 3** Suppose that  $\{h_i, i = 1, 2, ..., n\}$  is a set of Lipschitz functions on X. Define

$$h(x) := \max\{h_i(x) | i = 1, 2, ..., n\}, \quad \forall x \in X.$$

For each  $x \in X$ , let I(x) be the subset of  $\{1, 2, ..., n\}$  satisfying that  $h_i(x) = h(x)$ ,  $i \in I(x)$ . Then we have

$$\partial^{o} h(x) \subset co \left\{ \partial^{o} h_{i}(x) | i \in I(x) \right\}, \tag{11}$$

where coA is the convex hull of A.

## 3 Maximum principle for optimal investment problem

In this section, our aim is to derive a necessary condition for optimal problem (2). Just as usual, in order to use the method of non-smooth analysis, we need the following assumption:

the risk measure 
$$\rho$$
 is Lipschitz. (A.3)

The following proposition shows that  $\mathcal{E}_{0,T}^{g}$  is Lipschitz in  $L^{2}(\mathcal{F}_{T})$ . For more details, see [2, 22, 23].

**Proposition 3** Suppose that g satisfies (A.1) and (A.2),  $\xi_i \in L^2(\mathcal{F}_T)$ , i = 1, 2, and  $(y_t^i, z_t^i)$  are the solutions of BSDE (1) with terminal values  $\xi_i$ , respectively, then there exists a constant C > 0 such that

$$E\left[\sup_{t\in[0,T]} \left|y_t^2 - y_t^1\right|^2\right] + E\left[\int_0^T \left|z_t^2 - z_t^1\right|^2 \mathrm{d}t\right] \le CE\left[\left|\xi_2 - \xi_1\right|^2\right].$$
(12)

*Remark* 4 From Proposition 3, we know that, supposing *g* satisfies (A.1) and (A.2), then, for any  $\xi_1, \xi_2 \in L^2(\mathcal{F}_T)$ ,

$$\left|\mathcal{E}_{0,T}^{g}(\xi_{1}) - \mathcal{E}_{0,T}^{g}(\xi_{2})\right| \le C^{\frac{1}{2}} \|\xi_{1} - \xi_{2}\|$$

holds.

The following theorem is our main result.

**Theorem 1** Under the conditions (A.1)–(A.3), if  $\xi^*$  is an optimal objective of problem (2), then there exist a non-negative constant  $\lambda, \zeta \in \partial^o \rho(\xi^*)$  and  $\eta \in \partial^o \mathcal{E}^g_{0,T}(\xi^*)$  such that

$$\zeta + \lambda \eta = 0 \tag{13}$$

holds.

In order to prove Theorem 1, we need the following lemma.

**Lemma 4** Suppose that g satisfies (A.1) and (A.2). Then, for any  $\xi \in L^2(\mathcal{F}_T)$ , we have  $0 \notin \partial^o \mathcal{E}^g_{0,T}(\xi)$ .

*Proof* For notational simplicity, we only consider the case d = 1. We assume that there exists  $\xi^* \in L^2(\mathcal{F}_T)$  such that  $0 \in \partial^o \mathcal{E}^g_{0,T}(\xi^*)$ . For any  $\xi \in L^2(\mathcal{F}_T)$ , let  $h(\xi) := \mathcal{E}^g_{0,T}(\xi)$ , then

$$h^o(\xi^*;\eta) \ge 0$$

holds, for any  $\eta \in L^2(\mathcal{F}_T)$ . For each  $\xi \in L^2(\mathcal{F}_T)$ , suppose that  $(y_t, z_t)$  and  $(y_t^r, z_t^r)$  are the solutions of the following BSDEs on [0, T]:

$$y_t = \xi + \int_t^T g(s, y_s, z_s) \, \mathrm{d}s - \int_t^T z_s \, \mathrm{d}W_s,$$
$$y_t^r = \xi + r\eta + \int_t^T g(s, y_s^r, z_s^r) \, \mathrm{d}s - \int_t^T z_s^r \, \mathrm{d}W_s,$$

respectively. Thus, we have

$$y_{t}^{r} - y_{t} = r\eta + \int_{t}^{T} \left[ g(s, y_{s}^{r}, z_{s}^{r}) - g(s, y_{s}, z_{s}) \right] ds - \int_{t}^{T} \left( z_{s}^{r} - z_{s} \right) dW_{s}$$
  
=  $r\eta + \int_{t}^{T} \left[ \alpha_{s} \left( y_{s}^{r} - y_{s} \right) + \beta_{s} \left( z_{s}^{r} - z_{s} \right) \right] ds - \int_{t}^{T} \left( z_{s}^{r} - z_{s} \right) dW_{s},$ 

where

$$\begin{aligned} \alpha_s &= \frac{g(s, y_s^r, z_s) - g(s, y_s, z_s)}{y_s^r - y_s} \mathbf{1}_{\{y_s^r - y_s \neq 0\}}, \\ \beta_s &= \frac{g(s, y_s^r, z_s^r) - g(s, y_s^r, z_s)}{z_s^r - z_s} \mathbf{1}_{\{z_s^r - z_s \neq 0\}}, \end{aligned}$$

which imply  $\sup_{s \in [0,T]} |\alpha_s| \le M$  and  $\sup_{s \in [0,T]} |\beta_s| \le M$  from the Lipschitz condition (A.1). Applying Itô's formula, we have

$$\mathcal{E}_{0,T}^{g}(\xi+r\eta)-\mathcal{E}_{0,T}^{g}(\xi)=E\bigg[r\eta\cdot\exp\bigg\{\int_{0}^{T}\bigg[\alpha_{s}-\frac{1}{2}|\beta_{s}|^{2}\bigg]\mathrm{d}s+\int_{0}^{T}\beta_{s}\,\mathrm{d}W_{s}\bigg\}\bigg].$$

Let  $\eta = -1$ , then, by Definition 1, we can deduce that

$$h^{o}(\xi^{*};\eta) = \limsup_{\xi \to \xi^{*}, r \downarrow 0} \frac{\mathcal{E}^{g}_{0,T}(\xi + r\eta) - \mathcal{E}^{g}_{0,T}(\xi)}{r} < 0.$$

We get a contradiction with  $h^o(\xi^*; \eta) \ge 0$ . The proof of Lemma 4 is complete.

*Proof of Theorem* 1 By Lemma 4, we have  $0 \notin \partial^o \mathcal{E}^g_{0,T}(\xi)$ , for any  $\xi \in L^2(\mathcal{F}_T)$ . Hence by Definition 1, Remark 2, Definition 2 and Proposition 1, we can see that, if  $\xi^*$  is an optimal objective of problem (2), then the Fermat condition

$$0 \in \partial^{o} \rho\left(\xi^{*}\right) + N_{\mathcal{A}(x)}\left(\xi^{*}\right) \tag{14}$$

holds.

*Case* 1:  $\mathcal{E}_{0,T}^g(\xi^*) = x$ . In this case,  $\mathcal{A}(x) = \{\xi \in L^2(\mathcal{F}_T) | \mathcal{E}_{0,T}^g(\xi) \le \mathcal{E}_{0,T}^g(\xi^*)\}$ . Thus, by Proposition 2, there exist a non-negative constant  $\lambda, \zeta \in \partial^o \rho(\xi^*)$  and  $\eta \in \partial^o \mathcal{E}_{0,T}^g(\xi^*)$  such that

$$\zeta + \lambda \eta = 0 \tag{15}$$

holds.

*Case* 2:  $\mathcal{E}_{0,T}^g(\xi^*) = x_0 < x$ . In this case, denote

$$\widetilde{C} := \left\{ \xi \in L^2(\mathcal{F}_T) | \mathcal{E}_{0,T}^g(\xi) \le x_0 \right\}.$$

Since  $\xi^*$  is an optimal objective of problem (2),  $\xi^*$  is also an optimal objective of problem

$$\min_{\substack{\xi \in \widetilde{C}}} \rho(\xi). \tag{16}$$

Thus, by Definition 1, Remark 2, Definition 2 and Proposition 2 again, Equation (15) also holds. The proof of Theorem 1 is complete.  $\Box$ 

From Remark 4, we know that  $\mathcal{E}_{0,T}^g$  is Lipschitz in  $L^2(\mathcal{F}_T)$ . By Definition 1, we can see that  $\partial^o \mathcal{E}_{0,T}^g(\xi)$  is not empty for any  $\xi \in L^2(\mathcal{F}_T)$ . Indeed, we have following result.

**Lemma 5** Suppose that  $\xi \in L^2(\mathcal{F}_T)$  and g satisfies (A.1) and (A.2). Let  $(y_t, z_t)$  be the solution of BSDE (1) with terminal value  $\xi$ , then for any  $\eta \in \partial^o \mathcal{E}^g_{0,T}(\xi)$  and  $\zeta \in L^2(\mathcal{F}_T)$ , there exists  $(\varphi_t, \psi_t) \in \partial^o g(t, y_t, z_t)$  such that

 $\langle \eta, \zeta \rangle = \tilde{y}_0$ 

holds, where  $(\tilde{y}_t, \tilde{z}_t)$  is the solution of the following BSDE on [0, T]:

$$\tilde{y}_t = \zeta + \int_t^T (\varphi_s \tilde{y}_s + \psi_s \tilde{z}_s) \,\mathrm{d}s - \int_t^T \tilde{z}_s \cdot \,\mathrm{d}W_s. \tag{17}$$

*Proof* For any given  $\omega \in \Omega$  and  $t \in [0, T]$ , let  $\hat{g}(t, \hat{y}, \hat{z}) = g^o(t, y_t, z_t; \hat{y}, \hat{z}), \forall (\hat{y}, \hat{z}) \in \mathbb{R} \times \mathbb{R}^d$ . By Definition 1, we can see that  $\hat{g}(t, \hat{y}, \hat{z})$  is Lipschitz, positively homogeneous and convex in  $(\hat{y}, \hat{z})$ , and hence

$$\hat{g}(t,\hat{y},\hat{z}) = \max_{(\varphi_t,\psi_t)\in\partial^o g(t,y_t,z_t)} \langle (\varphi_t,\psi_t), (\hat{y},\hat{z}) \rangle.$$
(18)

For any  $\zeta \in L^2(\mathcal{F}_T)$ , we consider the following BSDE on [0, T]:

$$\hat{y}_t = \zeta + \int_t^T \hat{g}(s, \hat{y}_s, \hat{z}_s) \,\mathrm{d}s - \int_t^T \hat{z}_s \cdot \,\mathrm{d}W_s. \tag{19}$$

Let  $\hat{f}(\zeta) := \hat{y}_0$ . Denote

$$D := \left\{ \eta \in L^2(\mathcal{F}_T) | \langle \eta, \zeta \rangle = \tilde{y}_0, \forall (\varphi_t, \psi_t) \in \partial^o g(t, y_t, z_t) \right\},\tag{20}$$

where  $(\tilde{y}_t, \tilde{z}_t)$  is the solution of the following BSDE on [0, T]:

$$\tilde{y}_t = \zeta + \int_t^T (\varphi_s \tilde{y}_s + \psi_s \tilde{z}_s) \, \mathrm{d}s - \int_t^T \tilde{z}_s \cdot \, \mathrm{d}W_s.$$

By the comparison theorem of BSDE (for example, see El Karoui et al. [7]), we have  $\hat{f}(\zeta) \geq \langle \eta, \zeta \rangle$ , for any  $\eta \in D$ . Since  $\hat{g}$  is positively homogeneous and convex in  $(\hat{y}, \hat{z})$ , by Propositions 3.3 and 3.4 in Peng [16], we know that  $\hat{f}$  is positively homogeneous and convex in  $L^2(\mathcal{F}_T)$ , and hence

$$\hat{f}(\zeta) = \max_{\eta \in D} \langle \eta, \zeta \rangle.$$
(21)

By Definition 1, we can obtain that

$$\left(\mathcal{E}_{0,T}^{g}\right)^{o}(\xi;\zeta) = \hat{f}(\zeta) = \max_{\eta\in D} \langle \eta, \zeta \rangle.$$

This means that  $\partial^o \mathcal{E}^g_{0,T}(\xi) = D$ . The proof of Lemma 5 is complete.

In Ji and Peng [10], the authors assume that *g* is continuously differentiable in (y, z). In this special case, we can get an explicit form of  $\partial^o \mathcal{E}^g_{0,T}$ .

**Lemma 6** Suppose that  $\xi \in L^2(\mathcal{F}_T)$  and g satisfies (A.1) and (A.2). Let  $(y_t, z_t)$  be the solution of BSDE (1) with terminal value  $\xi$ . If g is continuously differentiable in  $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ , then we have

$$\partial^o \mathcal{E}^g_{0,T}(\xi) = \{q_T\}$$

and for any  $\eta \in L^2(\mathcal{F}_T)$ ,

$$\langle q_T, \eta \rangle = \tilde{y}_0,\tag{22}$$

where  $(\tilde{y}_t, \tilde{z}_t)$  is the solution of the following BSDE on [0, T]:

$$\tilde{y}_t = \eta + \int_t^T \left[ g'_y(s, y_s, z_s) \tilde{y}_s + g'_z(s, y_s, z_s) \tilde{z}_s \right] \mathrm{d}s - \int_t^T \tilde{z}_s \cdot \mathrm{d}W_s.$$

*Proof* For any given  $\omega \in \Omega$  and  $t \in [0, T]$ , let  $\tilde{g}(t, \tilde{y}, \tilde{z}) = g^o(t, y_t, z_t; \tilde{y}, \tilde{z}), \forall (\tilde{y}, \tilde{z}) \in \mathbb{R} \times \mathbb{R}^d$ . Since *g* satisfies (A.1), (A.2), and is continuously differentiable in  $(y, z) \in \mathbb{R} \times \mathbb{R}^d$ , we know that *g* is strictly differentiable in  $(y, z) \in \mathbb{R} \times \mathbb{R}^d$  and

$$\tilde{g}(t, \tilde{y}, \tilde{z}) = g'_{y}(t, y_t, z_t)\tilde{y} + g'_{z}(t, y_t, z_t)\tilde{z}$$

holds. For any  $\eta \in L^2(\mathcal{F}_T)$ , let  $(\tilde{y}_t, \tilde{z}_t)$  be the solution of the following BSDE on [0, T]:

$$\tilde{y}_t = \eta + \int_t^T \tilde{g}(s, \tilde{y}_s, \tilde{z}_s) \,\mathrm{d}s - \int_t^T \tilde{z}_s \cdot \,\mathrm{d}W_s.$$

Denote  $\mathcal{E}_{0,T}^{\tilde{g}}(\eta) := \tilde{y}_0$ . By Definition 1, we obtain

$$\left(\mathcal{E}_{0,T}^{g}\right)^{o}(\xi;\eta) = \mathcal{E}_{0,T}^{\tilde{g}}(\eta).$$
<sup>(23)</sup>

Since  $\tilde{g}$  is linear in  $(\tilde{y}, \tilde{z})$ , by Propositions 3.3 and 3.4 in Peng [16], we know that  $\mathcal{E}_{0,T}^{\tilde{g}}$  is linear in  $L^2(\mathcal{F}_T)$ . By Proposition 3, we can see that  $\mathcal{E}_{0,T}^{\tilde{g}}$  is continuous in  $L^2(\mathcal{F}_T)$ . Thus, by the Riesz representation theorem, there exists  $q_T \in L^2(\mathcal{F}_T)$  such that

$$\mathcal{E}_{0,T}^g(\eta) = \langle q_T, \eta \rangle \tag{24}$$

holds for any  $\eta \in L^2(\mathcal{F}_T)$ . From (23) and (24), we have  $\partial^o \mathcal{E}^g_{0,T}(\xi) = \{q_T\}$ . The proof of Lemma 6 is complete.

By Theorem 1 and Lemma 6, we immediately obtain the following theorem.

**Theorem 2** Suppose that  $\rho$  satisfies (A.3), and g satisfies (A.1), (A.2), and is continuously differentiable in (y, z). Assume that  $\xi^*$  is an optimal objective of problem (2), let  $(y_t^*, z_t^*)$  be the solution of BSDE (1) with terminal value  $\xi^*$ . Then there exist a non-negative constant  $\lambda$  and  $\zeta \in \partial^o \rho(\xi^*)$  such that

 $\zeta + \lambda q_T^* = 0$ 

holds, where  $q_T^*$  can be obtained by (22).

# 4 Some examples

In this section, we give three examples for our obtained result. In the sequel, suppose that g satisfies (A.1), (A.2), and is continuously differentiable in (y, z).

Problem A. Minimize

 $E[\varphi(\xi)] - c$ 

subject to

$$\begin{cases} E[u(\xi)] = c, \\ \mathcal{E}_{0,T}^{g}(\xi) = x, \\ \mathcal{E}_{0,T}^{g}(\xi) \ge 0, \\ \xi \in L^{2}(\mathcal{F}_{T}), \end{cases}$$

where  $\varphi : R \mapsto R$  is Lipschitz and continuously differentiable, and  $u : R \mapsto R$  has bounded continuous derivative u'.

We can extend Problem A to a more general framework.

Problem B. Minimize

$$E[\varphi(\xi)] - c$$

subject to

$$\begin{cases} E[u(\xi)] \leq c, \\ \mathcal{E}_{0,T}^g(\xi) \leq x, \\ \mathcal{E}_{0,T}^g(\xi) \geq 0, \\ \xi \in L^2(\mathcal{F}_T), \end{cases}$$

where  $\varphi : R \mapsto R$  is Lipschitz and continuously differentiable, and  $u : R \mapsto R$  has bounded continuous derivative u'.

In order to study the optimization problems A and B, we need the following lemma.

**Lemma** 7 If  $u : R \mapsto R$  has bounded continuous derivative u', then the function  $\rho(\xi) = E[u(\xi)]$  for any  $\xi \in L^2(\mathcal{F}_T)$  is strictly differentiable, and  $\partial^o \rho(\xi) = \{u'(\xi)\}$ .

By the Fubini theorem and the dominated convergence theorem, we can easily prove Lemma 7. So we omit it.

**Theorem 3** If the optimal objective  $\xi^*$  of Problem A exists, there exist two constants  $\lambda_1$  and  $\lambda_2$  such that

$$\varphi'(\xi^*) + \lambda_1 u'(\xi^*) + \lambda_2 q_T^* = 0$$

holds, where  $\partial^o \mathcal{E}^g_{0,T}(\xi^*) = \{q_T^*\}.$ 

*Proof* Denote  $\rho(\xi) := E[\varphi(\xi)] - c$ , for any  $\xi \in L^2(\mathcal{F}_T)$  and

 $U \coloneqq \{\xi \in L^2(\mathcal{F}_T) | h(\xi) \le 0\},\$ 

where  $h(\xi) := \max\{E[u(\xi)] - c, c - E[u(\xi)], \mathcal{E}_{0,T}^g(\xi) - x, x - \mathcal{E}_{0,T}^g(\xi), -\mathcal{E}_{0,T}^g(\xi)\}$ . If the optimal objective  $\xi^*$  of Problem A exists, then, by Lemmas 6 and 7, we have  $\partial^o \rho(\xi^*) = \{\varphi'(\xi^*)\}, \partial^o E[u(\xi^*)] = \{u'(\xi^*)\}$  and  $\partial^o \mathcal{E}_{0,T}^g(\xi^*) = \{q_T^*\}.$ 

By Theorem 2 and Lemma 3, there exist two constants  $\lambda_1$  and  $\lambda_2$  such that

$$\varphi'(\xi^*) + \lambda_1 u'(\xi^*) + \lambda_2 q_T^* = 0$$

holds. The proof of Theorem 3 is complete.

**Theorem 4** If the optimal objective  $\xi^*$  of Problem B exists, then there exist three constants  $\lambda \ge 0$  and  $a, \beta \in [0, 1]$  satisfying  $\alpha + \beta \le 1$ , such that

$$\begin{aligned} \varphi'(\xi^*) + \lambda q_T^* &= 0, \qquad \mathcal{E}_{0,T}^g(\xi^*) - x > \max\{E[u(\xi^*)] - c, -\mathcal{E}_{0,T}^g(\xi^*)\}, \\ \varphi'(\xi^*) - \lambda q_T^* &= 0, \qquad -\mathcal{E}_{0,T}^g(\xi^*) > \max\{E[u(\xi^*)] - c, \mathcal{E}_{0,T}^g(\xi^*) - x\}, \\ \varphi'(\xi^*) + \lambda u'(\xi^*) &= 0, \qquad E[u(\xi^*)] - c > \max\{\mathcal{E}_{0,T}^g(\xi^*) - x, -\mathcal{E}_{0,T}^g(\xi^*)\}, \\ \varphi'(\xi^*) + \lambda [\alpha u'(\xi^*) + (1 - \alpha)q_T^*] &= 0, \qquad E[u(\xi^*)] - c = \mathcal{E}_{0,T}^g(\xi^*) - x > -\mathcal{E}_{0,T}^g(\xi^*), \\ \varphi'(\xi^*) + \lambda [\alpha u'(\xi^*) + (\alpha - 1)q_T^*] &= 0, \qquad E[u(\xi^*)] - c = -\mathcal{E}_{0,T}^g(\xi^*) > \mathcal{E}_{0,T}^g(\xi^*) - x, \\ \varphi'(\xi^*) + \lambda (2\alpha - 1)q_T^* &= 0, \qquad \mathcal{E}_{0,T}^g(\xi^*) - x = -\mathcal{E}_{0,T}^g(\xi^*) > E[u(\xi^*)] - c, \\ \varphi'(\xi^*) + \lambda [(\alpha - \beta)q_T^* + (1 - \alpha - \beta)u'(\xi^*)] &= 0, \\ \mathcal{E}_{0,T}^g(\xi^*) - x &= -\mathcal{E}_{0,T}^g(\xi^*) = E[u(\xi^*)] - c, \end{aligned}$$

holds, where  $\partial^o \mathcal{E}_{0,T}^g(\xi^*) = \{q_T^*\}.$ 

*Proof* Denote  $\rho(\xi) := E[\varphi(\xi)] - c$ , for any  $\xi \in L^2(\mathcal{F}_T)$  and

$$U := \left\{ \xi \in L^2(\mathcal{F}_T) | h(\xi) \le 0 \right\},\$$

where  $h(\xi) := \max\{E[u(\xi)] - c, \mathcal{E}_{0,T}^g(\xi) - x, -\mathcal{E}_{0,T}^g(\xi)\}$ . If the optimal objective  $\xi^*$  of Problem B exists, then, by Lemmas 6 and 7, we have  $\partial^o \rho(\xi^*) = \{\varphi'(\xi^*)\}, \ \partial^o E[u(\xi^*)] = \{u'(\xi^*)\}$  and  $\partial^o \mathcal{E}_{0,T}^g(\xi^*) = \{q_T^*\}$ .

*Case* 1:  $\mathcal{E}_{0,T}^{g}(\xi^*) - x > \max\{E[u(\xi^*)] - c, -\mathcal{E}_{0,T}^{g}(\xi^*)\}$ . By Theorem 2, there exists a non-negative constant  $\lambda$  such that  $\varphi'(\xi^*) + \lambda q_T^* = 0$  holds.

*Case* 2:  $-\mathcal{E}_{0,T}^g(\xi^*) > \max\{E[u(\xi^*)] - c, \mathcal{E}_{0,T}^g(\xi^*) - x\}$ . By Theorem 2, there exists a non-negative constant  $\lambda$  such that  $\varphi'(\xi^*) - \lambda q_T^* = 0$  holds.

*Case* 3:  $E[u(\xi^*)] - c > \max\{\mathcal{E}_{0,T}^g(\xi^*) - x, -\mathcal{E}_{0,T}^g(\xi^*)\}$ . By Theorem 2, there exists a non-negative constant  $\lambda$  such that  $\varphi'(\xi^*) + \lambda u'(\xi^*) = 0$  holds.

*Case* 4:  $E[u(\xi^*)] - c = \mathcal{E}_{0,T}^g(\xi^*) - x > -\mathcal{E}_{0,T}^g(\xi^*)$ . By Theorem 2 and Lemma 3, there exist two constants  $\lambda \ge 0$  and  $\alpha \in [0, 1]$  such that  $\varphi'(\xi^*) + \lambda[\alpha u'(\xi^*) + (1 - \alpha)q_T^*] = 0$  holds.

*Case* 5:  $E[u(\xi^*)] - c = -\mathcal{E}_{0,T}^g(\xi^*) > \mathcal{E}_{0,T}^g(\xi^*) - x$ . By Theorem 2 and Lemma 3, there exist two constants  $\lambda \ge 0$  and  $\alpha \in [0, 1]$  such that  $\varphi'(\xi^*) + \lambda[\alpha u'(\xi^*) + (\alpha - 1)q_T^*] = 0$  holds.

*Case* 6:  $\mathcal{E}_{0,T}^g(\xi^*) - x = -\mathcal{E}_{0,T}^g(\xi^*) > E[u(\xi^*)] - c$ . By Theorem 2 and Lemma 3, there exist two constants  $\lambda \ge 0$  and  $\alpha \in [0, 1]$  such that  $\varphi'(\xi^*) + \lambda(2\alpha - 1)q_T^* = 0$  holds.

*Case* 7:  $\mathcal{E}_{0,T}^g(\xi^*) - x = -\mathcal{E}_{0,T}^g(\xi^*) = E[u(\xi^*)] - c$ . By Theorem 2 and Lemma 3, there exist three constants  $\lambda \ge 0$  and  $a, \beta \in [0, 1]$  satisfying  $\alpha + \beta \le 1$ , such that  $\varphi'(\xi^*) + \lambda[(\alpha - \beta)q_T^* + (1 - \alpha - \beta)u'(\xi^*)] = 0$  holds. The proof of Theorem 4 is complete.

In the classic investment problem, one often takes the variance as a risk measure and the mean-variance method is used in much of the literature. But by Delbaen [6], and Föllmer and Schied [8], such a kind of risk measure is not perfect. We often take  $\rho(\cdot)$  as a coherent or convex risk measure in (2). In Rosazza Gianin [19], the author proved that if g is a sub-additive and positively homogeneous function satisfying (A.1) and (A.2), define  $\rho(\xi) := \mathcal{E}_{0,T}^g(-\xi)$ , for any  $\xi \in L^2(\mathcal{F}_T)$ , then  $\rho(\cdot)$  is a convex risk measure.

*Problem* C. Suppose f is sub-additive and homogeneous function satisfying (A.1) and (A.2) and independent of y, and  $f'_z$  is continuous in z. Minimize

$$\mathcal{E}_{0,T}^f(-\xi)$$

subject to

$$\begin{cases} E[\xi] \ge c, \\ \mathcal{E}_{0,T}^{g}(\xi) \le x \\ \mathcal{E}_{0,T}^{g}(\xi) \ge 0 \\ \xi \in L^{2}(\mathcal{F}_{T}) \end{cases}$$

**Theorem 5** If the optimal objective  $\xi^*$  of Problem C exists, then there exist three constants  $\lambda \ge 0$  and  $a, \beta \in [0, 1]$  satisfying  $\alpha + \beta \le 1$ , such that

$$\begin{aligned} k_T^* + \lambda q_T^* &= 0, \qquad \mathcal{E}_{0,T}^g(\xi^*) - x > \max\{c - E[\xi^*], -\mathcal{E}_{0,T}^g(\xi^*)\}, \\ k_T^* - \lambda q_T^* &= 0, \qquad -\mathcal{E}_{0,T}^g(\xi^*) > \max\{c - E[\xi^*], \mathcal{E}_{0,T}^g(\xi^*) - x\}, \\ k_T^* - \lambda &= 0, \qquad c - E[\xi^*] > \max\{\mathcal{E}_{0,T}^g(\xi^*) - x, -\mathcal{E}_{0,T}^g(\xi^*)\}, \\ k_T^* + \lambda[-\alpha + (1 - \alpha)q_T^*] &= 0, \qquad c - E[\xi^*] = \mathcal{E}_{0,T}^g(\xi^*) - x > -\mathcal{E}_{0,T}^g(\xi^*), \\ k_T^* + \lambda[-\alpha + (\alpha - 1)q_T^*] &= 0, \qquad c - E[\xi^*] = -\mathcal{E}_{0,T}^g(\xi^*) > \mathcal{E}_{0,T}^g(\xi^*) - x, \\ k_T^* + \lambda(2\alpha - 1)q_T^* &= 0, \qquad \mathcal{E}_{0,T}^g(\xi^*) - x &= -\mathcal{E}_{0,T}^g(\xi^*) > c - E[\xi^*], \\ k_T^* + \lambda[(\alpha - \beta)q_T^* + (\alpha + \beta - 1)] &= 0, \qquad \mathcal{E}_{0,T}^g(\xi^*) - x &= -\mathcal{E}_{0,T}^g(\xi^*) = c - E[\xi^*], \end{aligned}$$

holds, where  $\partial^o \mathcal{E}^g_{0,T}(\xi^*) = \{q_T^*\}$  and  $\partial^o \mathcal{E}^f_{0,T}(-\xi^*) = \{k_T^*\}$ . Let  $(y_t^*, z_t^*)$  be the solution of BSDE (1) with terminal value  $-\xi^*$ . For any  $\eta \in L^2(\mathcal{F}_T)$ ,  $\langle k_T^*, -\eta \rangle = \bar{y}_0$ .  $(\bar{y}_t, \bar{z}_t)$  is the solution of the following BSDE on [0, T]:

$$\bar{y}_t = -\eta + \int_t^T f_z'(s, z_s^*) \bar{z}_s \,\mathrm{d}s - \int_t^T \bar{z}_s \cdot \,\mathrm{d}W_s$$

The proof of Theorem 5 is very similar to that of Theorem 4. So we omit the proof.

#### Acknowledgements

The authors would like to thank the anonymous referees for their constructive suggestions and valuable comments.

#### Funding

The work of Helin Wu is supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1400922). The work of Yong Ren is supported by the National Natural Science Foundation of China (11871076). The work of Feng Hu is supported by the National Natural Science Foundation of China (11801307) and the Natural Science Foundation of Shandong Province (ZR2016JL002 and ZR2017MA012).

#### **Competing interests**

The authors declare that they have no competing interests.

#### Authors' contributions

The authors contributed equally and significantly in writing this article. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup>School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, China. <sup>2</sup>Department of Mathematics, Anhui Normal University, Wuhu, China. <sup>3</sup>School of Statistics, Qufu Normal University, Qufu, China.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Received: 5 June 2018 Accepted: 4 December 2018 Published online: 12 December 2018

#### References

- Bielecki, T.R., Jin, H.Q., Pliska, S.R., Zhou, X.: Continuous time mean-variance portfolio selection with bankruptcy prohibition. Math. Finance 15, 213–244 (2005)
- Briand, P., Coquet, F., Hu, Y., Mémin, J., Peng, S.: A converse comparison theorem for BSDEs and related properties of g-expectation. Electron. Commun. Probab. 5, 101–117 (2000)
- Chen, S., Li, X., Zhou, X.: Stochastic linear quadratic regulators with indefinite control weight costs. SIAM J. Control Optim. 36, 1685–1702 (1998)
- 4. Clarke, F.H.: Optimization and Nonsmooth Analysis, 2nd edn. Classics in Applied Mathematics, vol. 5. SIAM, Philadelphia (1990)
- Clarke, F.H., Ledyaev, Y.S., Stern, R.J., Wolenski, P.R.: Nonsmooth Analysis and Control Theory. Graduate Texts in Mathematics, vol. 178. Springer, New York (1998)
- Delbaen, F.: Coherent risk measures on general probability space. In: Sandmann, K., Schonbucher, P.J. (eds.) Advance in Finance and Stochastics, pp. 1–37. Springer, Berlin (2002)
- El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equations in finance. Math. Finance 7, 1–71 (1997)
- 8. Föllmer, H., Schied, A.: Stochastic Finance: An Introduction in Discrete Time. De Gruyter, Berlin (2004)
- 9. Harrison, J.M., Kreps, D.: Martingales and multiperiod securities market. J. Econ. Theory 20, 381–408 (1979)
- Ji, S., Peng, S.: Terminal perturbation method for the backward approach to continuous time mean-variance portfolio selection. Stoch. Process. Appl. 118, 952–967 (2008)
- Li, D., Ng, W.L.: Optimal dynamic portfolio selection: multiperiod mean variance formulation. Math. Finance 10, 387–406 (2000)
- 12. Li, X., Zhou, X., Lim, A.E.B.: Dynamic mean variance portfolio selection with no shorting constraints. SIAM J. Control Optim. 40, 1540–1555 (2001)
- 13. Lim, A.E.B., Zhou, X.: Mean variance portfolio selection with random parameters. Math. Oper. Res. 27, 101–120 (2002)
- 14. Markowitz, H.: Portfolio selection. J. Finance 7, 77-91 (1952)
- 15. Pardoux, E., Peng, S.: Adapted solution of a backward stochastic differential equation. Syst. Control Lett. 14, 55–61 (1990)
- 16. Peng, S.: Modelling derivatives pricing with their generating functions (2006) arXiv:math/0605599 [math.PR]
- Pliska, S.R.: A discrete time stochastic decision model. In: Fleming, W.H., Gorostiza, L.G. (eds.) Advances in Filtering and Optimal Stochastic Control, in. Lecture Notes in Control and Information Sciences, vol. 42, pp. 290–304. Springer, New York (1982)
- 18. Pliska, S.R.: Introduction to Mathematical Finance. Blackwell, Oxford (1997)
- 19. Rosazza Gianin, E.: Risk measures via *g*-expectations. Insur. Math. Econ. **39**, 19–34 (2006)
- 20. Yong, J., Zhou, X.: Stochastic Controls—Hamiltonian Systems and HJB Equations. Springer, New York (1999)
- Zhou, X., Li, D.: Continuous time mean variance portfolio selection: a stochastic LQ framework. Appl. Math. Optim. 42, 19–33 (2000)
- 22. Zong, Z.: A comonotonic theorem for backward stochastic differential equations in L<sup>p</sup> and its applications. Ukr. Math. J. **64**, 857–874 (2012)
- Zong, Z., Hu, F.: BSDEs under filtration-consistent nonlinear expectations and the corresponding decomposition theorem for ε-supermartingales in L<sup>p</sup>. Rocky Mt. J. Math. 43, 677–695 (2013)

# Submit your manuscript to a SpringerOpen<sup>o</sup> journal and benefit from:

- Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at > springeropen.com