Duality in convex stochastic optimization

Teemu Pennanen* Ari-Pekka Perkkiö†

May 11, 2022

Abstract

This paper studies duality and optimality conditions in general convex stochastic optimization problems introduced by Rockafellar and Wets in [28]. We derive an explicit dual problem in terms of two dual variables, one of which is the shadow price of information while the other one gives the marginal cost of a perturbation much like in classical Lagrangian duality. Existence of primal solutions and the absence of duality gap are obtained without compactness or boundedness assumptions. In the context of financial mathematics, the relaxed assumptions are satisfied under the well-known no-arbitrage condition and the reasonable asymptotic elasticity condition of the utility function. We extend classical portfolio optimization duality theory to problems of optimal semi-static hedging. Besides financial mathematics, we obtain several new results in stochastic programming and stochastic optimal control.

Keywords. Convex duality, stochastic programming, stochastic optimal control, financial mathematics

AMS subject classification codes. 90C15, 90C46, 46N10, 93E20, 46N10, 91G80

1 Introduction

Given a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t=0}^T$ (an increasing sequence of sub-σ-algebras of $\mathcal{F}$), consider the problem

$$\text{minimize} \qquad Ef(x,\bar{u}) := \int f(x(\omega),\bar{u}(\omega),\omega)dP(\omega) \quad \text{over } x \in \mathcal{N} \qquad (SP)$$

where $\mathcal{N}$ is a linear space of stochastic processes $x = (x_t)_{t=0}^T$ adapted to $(\mathcal{F}_t)_{t=0}^T$ (i.e., x_t is $\mathcal{F}_t$-measurable) and $\bar{u}$ is a $\mathbb{R}^m$-valued random variable. We assume

*Department of Mathematics, King's College London, Strand, London, WC2R 2LS, United Kingdom, teemu.pennanen@kcl.ac.uk

†Mathematics Institute, Ludwig-Maximilian University of Munich, Theresienstr. 39, 80333 Munich, Germany, a.perkkioe@lmu.de. Corresponding author

that f is a *convex normal integrand* on $\mathbb{R}^n \times \mathbb{R}^m \times \Omega$, i.e. $f(\cdot, \omega)$ is a closed convex function for every $\omega \in \Omega$ and $\omega \mapsto \operatorname{epi} f(\cdot, \omega)$ is an $\mathcal{F}$-measurable set-valued mapping; see [31, Chapter 14]. Here and in what follows, we define the integral of an extended real-valued random variable as $+\infty$ unless its positive part is integrable. The integral of any extended real-valued measurable function is then a well defined extended real number so it follows that Ef is a well-defined convex function on $L^0(\mathbb{R}^n \times \mathbb{R}^m)$.

Problems of the form (SP) were first studied in [28] where it was observed that many more specific stochastic optimization problems can be written in this unified format. Examples include more traditional formulations of stochastic programming, convex stochastic control and various problems in financial mathematics; see Section 6 below. In [28], problem (SP) was analyzed through dynamic programming and convex duality. Soon after, [7] extended the dynamic programming principle by removing the convexity assumption but, like [28], assumed the set of feasible solutions to be bounded. The boundedness assumptions were removed in [13, 16, 22, 2, 21, 19].

Like [28, 23, 13, 2], the present paper studies problem (SP) with the functional analytic techniques of convex duality. This will yield dual problems whose optimum values coincide with that of (SP) and whose optimal solutions can be used to characterize those of (SP). We extend the classic results of [28, 29, 30] so as to cover various duality results developed independently in stochastics and financial mathematics e.g. in [6, 33, 9, 34, 15]. The new results allow also for significant extensions to central models in stochastic programming, stochastic optimal control and financial mathematics. In particular, we extend the inequality constrained models of [29] by including equality constraints and allowing for unbounded strategies. In stochastic optimal control, we obtain a scenariowise maximum principle. We also extend the classical duality results of financial mathematics to optimal semistatic hedging problems where one optimizes over dynamic trading strategies as well as statically held derivative portfolios. In each application, we establish the existence of primal solutions and the absence of a duality gap.

Much like in [28, 29, 13, 2], our strategy is to analyze (SP) through the general duality framework of [27]. We deviate from the above references, however, in that we employ two dualizing parameters: the random vector $\bar{u}$ in (SP) and another one that perturbs the adaptedness constraint on x. This yields an explicit dual problem for (SP) in terms of two dual variables: one is the "shadow price of information" studied e.g. in [35, 28, 23, 4, 5, 18] and the other one gives the marginal cost of changing $\bar{u}$. As a special case, we obtain the dual problem of [29] for stochastic optimization problems with inequality constraints. We find new duality frameworks for many other problem classes including optimal stoping, stochastic optimal control and portfolio optimization. Moreover, our results apply without the compactness and boundedness assumptions made in [29].

Without the boundedness assumptions, problem (SP) does not directly fit the framework of [27] which assumes that the optimal solutions are sought from a locally convex vector space. We will thus first, in Section 3, restrict the decision strategies x to a locally convex space $\mathcal{X}$ of $\mathbb{R}^n$-valued random variables. Straightforward application of the functional analytic duality theory then yields a dual problem and optimality conditions for the restricted problem. We return to the original problem (SP) in Section 4 and find that its optimum value as a function of the parameters (z, u) has the same lower semicontinuous hull as that of the restricted problem. It follows that their dual problems coincide and, by an application of Fenchel inequality and Lemma 22, we find scenariowise optimality conditions for (SP). Section 5 recalls sufficient conditions for the lower semicontinuity of the optimum value function of (SP). Section 6 illustrates the new results with applications to more specific problems classes.

2 Integral functionals in duality

Convex duality is based on the theory of conjugate functions on dual pairs of locally convex topological vector spaces; see [27]. The first part of this section reviews spaces of random variables in separating duality with each other while the second part reviews conjugation of integral functionals on such spaces. This forms the functional analytic setting for the duality theory of stochastic optimization developed in the followup sections. For full generality, we make minimal assumptions on the spaces of random variables. The classical Lebesgue and Orlicz spaces, L^p and L^Φ are covered as special cases but also many others that come up naturally e.g. in engineering and finance.

2.1 Dual spaces of random variables

Let $\mathcal{U}$ and $\mathcal{Y}$ be linear spaces of $\mathbb{R}^m$-valued random variables in separating duality under the bilinear form

$$\langle u, y\rangle := E[u \cdot y].$$

This means that $u \cdot y \in L^1$ for all $u \in \mathcal{U}$ and $y \in \mathcal{Y}$ and that for every nonzero $u \in \mathcal{U}$, there exists a $y \in \mathcal{Y}$ such that $\langle u, y\rangle \neq 0$ and vice versa. As usual, we identify random variables that coincide almost surely so the elements of $\mathcal{U}$ and $\mathcal{Y}$ are actually equivalence classes of random variables that coincide almost surely. We will also assume that the spaces are *decomposable* and *solid*. Decomposability means that

$$1_A u + 1_{\Omega\setminus A} u' \in \mathcal{U}$$

for every $u \in \mathcal{U}$ and $u' \in L^\infty$ while solidity means that if $\bar{u} \in \mathcal{U}$ and $u \in L^0$ are such that $|u^i| \leq |\bar{u}^i|$ almost surely for every $i = 1, \dots, m$, then $u \in \mathcal{U}$; similarly for $\mathcal{Y}$. Solidity implies that

$$\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_m \quad \text{and} \quad \mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_m,$$

where $\mathcal{U}_i$ and $\mathcal{Y}_i$ are solid decomposable spaces of real-valued random variables in separating duality under the bilinear form $(u_i, y_i) \mapsto E[u_i y_i]$. In particular,

$$u_i y_i \in L^1 \quad \text{and} \quad \langle u, y\rangle = \sum_{i=1}^m E[u_i y_i] \quad \forall u \in \mathcal{U}, y \in \mathcal{Y}. \tag{1}$$

Given a solid space of real-valued random variables $\mathcal{U}_0$, the space $\{u \in L^0(\mathbb{R}^m) \mid |u| \in \mathcal{U}_0\}$ is solid and it can be written as $\mathcal{U}_0^m$, the m-fold Cartesian product of $\mathcal{U}_0$. A solid space containing all constant functions is decomposable. The following shows that the converse does not hold.

Counterexample 1. *Let $x \geq 1$ be an unbounded real-valued random variable and $\mathcal{X} := L^\infty + Lin(x1_A \mid A \in \mathcal{F})$. Then $\mathcal{X}$ is decomposable, by construction, but not solid, since it does not contain $\sqrt{x}$.*

Decomposable solid spaces of random variables in separating duality include Lebesgue spaces, Orlicz spaces, Marcinkiewich spaces paired with Lorentz spaces, spaces of finite moments $\|u\|_{L^p}$ for all $p \in (1, \infty)$ as well as the general class of Banach Function Spaces or, even more generally, locally convex function spaces; see [20] and its references. The spaces of continuous functions or various Sobolev spaces of functions on $\mathbb{R}^n$ fail to be decomposable or solid. The space L^0 of all random variables is decomposable and solid but if $(\Omega, \mathcal{F}, P)$ is atomless, it cannot be paired with a nontrivial space of random variables. Indeed, if $y \in L^0$ is nonzero, then there exists $\epsilon > 0$ and $A \in \mathcal{F}$ such that $|y|1_A > \epsilon$ and $P(A)$. Since the space is atomless, there exists $\eta < 0$ with $E[1_A\eta] = -\infty$. Choosing $u = 1_A y\eta$, we get $E[u \cdot y] = -\infty$.

Given a topology on $\mathcal{U}$, the corresponding topological dual of $\mathcal{U}$ is the linear space of all continuous linear functionals on $\mathcal{U}$. A topology is *compatible* with the bilinear form on $\mathcal{U} \times \mathcal{Y}$ if every continuous linear functional can be expressed in the form

$$u \mapsto \langle u, y\rangle$$

for some $y \in \mathcal{Y}$. Such topologies can be characterized in terms of the "weak" and "Mackey" topologies associated with the bilinear form. The *weak topology* $\sigma(\mathcal{U}, \mathcal{Y})$ on $\mathcal{U}$ is the topology generated by linear functionals $u \mapsto \langle u, y\rangle$ where $y \in \mathcal{Y}$. Similarly for $\mathcal{Y}$. The *Mackey topology* is the topology generated by the sublinear functionals

$$\sigma_D(u) := \sup_{y\in D}\langle u, y\rangle,$$

where $D \subset \mathcal{Y}$ is $\sigma(\mathcal{Y},\mathcal{U})$-compact. Similarly for $\mathcal{Y}$. Given a topology on $\mathcal{U}$, the corresponding topological dual can be identified with $\mathcal{Y}$ if and only if the topology is between $\sigma(\mathcal{U}, \mathcal{Y})$ and $\tau(\mathcal{U}, \mathcal{Y})$. If $\mathcal{U}$ is Fréchet (e.g. Banach) and $\mathcal{Y}$ is its topological dual, then the $\sigma(\mathcal{Y},\mathcal{U})$-compact sets are the bounded sets in $\mathcal{Y}$, so $\tau(\mathcal{U}, \mathcal{Y})$ is the strong topology; see [10].

The following is from [16].

Lemma 2. *We have $L^\infty \subseteq \mathcal{U} \subseteq L^1$ and $L^\infty \subseteq \mathcal{Y} \subseteq L^1$ and*

$$\begin{aligned}
\sigma(L^1, L^\infty)|_{\mathcal{U}} \subseteq \sigma(\mathcal{U}, \mathcal{Y}), &\quad \sigma(\mathcal{U}, \mathcal{Y})|_{L^\infty} \subseteq \sigma(L^\infty, L^1),\\
\tau(L^1, L^\infty)|_{\mathcal{U}} \subseteq \tau(\mathcal{U}, \mathcal{Y}), &\quad \tau(\mathcal{U}, \mathcal{Y})|_{L^\infty} \subseteq \tau(L^\infty, L^1).
\end{aligned}$$

The L^0-topology on $\mathcal{U}$ is weaker than $\tau(\mathcal{U}, \mathcal{Y})$.

Given a decomposable space $\mathcal{U}$, its *Köthe dual* is the linear space

$$\{y \in L^0 \mid u \cdot y \in L^1 \quad \forall u \in \mathcal{U}\}.$$

This is the largest space of random variables that can be paired with $\mathcal{U}$ with the bilinear form $(u, y) \mapsto E[u \cdot y]$. Clearly, the Köthe dual of a solid space is solid. The following is well-known, e.g., in Lebesgue and Orlicz spaces.

Lemma 3. *Let $\mathcal{U}$ and $\mathcal{Y}$ be solid and $\mathcal{G} \subset \mathcal{F}$ a σ-algebra such that $E^{\mathcal{G}}\mathcal{U} \subset \mathcal{U}$. The mapping $E^{\mathcal{G}} : \mathcal{U} \to \mathcal{U}$ is weakly continuous if and only if $E^{\mathcal{G}}\mathcal{Y} \subset \mathcal{Y}$ and in this case,*

$$\langle E^{\mathcal{G}}u, y\rangle = \langle u, E^{\mathcal{G}}y\rangle \quad \forall u \in \mathcal{U},\ y \in \mathcal{Y}.$$

If $\mathcal{Y}$ is the Köthe dual of $\mathcal{U}$, then $E^{\mathcal{G}}\mathcal{Y} \subset \mathcal{Y}$.

Proof. If u^i, y^i, $(E^{\mathcal{G}}u)^i y^i$ and $u^i(E^{\mathcal{G}}y)^i$ are integrable, Lemma 63 gives

$$E[E^{\mathcal{G}}u \cdot y] = E[(E^{\mathcal{G}}u) \cdot E^{\mathcal{G}}y] = E[u \cdot E^{\mathcal{G}}y]. \tag{2}$$

Thus, if $E^{\mathcal{G}}\mathcal{U} \subset \mathcal{U}$ and $E^{\mathcal{G}}\mathcal{Y} \subset \mathcal{Y}$, then, by (1), the function $u \mapsto E^{\mathcal{G}}u$ is weakly continuous. On the other hand, if $E^{\mathcal{G}} : \mathcal{U} \to \mathcal{U}$ is weakly continuous, then $u \mapsto E[E^{\mathcal{G}}u \cdot y]$ is $\sigma(\mathcal{U}, \mathcal{Y})$-continuous for $y \in \mathcal{Y}$. Thus, there exists a $y' \in \mathcal{Y}$ such that $E[E^{\mathcal{G}}u \cdot y] = E[u \cdot y']$ for all $u \in \mathcal{U}$. Since $y \in L^1$, (2) gives

$$E[E^{\mathcal{G}}u \cdot y] = E[u \cdot E^{\mathcal{G}}y] \quad \forall u \in L^\infty.$$

Thus, $y' = E^{\mathcal{G}}y$ almost surely.

Assume now that $\mathcal{Y}$ is the Köthe dual of $\mathcal{U}$ and let $y \in \mathcal{Y}$. It suffices to show $E^{\mathcal{G}}y \in \mathcal{Y}$. By solidity and linearity, we may assume that at most one component y^i of y is nonzero and that it is nonnegative. Then $E^{\mathcal{G}}y^i$ is nonnegative. Since $\mathcal{Y}$ is the Köthe dual, it suffices to show that $E[u^i(E^{\mathcal{G}}y^i)] < \infty$ for every nonnegative $u \in \mathcal{U}$. By Lemma 63, $E[u^i(E^{\mathcal{G}}y^i)] = E[E^{\mathcal{G}}(u^i)y^i]$, where the right side is finite, since $E^{\mathcal{G}}\mathcal{U} \subset \mathcal{U}$. $\square$

Let $\mathcal{X}$ and $\mathcal{V}$ be decomposable solid spaces of $\mathbb{R}^n$-valued random variables in separating duality under the bilinear form

$$(x, v) \mapsto E[x \cdot v].$$

A linear mapping $\mathcal{A} : \mathcal{X} \to \mathcal{U}$ is *weakly continuous* if it is continuous with respect to the weak topologies. This means that $x \mapsto \langle \mathcal{A}x, y\rangle$ is $\sigma(\mathcal{X}, \mathcal{V})$-continuous for all $y \in \mathcal{Y}$, or equivalently, there exists a linear mapping $\mathcal{A}^* : \mathcal{Y} \to \mathcal{V}$ such that

$$\langle \mathcal{A}x, y\rangle = \langle x, \mathcal{A}^*y\rangle \quad \forall x \in \mathcal{X},\ y \in \mathcal{Y}.$$

The mapping $\mathcal{A}^*$ is known as the *adjoint* of $\mathcal{A}$.

Lemma 4. *Let $A \in L^0(\mathbb{R}^{m\times n})$ be a random matrix such that $Ax \in \mathcal{U}$ for all $x \in \mathcal{X}$. The linear mapping $\mathcal{A} : \mathcal{X} \to \mathcal{U}$ defined pointwise by*

$$\mathcal{A}x = Ax \quad a.s.$$

is weakly continuous if and only if $A^*y \in \mathcal{V}$ for all $y \in \mathcal{Y}$, and in this case its adjoint is given pointwise by

$$\mathcal{A}^*y = A^*y \quad a.s.$$

If $\mathcal{V}$ is the Köthe dual of $\mathcal{X}$, then $A^*y \in \mathcal{V}$ for all $y \in \mathcal{Y}$,

Proof. For any $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

$$\langle \mathcal{A}x, y\rangle = E[(Ax) \cdot y] = E[x \cdot A^*y],$$

which proves the equivalence and the adjoint formula. The above equation implies that $x \cdot A^*y \in L^1$, so $A^*y \in \mathcal{V}$ when $\mathcal{V}$ is the Köthe dual of $\mathcal{X}$. □

2.2 Conjugates of integral functionals

This section studies convex integral functionals on paired decomposable spaces $\mathcal{U}$ and $\mathcal{Y}$ of random variables. More precisely, we take a normal integrand h and study the integral functionals $Eh : \mathcal{U} \to \overline{\mathbb{R}}$ and $Eh^* : \mathcal{Y} \to \overline{\mathbb{R}}$ defined by

$$Eh(u) := \int_\Omega h(u(\omega), \omega)dP(\omega)$$

and

$$Eh^*(y) := \int_\Omega h^*(y(\omega), \omega)dP(\omega).$$

The following two theorems are essentially reformulations of the main results in [25]. We give the simple proofs for completeness.

Theorem 5. *If h is a convex normal integrand with $\operatorname{dom} Eh \neq \emptyset$, then*

$$(Eh)^* = Eh^*.$$

Moreover, $y \in \partial Eh(u)$ if and only if $Eh(u)$ is finite and $y \in \partial h(u)$ almost surely.

Proof. The first claim follows by applying [31, Theorem 14.60] to the normal integrand $h_y(u, \omega) := h(u, \omega) - u \cdot y(\omega)$, where $y \in \mathcal{Y}$. As to the second, we have $y \in \partial Eh(u)$ if and only if $Eh(u)$ is finite and $Eh(u) + (Eh)^*(y) = \langle u, y\rangle$. Since $(Eh)^* = Eh^*$ by the first part, the equality holds, by Fenchel's inequality, if and only if

$$h(u) + h^*(y) = u \cdot y$$

almost surely. This means that $y \in \partial h(u)$ almost surely. □

Corollary 6. *Let h be a convex normal integrand. The following are equivalent*

1. $\operatorname{dom} Eh \neq \emptyset$ *and* $\operatorname{dom} Eh^* \neq \emptyset$,
2. *Eh is proper and closed,*

3. $\operatorname{dom} Eh \neq \emptyset$ *and there exists* $y \in \mathcal{Y}$ *and* $\alpha \in L^1$ *such that*

$$h(u,\omega) \geq u \cdot y(\omega) - \alpha(\omega)$$

and imply that Eh *and* Eh^* *are conjugates of each other and that* $y \in \partial Eh(u)$ *if and only if* $y \in \partial h(u)$ *almost surely.*

Proof. By Theorem 5, 1 implies 2. Assuming 2, there exists $u \in \operatorname{dom} Eh$, $y \in \mathcal{Y}$ and $a \in \mathbb{R}$ such that

$$Eh(u) \geq \langle u, y\rangle - a \quad \forall u \in \mathcal{U}$$

Thus, by Theorem 5,

$$a \geq (Eh)^*(y) = Eh^*(y).$$

By Fenchel's inequality

$$h(u,\omega) + h^*(y,\omega) \geq u \cdot y,$$

so 3 holds with $\alpha(\omega) = h^*(y(\omega),\omega)$. If 3 holds, $Eh^*(y) \leq E\alpha$, so 1 holds. By Theorem 5, Eh and Eh^* are conjugates of each other and $y \in \partial Eh(u)$ implies $y \in \partial h(u)$ almost surely. If $y \in \partial h(u)$, then $h(u) + h^*(y) = u \cdot y$ almost surely, where each summand is integrable, since Eh and Eh^* are proper, so $Eh(u) + Eh^*(y) = \langle u, y\rangle$, which means that $y \in \partial Eh(u)$. □

Corollary 7. *Given a closed convex valued measurable mapping* $S : \Omega \rightrightarrows \mathbb{R}^m$, *the set*

$$\mathcal{S} := \{u \in \mathcal{U} \mid u \in S \text{ a.s.}\}$$

is closed and convex.

Proof. This follows by applying Theorem 5 to the conjugate of $h(u,\omega) := \delta_S(u,\omega)$. □

By symmetry, one can add obvious dual versions of 2 and 3 in the list of equivalent conditions in Corollary 6. The following gives a general form of the classical Jensen's inequality for conditional expectations.

Theorem 8 (Jensen's inequality)**.** *Assume that* $\mathcal{U}$ *and* $\mathcal{Y}$ *are solid with* $E^{\mathcal{G}}\mathcal{U} \subset \mathcal{U}$ *and* $E^{\mathcal{G}}\mathcal{Y} \subset \mathcal{Y}$ *and let* h *be a* $\mathcal{G}$*-measurable convex normal integrand such that* Eh^* *is proper on* $\mathcal{Y}$*. Then*

$$Eh(E^{\mathcal{G}}u) \leq Eh(u)$$

for every $u \in \mathcal{U}$*.*

Proof. Assume first that Eh^* is proper on $\mathcal{Y} \cap L^0(\mathcal{G})$. Corollary 6 then gives

$$\begin{aligned}
Eh(E^{\mathcal{G}}u) &= \sup_{y \in \mathcal{Y} \cap L^0(\mathcal{G})} \{E[(E^{\mathcal{G}}u) \cdot y] - Eh^*(y)\} \\
&= \sup_{y \in \mathcal{Y} \cap L^0(\mathcal{G})} E[u \cdot y - h^*(y)] \\
&\leq E \sup_{y \in \mathbb{R}^m} \{u \cdot y - h^*(y)\} \\
&= Eh(u)
\end{aligned}$$

for any $u \in \mathcal{U}$. If Eh^* is proper merely on $\mathcal{Y}$, then $E(h^*)^+$ is proper on $\mathcal{Y}$ as well. The function $E[(h^*)^+]^*$ is finite at the origin so, by the first part of the proof, $E(h^*)^+(E^{\mathcal{G}}y) \le E(h^*)^+(y)$, so Eh^* is finite on $\mathcal{Y} \cap L^0(\mathcal{G})$. $\square$

3 Duality for integrable strategies

We will develop a duality theory for (SP) by applying the general conjugate duality framework of Rockafellar [27] first to the parametric optimization problem

$$\text{minimize} \qquad Ef(x,\bar{u}) := \int f(x(\omega),\bar{u}(\omega),\omega)dP(\omega) \quad \text{over } x \in \mathcal{X}_a \qquad (SP_{\mathcal{X}})$$

where $\mathcal{X} \subset L^0(\Omega,\mathcal{F},P;\mathbb{R}^n)$ is a solid decomposable space of random paths and

$$\mathcal{X}_a := \mathcal{X} \cap \mathcal{N}.$$

We will assume that the parameter $\bar{u}$ belongs to another solid decomposable space $\mathcal{U} \subset L^0(\Omega,\mathcal{F},P;\mathbb{R}^m)$ of random variables. The general theory of convex duality will give a dual problem and optimality conditions for $(SP_{\mathcal{X}})$. Section 4 will then extend these results to the original problem (SP) where we optimize over general adapted strategies in L^0.

We embed $(SP_{\mathcal{X}})$ into the conjugate duality framework by introducing an additional parameter $z \in \mathcal{X}$ and the extended real-valued convex function F on $\mathcal{X} \times \mathcal{X} \times \mathcal{U}$ defined by

$$F(x,z,u) := Ef(x,u) + \delta_{\mathcal{N}}(x-z).$$

We denote the associated optimum value function by

$$\varphi(z,u) := \inf_{x\in\mathcal{X}}\{Ef(x,u) \,|\, x-z \in \mathcal{N}\}.$$

We assume that $\mathcal{X}$ is in separating duality with a solid decomposable space $\mathcal{V} \subset L^0(\Omega,\mathcal{F},P;\mathbb{R}^n)$ and that $\mathcal{U}$ is in separating duality with a solid decomposable space $\mathcal{Y} \subset L^0(\Omega,\mathcal{F},P;\mathbb{R}^m)$. The bilinear forms are the usual ones, i.e.

$$\langle x,v\rangle := E[x\cdot v] \quad \text{and} \quad \langle u,y\rangle := E[u\cdot y].$$

Solidity implies that

$$\mathcal{X} = \mathcal{X}_0 \times \cdots \times \mathcal{X}_T \quad \text{and} \quad \mathcal{V} = \mathcal{V}_0 \times \cdots \times \mathcal{V}_T,$$

where $\mathcal{X}_t$ and $\mathcal{V}_t$ are solid decomposable spaces of $\mathbb{R}^{n_t}$-valued random variables in separating duality under the bilinear form $(x_t,v_t) \mapsto E[x_t \cdot v_t]$. It follows that

$$\langle x,v\rangle = \sum_{t=0}^T E[x_t\cdot v_t] \quad \forall x\in\mathcal{X},\ v\in\mathcal{V}$$

and

$$\mathcal{X}_a = \mathcal{X}_0(\mathcal{F}_0) \times \cdots \times \mathcal{X}_T(\mathcal{F}_T).$$

According to the general conjugate duality framework of [27], the *dual problem* is the concave maximization problem

$$\text{maximize} \quad \langle \bar{u}, y \rangle - F^*(0, p, y) \quad \text{over } (p, y) \in \mathcal{V} \times \mathcal{Y}. \tag{D}$$

More explicit forms will be given below. By definition, $\varphi^*(p, y) = F^*(0, p, y)$, so the dual problem can be written as

$$\text{maximize} \quad \langle \bar{u}, y \rangle - \varphi^*(p, y) \quad \text{over } (p, y) \in \mathcal{V} \times \mathcal{Y}.$$

By Fenchel's inequality,

$$F(x, 0, u) \geq \langle u, y \rangle - F^*(0, p, y) \quad \forall x \in \mathcal{X}, u \in \mathcal{U}, p \in \mathcal{V}, y \in \mathcal{Y}.$$

Denoting the optimal values of primal and dual problem, respectively, as $\inf(SP_{\mathcal{X}})$ and $\sup(D)$, we thus have

$$\inf\,(SP_{\mathcal{X}}) \geq \sup\,(D)$$

A *duality gap* is said to exist if the inequality is strict. Conversely, we say that there is no duality gap if the above holds as an equality.

The associated *Lagrangian* is the convex-concave function L on $\mathcal{X} \times \mathcal{V} \times \mathcal{Y}$ given by

$$L(x, p, y) := \inf_{(z,u) \in \mathcal{V} \times \mathcal{U}} \{F(x, z, u) - \langle z, p \rangle - \langle u, y \rangle\}.$$

By definition, the conjugate of F can be expressed as

$$F^*(v, p, y) = \sup_{x \in \mathcal{X}} \{\langle x, v \rangle - L(x, p, y)\}.$$

The associated minimax problem is to find a saddle-value and/or a saddle-point of the concave-convex function

$$L_{\bar{u}}(x, p, y) := L(x, p, y) + \langle \bar{u}, y \rangle,$$

when minimizing over x and maximizing over (p, y). If

$$\inf_x \sup_{p,y} L_{\bar{u}}(x, p, y) = \sup_{p,y} \inf_x L_{\bar{u}}(x, p, y),$$

the common value is called the *minimax* or the *saddle-value*, and (x, p, y) is called a *saddle-point* if

$$L_{\bar{u}}(x, p', y') \leq L_{\bar{u}}(x, p, y) \leq L_{\bar{u}}(x', p, y) \quad \forall x', p', y'.$$

Existence of a saddle-point implies the existence of a saddle-value.

Since $\mathcal{N}$ is closed in probability, Lemma 2 gives the following.

Lemma 9. *$\mathcal{X}_a$ is $\sigma(\mathcal{X}, \mathcal{V})$-closed.*

The following three theorems are restatements of the main duality results in [27] in the present setting. They all involve the assumption that the integral functional Ef be closed in u. This means that $Ef(x, \cdot)$ is closed in $\mathcal{U}$ for each $x \in \mathcal{X}$. Combined with Lemma 9, this implies that the function F is closed in (z, u).

The following characterizes the absence of duality gap.

Theorem 10. *The following are equivalent,*

1. $\inf(SP_{\mathcal{X}}) = \sup(D)$,
2. *φ is closed at $(0, \bar{u})$.*

If Ef is closed in u, the above are equivalent to

3. *The function $L_{\bar{u}}$ has a saddle-value.*

The next one characterizes situations where there is no duality gap and, furthermore, the dual admits solutions.

Theorem 11. *If $\varphi(0, u) < \infty$, the following are equivalent*

1. *(p, y) solves (D) and $\inf(SP_{\mathcal{X}}) = \sup(D)$.*
2. *either $\varphi(0, \bar{u}) = -\infty$ or $(p, y) \in \partial\varphi(0, \bar{u})$,*

If Ef is closed in u, the above are equivalent to

3. $\inf_x \sup_{p,y} L_{\bar{u}}(x, p, y) = \inf_x L_{\bar{u}}(x, p, y)$.

The following characterizes the situations where both primal and dual solutions exist and there is no duality gap.

Theorem 12. *The following are equivalent,*

1. *x solves $(SP_{\mathcal{X}})$, (p, y) solves (D) and $\inf(SP_{\mathcal{X}}) = \sup(D) \in \mathbb{R}$,*
2. $(0, p, y) \in \partial F(x, 0, \bar{u})$.

If Ef is closed in u, the above are equivalent to

3. *$0 \in \partial_x L(x, p, y)$ and $(0, \bar{u}) \in \partial_{(p,y)}[-L](x, p, y)$.*

In order to write the dual problem and the optimality conditions more explicitly in terms of the problem data, we will first derive explicit expressions for F^* and φ^*. The rest of the section will then focus on the Lagrangian, the associated minimax problem and optimality conditions. We will denote the orthogonal complement of $\mathcal{X}_a$ by

$$\mathcal{X}_a^{\perp} := \{v \in \mathcal{V} \mid \langle x, v \rangle = 0 \quad \forall x \in \mathcal{X}_a\}.$$

Lemma 13. *If* $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$*, then*

$$F^*(v,p,y) = Ef^*(v+p,y) + \delta_{\mathcal{X}_a^\perp}(p),$$

and, in particular,

$$\varphi^*(p,y) = Ef^*(p,y) + \delta_{\mathcal{X}_a^\perp}(p).$$

If, in addition, $\operatorname{dom} Ef^* \cap (\mathcal{V} \times \mathcal{Y}) \neq \emptyset$*, then* F *is proper and closed.*

Proof. By the interchange rule [31, Theorem 14.60],

$$\begin{aligned} F^*(v,p,y) &= \sup_{x\in\mathcal{X},z\in\mathcal{X},u\in\mathcal{U}} \{\langle x,v\rangle + \langle z,p\rangle + \langle u,y\rangle - Ef(x,u) \,|\, x-z \in \mathcal{X}_a\} \\ &= \sup_{x\in\mathcal{X},z'\in\mathcal{X},u\in\mathcal{U}} \{E[x\cdot(v+p) + u\cdot y - f(x,u) - z'\cdot p] \,|\, z' \in \mathcal{X}_a\} \\ &= Ef^*(v+p,y) + \delta_{\mathcal{X}_a^\perp}(p). \end{aligned}$$

When $\operatorname{dom} Ef^* \neq \emptyset$, Ef is proper and closed, by Corollary 6, so F is closed as a sum of proper and closed functions. Clearly, F is proper. □

As an immediate corollary, we get the following.

Theorem 14. *If* $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$*, the dual problem* (D) *can be written as*

$$\text{maximize} \quad \langle \bar{u}, y\rangle - Ef^*(p,y) \quad \text{over} \quad (p,y) \in \mathcal{X}_a^\perp \times \mathcal{Y}. \qquad (D)$$

Much as in [28, Section 4] and [23, Proposition 1], the orthogonal complement of $\mathcal{X}_a$ can be expressed in terms of the set

$$\mathcal{N}^\perp := \{v \in L^1 \mid \langle x,v\rangle = 0 \quad \forall x \in \mathcal{N} \cap L^\infty\}.$$

Lemma 15. *The set* $\mathcal{X}_a$ *is* $\sigma(\mathcal{X},\mathcal{V})$*-closed and*

$$\mathcal{X}_a^\perp = \mathcal{N}^\perp \cap \mathcal{V} = \{v \in \mathcal{V} \,|\, E_t v_t = 0 \quad t = 0,\dots,T\}.$$

Proof. Since $\mathcal{N}$ $\mathcal{N}$ is closed in L^0, Lemma 2 implies that $\mathcal{X}_a$ is closed in $\tau(\mathcal{X},\mathcal{V})$ and thus, by convexity, also in $\sigma(\mathcal{X},\mathcal{V})$. Since

$$\mathcal{X}_a = \mathcal{X}_0(\mathcal{F}_0) \times \cdots \times \mathcal{X}_T(\mathcal{F}_T),$$

we have $v \in \mathcal{X}_a^\perp$ if and only if $E[x_t \cdot v_t] = 0$ for every $x_t \in \mathcal{X}_t(\mathcal{F}_t)$. Here, $E[x_t \cdot v_t] = E[x_t \cdot (E_t v_t)]$, by Lemma 63. □

Note that the dual objective can be written also as

$$\langle \bar{u}, y\rangle - Ef^*(p,y) = E \inf_{(x,u)\in\mathbb{R}^n\times\mathbb{R}^m} [f(x,u) - x\cdot p + (\bar{u}-u)\cdot y].$$

This is the optimum value in a relaxed version of the primal problem (SP) where we are now allowed to optimize over both x and u and the information constraint $x \in \mathcal{N}$ has been removed so the minimization can be done. The constraints have been replaced by linear penalties given by the dual variables p and y. The optimum value of (D) is less than or equal to that of $(SP_{\mathcal{X}})$. If the value function φ is closed at $(0, \bar{u})$, then, by Theorem 10, the optimum values can be made arbitrarily close by an appropriate choice of (p, y). If $(p, y) \in \partial\varphi(0, \bar{u})$, then, by Theorem 11, there is no duality gap and (p, y) solves the dual. This implies, in particular, that p is a subgradient of φ with respect to the first argument at $(0, \bar{u})$, i.e.,

$$Ef(x + z, \bar{u}) - \langle z, p\rangle \geq \varphi(0, \bar{u}) \quad \forall x \in \mathcal{X}_a, z \in \mathcal{X}.$$

In other words, one cannot improve the optimum value of $(SP_{\mathcal{X}})$ by adding a nonadapted perturbation z to the strategy x when one has to pay $\langle z, p\rangle$. Such an element $p \in \mathcal{X}_a^\perp$ is known as a *shadow price of information* of $(SP_{\mathcal{X}})$. In the deterministic setting, $\mathcal{X}_a^\perp = \{0\}$ so the dual problem becomes

$$\text{maximize} \quad \bar{u} \cdot y - f^*(0, y) \quad \text{over} \quad y \in \mathbb{R}^m$$

and we recover the classical duality framework in finite-dimensional spaces.

Theorem 14 can be used to restate Theorems 10, 11, and 12 more explicitly. In particular, the first two equivalences in Theorem 12 can be restated as follows.

Theorem 16. *If $(SP_{\mathcal{X}})$ and (D) are feasible, then the following are equivalent*

1. *x solves $(SP_{\mathcal{X}})$, (p, y) solves (D) and* $\inf(SP_{\mathcal{X}}) = \sup(D)$,
2. *$x \in \mathcal{X}_a$, $(p, y) \in \mathcal{X}_a^\perp \times \mathcal{Y}$ and*

$$(p, y) \in \partial f(x, \bar{u}) \quad \textit{a.s.}$$

Proof. By Theorem 12, 1 is equivalent to $(0, p, y) \in \partial F(x, 0, \bar{u})$ which means that $F(x, 0, \bar{u}) + F^*(0, p, y) = \langle \bar{u}, y\rangle$. By Lemma 13, this means that $x \in \mathcal{X}_a$, $p \in \mathcal{X}_a^\perp$ and

$$Ef(x, \bar{u}) + Ef^*(p, y) = E[x \cdot p] + E[\bar{u} \cdot y]. \tag{3}$$

Given $(x', u') \in \mathcal{X} \times \mathcal{U}$ and $(p', y') \in \mathcal{V} \times \mathcal{Y}$, we have

$$f(x', u') + f^*(p', y') \geq x' \cdot p' + u \cdot y', \tag{4}$$

by Fenchel's inequality, so the feasibility assumptions imply that the negative parts of $f(x', u')$ and $f^*(p', y')$ are integrable and thus, by Lemma 62,

$$Ef(x', u') + Ef^*(p', y') = E[f(x', u') + f^*(p', y')].$$

Thus, (3) means that $(x, \bar{u})$ and (p, y) satisfy (4) as an equality, i.e., $(p, y) \in \partial f(x, \bar{u})$. $\square$

Note that the dual is feasible e.g. if F is bounded from below since then $F^*(0, 0)$ is finite. If $\partial\varphi(0, \bar{u})$ is nonempty, then by Theorem 11, there is no duality gap and a dual has a solution. Theorem 16 thus implies the following.

Corollary 17. *If $\partial\varphi(0,\bar{u}) \neq \emptyset$, then* $\inf(SP_{\mathcal{X}}) = \sup(D)$ *and the following are equivalent for an $x \in \mathcal{X}_a$,*

1. *x solves $(SP_{\mathcal{X}})$,*
2. *there exists $(p,y) \in \mathcal{X}_a^\perp \times \mathcal{Y}$ with*

$$(p,y) \in \partial f(x,\bar{u}) \quad a.s.$$

The rest of the section focuses on the Lagrangian L and the associated minimax problem. The Lagrangian L itself has a somewhat cumbersome expression but it turns out that it is "equivalent" to a simpler function that has the same saddle-value and saddle-points. The expressions derived below, involve the *Lagrangian integrand* $l : \mathbb{R}^n \times \mathbb{R}^m \times \Omega \to \overline{\mathbb{R}}$ defined by

$$l(x,y,\omega) := \inf_{u\in\mathbb{R}^m} \{f(x,u,\omega) - u\cdot y\}.$$

For any (x,y,ω), the function $l(\cdot,y,\omega)$ is convex and $l(x,\cdot,\omega)$ is upper semicontinuous and concave. Clearly,

$$f^*(v,y,\omega) = \sup_{x\in\mathbb{R}^n} \{x\cdot v - l(x,y,\omega)\}$$

so, by the biconjugate theorem,

$$(\mathrm{cl}_x\, l)(x,y,\omega) = \sup_{v\in\mathbb{R}^n} \{x\cdot v - f^*(v,y,\omega)\}.$$

Given $x\in\mathcal{X}$ and $y\in\mathcal{Y}$, the functions

$$(y,\omega) \mapsto -l(x(\omega),y,\omega) = \sup_{u\in\mathbb{R}^m} \{u\cdot y - f(x(\omega),u,\omega)\}$$

and

$$(x,\omega) \mapsto (\mathrm{cl}_x\, l)(x,y(\omega),\omega) = \sup_{v\in\mathbb{R}^n} \{x\cdot v - f^*(v,y(\omega),\omega)\}$$

are normal integrands, by Proposition 14.45 and Theorem 14.50 of [31]. In particular, the functions

$$\omega \mapsto l(x(\omega),y(\omega),\omega) \quad \text{and} \quad \omega \mapsto (\mathrm{cl}_x\, l)(x(\omega),y(\omega),\omega)$$

are measurable, by [31, Proposition 14.28],

We will denote the projection of $\mathrm{dom}\, Ef$ to the x component by

$$\mathrm{dom}_x\, Ef := \{x\in\mathcal{X} \mid \exists u\in\mathcal{U} :\ Ef(x,u) < \infty\}$$

and the projection of $\mathrm{dom}\, Ef^*$ to the y component by

$$\mathrm{dom}_y\, Ef^* := \{y\in\mathcal{Y} \mid \exists v\in\mathcal{U} :\ Ef^*(v,y) < \infty\}.$$

Lemma 18. *We have*

$$L(x,p,y) = \begin{cases} +\infty & \text{if } x \notin \operatorname{dom}_x Ef, \\ El(x,y) - \langle x,p\rangle & \text{if } x \in \operatorname{dom}_x Ef \text{ and } p \in \mathcal{X}_a^\perp, \\ -\infty & \text{otherwise.} \end{cases}$$

If $\operatorname{dom} Ef \neq \emptyset$*, then*

$$(\operatorname{cl}_x L)(x,p,y) = \begin{cases} E(\operatorname{cl}_x l)(x,y) - \langle x,p\rangle & \text{if } y \in \operatorname{dom}_y Ef^* \text{ and } p \in \mathcal{X}_a^\perp, \\ -\infty & \text{otherwise.} \end{cases}$$

If $\operatorname{dom} Ef \neq \emptyset$ *and* $\operatorname{dom} Ef^* \neq \emptyset$*, then all convex-concave functions between* L *and* $\operatorname{cl}_x L$ *have the same saddle-value, saddle-points and subdifferentials. In this case,*

$$v \in \partial_x L(x,p,y), \quad (z,u) \in \partial_{p,y}[-L](x,p,y)$$

if and only if $x - z \in \mathcal{X}_a$*,* $p \in \mathcal{X}_a^\perp$ *and*

$$p + v \in \partial_x l(x,y), \quad u \in \partial_y[-l](x,y) \quad a.s.$$

Proof. By definition,

$$\begin{aligned} L(x,p,y) &= \inf_{(z,u)\in\mathcal{X}\times\mathcal{U}} \{F(x,z,u) - \langle z,p\rangle - \langle u,y\rangle\} \\ &= \inf_{(z,u)\in\mathcal{X}\times\mathcal{U}} \{E[f(x,u) - z\cdot p - u\cdot y] \mid x - z \in \mathcal{X}_a\} \\ &= \inf_{(z',u)\in\mathcal{X}\times\mathcal{U}} \{E[f(x,u) - (x-z')\cdot p - u\cdot y] \mid z' \in \mathcal{X}_a\}, \end{aligned}$$

so the expression for L follows from [31, Theorem 14.60]. By Lemma 13,

$$\begin{aligned} (\operatorname{cl}_x L)(x,p,y) &= \sup_{v\in\mathcal{V}} \{\langle x,v\rangle - F^*(v,p,y)\} \\ &= \begin{cases} \sup_{v\in\mathcal{V}}\{\langle x,v\rangle - Ef^*(v+p,y)\} & \text{if } p \in \mathcal{X}_a^\perp, \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

so the expression for $\operatorname{cl}_x L$ follows from [31, Theorem 14.60] again. When $\operatorname{dom} Ef \neq \emptyset$ and $\operatorname{dom} Ef^* \neq \emptyset$, the function F is proper and closed by Lemma 13, so the saddle-values, saddle-points and subdifferentials of L and $\operatorname{cl}_x L$ coincide, by [32, Theorem 2 and 7].

When $\operatorname{dom} Ef \neq \emptyset$ and $\operatorname{dom} Ef^* \neq \emptyset$, F is closed, by Corollary 6, and then, $(v,z,u) \in \partial L(x,p,y)$ if and only if $(v,p,y) \in \partial F(x,z,u)$. By Lemma 13, this means that $x - z \in \mathcal{X}_a$, $p \in \mathcal{X}_a^\perp$ and

$$Ef(x,u) + Ef^*(v+p,y) = E[x\cdot v] + E[z\cdot p] + E[u\cdot y]$$

or, equivalently,

$$Ef(x,u) + Ef^*(v+p,y) = E[x\cdot(v+p)] + E[u\cdot y]$$

Since, by Fenchel's inequality,

$$f(x,u,\omega)+f^*(v+p,y,\omega)\geq x\cdot(v+p)+u\cdot y,$$

this means that $(v+p,y)\in\partial f(x,u)$ almost surely. Since f is closed, this is equivalent to $v+p\in\partial_x l(x,y)$ and $u\in\partial_y[-l](x,y)$. □

Corollary 19. *If $(SP_{\mathcal{X}})$ and (D) are feasible, the following are equivalent with the conditions in Theorem 16,*

1. *(x,p,y) is a saddle-point of*

$$(x,p,y)\mapsto El(x,y)-\langle x,p\rangle+\langle\bar{u},y\rangle,$$

when minimizing over $x\in\mathcal{X}$ and maximizing over $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$,

2. *$x\in\mathcal{X}_a$, $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$ and*

$$p\in\partial_x l(x,y),\quad \bar{u}\in\partial_y[-l](x,y)\quad a.s.$$

Proof. The claim follows from Theorem 12 and Lemma 18, since the convex-concave function in the first condition lies between L and $\operatorname{cl}_x L$. □

Similarly, we can augment Corollary 17 as follows.

Corollary 20. *If $\partial\varphi(0,\bar{u})\neq\emptyset$, then the following are equivalent,*

1. *x solves $(SP_{\mathcal{X}})$,*
2. *$x\in\mathcal{X}_a$ and there exists $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$ with*

$$(p,y)\in\partial f(x,\bar{u})\quad a.s.$$

3. *$x\in\mathcal{X}_a$ and there exists $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$ with*

$$p\in\partial_x l(x,y),\quad \bar{u}\in\partial_y[-l](x,y)\quad a.s.$$

In the deterministic setting, $\mathcal{X}_a^\perp=\{0\}$ so condition 3 in Corollary 20 becomes the Karush-Kuhn-Tucker (KKT) condition in finite-dimensional convex optimization. In the stochastic setting, the shadow price of information $p\in\mathcal{X}_a^\perp$ allows us to write the KKT-conditions scenariowise.

4 Duality for (SP)

While problem $(SP_{\mathcal{X}})$ in Section 3 allows for a convenient dualization within the purely functional analytic conjugate duality framework, there are interesting applications where $\inf(SP_{\mathcal{X}})>\inf(SP)$ or the infimum in (SP) is attained in L^0 but not in $\mathcal{X}$; see Counterexample 28 for a simple illustration. It may even happen that $(SP_{\mathcal{X}})$ is infeasible while (SP) is not; see Counterexample 29. This section shows that many of the duality relations between $(SP_{\mathcal{X}})$ and (D) derived in Section 3 also hold between (SP) and (D).

The function

$$\bar{\varphi}(z,u) := \inf_{x\in L^0}\{Ef(x,u) \mid x-z\in\mathcal{N}\}$$

on $\mathcal{X}\times\mathcal{U}$ gives the optimum value of (SP) when we perturb the strategies x by $z\in\mathcal{X}$ and vary the parameter $\bar{u}$ in the space $\mathcal{U}$. In particular, $\bar{\varphi}(0,\bar{u}) = \inf(SP)$. Clearly, $\varphi\ge\bar{\varphi}$ since the latter is defined by optimizing over a larger class of strategies. However, under a mild condition, their conjugates coincide.

Lemma 21. *If* $\operatorname{dom} Ef\cap(\mathcal{X}\times\mathcal{U})\neq\emptyset$*, then* $\varphi^*=\bar{\varphi}^*$.

Proof. Since $\varphi\ge\bar{\varphi}$, we have $\varphi^*\le\bar{\varphi}^*$. To prove the converse, let $(p,y)\in\operatorname{dom}\varphi^*$. By Lemma 13,

$$\varphi^*(p,y) = Ef^*(p,y)+\delta_{\mathcal{X}_a^\perp}(p),$$

so $p\in\mathcal{X}_a^\perp$. By Fenchel's inequality,

$$Ef(x,u)+\delta_{\mathcal{N}}(x-z)+Ef^*(p,y)\ge E[(x-z)\cdot p]+E[z\cdot p]+E[u\cdot y]$$

for all $(x,z,u)\in L^0\times\mathcal{X}\times\mathcal{U}$, so Lemma 22 below implies

$$Ef(x,u)+\delta_{\mathcal{N}}(x-z)+Ef^*(p,y)\ge E[z\cdot p]+E[u\cdot y].$$

Thus $\bar{\varphi}(z,u)+\varphi^*(p,y)\ge\langle z,p\rangle+\langle u,y\rangle$ for all $(z,u)\in\mathcal{X}\times\mathcal{U}$, which means that $\bar{\varphi}^*(p,y)\le\varphi^*(p,y)$. □

The above proof used the following from [22]; see also [19].

Lemma 22. *Let* $x\in\mathcal{N}$ *and* $v\in\mathcal{N}^\perp$*. If* $E[x\cdot v]^+\in L^1$*, then* $E[x\cdot v]=0$.

Note that, if $\operatorname{dom} Ef\cap(\mathcal{X}\times\mathcal{U})\neq\emptyset$ and φ^* is proper, then Ef is proper on $\mathcal{X}\times\mathcal{U}$.

Corollary 23. *We have* $\partial\varphi(z,u)=\partial\bar{\varphi}(z,u)$ *whenever the left side is nonempty. In particular, if* $\partial\varphi(0,\bar{u})\neq\emptyset$*, then*

$$\inf(SP)=\inf(SP_{\mathcal{X}})=\sup(D)$$

and the dual optimum is attained.

Proof. If $\partial\varphi(z,u)\neq\emptyset$, we have $(z,u)\in\operatorname{dom}\varphi$ and thus, $\operatorname{dom} Ef\cap(\mathcal{X}\times\mathcal{U})\neq\emptyset$. By the biconjugate theorem and Lemma 21, $\operatorname{cl}\varphi=\operatorname{cl}\bar{\varphi}$. In particular, $\varphi\ge\bar{\varphi}\ge\operatorname{cl}\varphi$. When $\partial\varphi(z,u)\neq\emptyset$, we have $\varphi(z,u)=\operatorname{cl}\varphi(z,u)$ so $\bar{\varphi}(z,u)=\varphi(z,u)$ and thus, $\bar{\varphi}(z,u)+\bar{\varphi}^*(p,y)=\langle x,p\rangle+\langle u,y\rangle$ if and only if $\varphi(z,u)+\varphi^*(p,y)=\langle x,p\rangle+\langle u,y\rangle$. In other words, $(p,y)\in\partial\bar{\varphi}(z,u)$ if and only if $(p,y)\in\partial\varphi(z,u)$. The second claim follows from Theorem 11 and Lemma 21 since subdifferentiability implies closedness. □

By Lemma 21,

$$\bar{\varphi}(0,\bar{u}) \geq \langle \bar{u}, y\rangle - \varphi^*(p,y) \quad \forall (p,y) \in \mathcal{V}\times\mathcal{Y},$$

where the right side is the dual objective from Section 3. Thus, the optimal value of (SP) is always bounded from below by the dual objective so the duality gap between (SP) and (D) is nonnegative. The duality gap is zero if and only if $\bar{\varphi}^{**}(0,\bar{u}) = \bar{\varphi}(0,\bar{u})$. Thus, we have the following, which gives the analogue of the first equivalence in Theorem 10.

Theorem 24. *If* $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$*, then the following are equivalent,*

1. $\inf (SP) = \sup (D)$,
2. $\bar{\varphi}$ *is closed at* $(0,\bar{u})$.

The following gives the analogue of the first equivalence in Theorem 11.

Theorem 25. *If* $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$ *and* $\bar{\varphi}(0,u) < \infty$*, then the following are equivalent*

1. (p,y) *solves* (D) *and* $\inf (SP) = \sup (D)$,
2. *either* $\bar{\varphi}(0,\bar{u}) = -\infty$ *or* $(p,y) \in \partial\bar{\varphi}(0,\bar{u})$,

Proof. Condition 2 means that either $\bar{\varphi}(0,\bar{u}) = -\infty$ or $\bar{\varphi}(0,z)+\bar{\varphi}^*(p,y) = \langle \bar{u}, y\rangle$, where, by Lemma 21, $\bar{\varphi}^* = \varphi^*$. Thus the claim follows from Lemma 13. □

Theorem 26. *If* $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$ *and* (SP) *and* (D) *are feasible, then following are equivalent*

1. x *solves* (SP), (p,y) *solves* (D) *and* $\inf (SP) = \sup (D)$,
2. x *is feasible in* (SP), (p,y) *is feasible in* (D) *and*
$$(p,y) \in \partial f(x,\bar{u}) \quad P\text{-a.s.} \tag{5}$$
3. x *is feasible in* (SP), (p,y) *is feasible in* (D) *and*
$$p \in \partial_x l(x,y), \quad \bar{u} \in \partial_y[-l](x,y) \quad P\text{-a.s.}$$

Proof. The assumptions imply that f is proper so the equivalence of 2 and 3 follows from [26, Theorem 37.5]. Let $x \in \mathcal{N}$ and $(p,y) \in \mathcal{V}\times\mathcal{Y}$ be feasible. By Fenchel's inequality,

$$f(x,u) + f^*(p,y) - \bar{u}\cdot y \geq x\cdot p \quad P\text{-a.s.}$$

so

$$Ef(x,u) + E[f^*(p,y) - \bar{u}\cdot y] \geq E[x\cdot p]$$

and one holds as an equality if and only if the other one does. Equality in the former means that 2 holds. By Lemma 22, $E[x\cdot p] = 0$, so equality in the latter means that 1 holds. □

If $\partial\bar\varphi(0,\bar u)\neq\emptyset$, then, by Theorem 25, $\inf(SP)=\sup(D)$ and the dual has a solution. Theorem 26 thus gives the following.

Corollary 27. *If* $\operatorname{dom} Ef\cap(\mathcal{X}\times\mathcal{U})\neq\emptyset$ *and* $\partial\bar\varphi(0,\bar u)\neq\emptyset$*, then* $\inf(SP)=\sup(D)$*, the optimum is attained in the dual and the following are equivalent,*

1. *x solves (SP),*
2. *x is feasible and there exists a dual feasible (p,y) with $(p,y)\in\partial f(x,\bar u)$ almost surely.*

If (SP) has a solution, (p,y) is dual optimal and $\inf(SP)=\sup(D)$, then, by Theorem 26, solutions of (SP) are scenariowise minimizers of the function

$$x\mapsto l(x,y)-x\cdot p$$

and, in particular, if the scenariowise minimizer is unique, then it is necessarily adapted and solves (SP).

Counterexample 28. *It is possible that* $\inf(SP)=\sup(D)$ *while* $\inf(SP_{\mathcal{X}})>\sup(D)$*. Indeed, let*

$$f(x,u,\omega)=(x_0-1)^2+\delta_{\{0\}}(x_0\xi(\omega)-x_1),$$

$\mathcal{F}_0$ be trivial and $\xi\in L^0(\mathcal{F}_1)$ with $\xi\notin\mathcal{X}$. Since f is nonnegative, $(1,\xi)$ is optimal for (SP) and the optimal value is zero. Here Ef is proper on $\mathcal{X}\times\mathcal{U}$, and, by a direct verification, $f^*(0,0)=0$, so the origin is a dual solution and $\inf(SP)=\sup(D)=0$*. On the other hand, the only feasible solution of $(SP_{\mathcal{X}})$ is the origin, so* $\inf(SP_{\mathcal{X}})=1$*.*

Counterexample 29. *It may happen that $(SP_{\mathcal{X}})$ is infeasible, but nevertheless, (SP) is feasible and* $\inf(SP)=\sup(D)$*. Indeed, let*

$$f(x,u,\omega)=\delta_{\{0\}}(x_T-u\xi(\omega)).$$

If $\xi\notin\mathcal{X}$ and $\bar u=1$, then (SP) is feasible while $(SP_{\mathcal{X}})$ is not. Clearly Ef is proper on $\mathcal{X}\times\mathcal{U}$ and $f^*(0,0)=0$, so $\inf(SP)=\sup(D)=0$*.*

The optimum value of the dual problem

$$\text{maximize}\quad \langle u,y\rangle-\varphi^*(p,y)\qquad\text{over}\quad (p,y)\in\mathcal{V}\times\mathcal{Y}\tag{D}$$

clearly coincides with that of

$$\text{maximize}\quad \langle u,y\rangle-g(y)\qquad\text{over}\quad (p,y)\in\mathcal{V}\times\mathcal{Y},\tag{rD}$$

where

$$g(y):=\inf_{p\in\mathcal{V}}\varphi^*(p,y).$$

Problem (rD) is called the *reduced dual problem.* A pair (p,y) solves (D) if and only if y solves (rD) and p attains the infimum in the definition of g. In many applications, the infimum and the minimizing p can be found analytically.

5 Absence of a duality gap

This section recalls the main result of [22] on the lower semicontinuity of the optimum value function of (SP). As we have seen, the lower semicontinuity implies the absence of a duality gap. Besides the lower semicontinuity, Theorem 31 below establishes the existence of optimal solutions to (SP).

Assumption 30. (SP) is feasible,

$$\{x \in \mathcal{N} \mid f^\infty(x,0) \le 0\}$$

is a linear space and there exists $p \in \mathcal{X}_a^\perp$ and $\epsilon > 0$ such that

$$\inf_{y \in \mathcal{Y}} Ef^*(\lambda p, y) < \infty$$

for $\lambda \in [1-\epsilon, 1+\epsilon]$.

The linearity condition in Assumption 30 holds trivially if $f(\cdot,0)$ is inf-compact since then, its recession function is strictly positive except at the origin; see [26, Theorem 8.6]. If f is bounded from below by an integrable random variable, then $Ef^*(0,0) < \infty$ so the second condition in Assumption 30 holds. The second condition holds also e.g. if $\operatorname{dom} Ef^* \cap (\mathcal{V} \times \mathcal{Y})$ is a nonempty cone. In certain models of financial mathematics, it is implied by the well-known asymptotic elasticity conditions on the utility function; see [19, Section 5.5].

Theorem 31. *Under Assumption* 30*, the function*

$$\bar{\varphi}(z,u) = \inf_{x \in L^0} \{Ef(x,u) \mid x - z \in \mathcal{N}\}$$

is lower semicontinuous on $\mathcal{X} \times \mathcal{U}$*,*

$$\bar{\varphi}^\infty(z,u) = \inf_{x \in L^0} E\{f^\infty(x,u) \mid z - z \in \mathcal{N}\},$$

and the infimums are attained for every $(z,u) \in \mathcal{X} \times \mathcal{U}$*.*

Proof. We have

$$\bar{\varphi}(z,u) = \inf_{x \in \mathcal{N}} Ef(x+z,u) = \inf_{x \in \mathcal{N}} E\bar{f}(x,z,u),$$

where $\bar{f}(x,z,u,\omega) = f(x+z,u,\omega)$. The claim thus follows from the main result of [22] as soon as

$$\{x \in \mathcal{N} \mid \bar{f}^\infty(x,0,0) \le 0\}$$

is linear and there exists $p \in \mathcal{X}_a^\perp$ and $\epsilon > 0$

$$\inf_{(p',y) \in \mathcal{V} \times \mathcal{Y}} E\bar{f}^*(\lambda p, p', y) < \infty.$$

Since $\bar{f}^\infty(x,z,u,\omega) = f^\infty(x+z,u,\omega)$, the former is clear from the linearity condition in Assumption 30. We have

$$\bar{f}^*(v,p,y,\omega) = f^*(v,y,\omega) + \delta_{\{0\}}(p-v)$$

so the latter follows from Assumption 30 as well. □

Combining the above with Theorem 24 gives the following.

Corollary 32. *Under Assumption* 30*,* $\inf(SP) = \sup(D)$*, and* (SP) *has a solution.*

6 Applications

This section applies the general duality result to specific instances of (SP). In the following applications, we give more explicit expressions for the involved functions and conditions but only give selected statements as examples of how the general theory can be applied.

6.1 Mathematical programming

Consider the problem

$$\begin{array}{llll} \text{minimize} & Ef_0(x) & \text{over} & x \in \mathcal{N}, \\ \text{subject to} & f_j(x) \le 0 & j = 1, \ldots, l \ a.s., & \\ & f_j(x) = 0 & j = l+1, \ldots, m \ a.s. & \end{array} \tag{MP}$$

where f_j are convex normal integrands with f_j affine for $j > l$. This fits the general duality framework with $\bar{u} = 0$ and

$$f(x,u,\omega) = \begin{cases} f_0(x,\omega) & \text{if } x \in \operatorname{dom} H,\ H(x) + u \in K, \\ +\infty & \text{otherwise,} \end{cases}$$

where $K = \mathbb{R}^l_- \times \{0\}$ and H is the K-convex random function defined by

$$\operatorname{dom} H(\cdot,\omega) = \bigcap_{j=1}^{m} \operatorname{dom} f_j(\cdot,\omega) \quad \text{and} \quad H(x,\omega) = (f_i(x,\omega))_{j=1}^m.$$

The Lagrangian integrand becomes

$$\begin{aligned} l(x,y,\omega) &= \inf\{f(x,u,\omega) - u \cdot y\} \\ &= \begin{cases} +\infty & \text{if } x \notin \operatorname{dom} H(\cdot,\omega), \\ f_0(x,\omega) + y \cdot H(x,\omega) & \text{if } x \in \operatorname{dom} H(\cdot,\omega) \text{ and } y \in K^*, \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

and the conjugate of f

$$\begin{aligned} f^*(p,y) &= \sup_{x \in \mathbb{R}^n} \{x \cdot p - l(x,y)\} \\ &= \begin{cases} \sup_{x \in \mathbb{R}^n} \{x \cdot p - f_0(x) - y \cdot H(x) \mid x \in \operatorname{dom} H(\cdot,\omega)\} & \text{if } y \in K^*, \\ +\infty & \text{if } y \notin K^*. \end{cases} \end{aligned}$$

If $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$, Lemma 13 says that the dual problem can be written as

$$\begin{array}{ll} \text{maximize} & E \inf_{x \in \mathbb{R}^n} \{f_0(x) + y \cdot H(x) - x \cdot p\} \ \text{ over } (p, y) \in \mathcal{X}_a^\perp \times \mathcal{Y} \\ \text{subject to} & \qquad y \in K^* \quad a.s. \end{array} \tag{D_{MP}}$$

To get more explicit expressions for f^* and the dual problem, additional structure is needed; see Example 36 below.

Recall that the *normal cone* of a convex set C at at point x is given by

$$N_C(x) := \{v \in \mathbb{R}^n \mid (x' - x) \cdot v \leq 0 \quad \forall x' \in C\}.$$

When C is a convex cone, then

$$v \in N_C(x) \quad \Leftrightarrow \quad x \in X,\ v \in C^*,\ x \cdot v = 0, \tag{6}$$

where $C^* := \{v \in \mathbb{R}^n \mid x \cdot v \leq 0\ \forall x \in C\}$ is the *polar cone* of C; see the end of [26, Section 23]. Theorem 26 gives the following.

Theorem 33. *If* $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$ *and* (MP) *and* (D_{MP}) *are feasible, then the following are equivalent*

1. *x solves (MP), (p, y) solves (D_{MP}) and* $\inf(MP) = \sup(D_{MP})$,
2. *x is feasible in (MP), (p, y) is feasible in (D_{MP}) and*

$$p \in \partial_x [f_0 + y \cdot H](x),$$
$$H(x) \in K, \quad y \in K^*, \quad y \cdot H(x) = 0$$

almost surely.

Proof. It suffices to note that, when $(x, y) \in \operatorname{dom} l$, we have

$$0 \in \partial_y [-l](x, y) = -H(x) + N_{K^*}(y)$$

if and only if $H(x) \in N_{K^*}(y)$. This is equivalent with the given complementarity condition by (6). □

Assumption 34.

1. (MP) is feasible,
2. $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$,
3. $\{x \in \mathcal{N} \mid f_j^\infty(x) \leq 0\ j = 0, \ldots l, f_j^\infty(x) = 0\ j = l + 1, \ldots, m\}$ is a linear space,
4. there exists a $p \in \mathcal{X}_a^\perp$ and an $\epsilon > 0$ such that for all $\lambda \in (1 - \epsilon, 1 + \epsilon)$ there exist a $y \in \mathcal{Y}$ and $\beta \in L^1$ such that $y \in K^*$ and

$$f_0(x, \omega) + y(\omega) \cdot H(x, \omega) \geq \lambda x \cdot p(\omega) - \beta(\omega) \quad \forall x \in \mathbb{R}^n.$$

Theorems 31 and 33 give the following.

Theorem 35. *Under Assumption* 34, $\inf(MP) = \sup(D_{MP})$ *and* (MP) *has a solution. In this case, a dual feasible* (p, y) *solves* (D_{MP}) *if and only if there exists a primal feasible* x *with*

$$\begin{gathered} p \in \partial_x[f_0 + y \cdot H](x), \\ H(x) \in K, \quad y \in K^*, \quad y \cdot H(x) = 0 \end{gathered}$$

almost surely.

In case of linear stochastic programming, the dual can be written down explicitly in terms of the problem data.

Example 36 (Linear stochastic programming)**.** *Consider the problem*

$$\begin{array}{ll} \text{minimize} & E[x \cdot c] \quad \text{over } x \in \mathcal{N} \\ \text{subject to} & Ax - b \in K \quad a.s. \end{array}$$

and assume that there exists $(x, u) \in \mathcal{X} \times \mathcal{U}$ *such that* $E[x \cdot c] < \infty$ *and* $Ax + u - b \in K$ *almost surely. The dual problem becomes*

$$\begin{array}{ll} \text{minimize} & E[b \cdot y] \quad \text{over } p \in \mathcal{X}_a^{\perp},\ y \in \mathcal{Y}, \\ \text{subject to} & c + A^* y = p,\ y \in K^* \quad a.s. \end{array}$$

and the scenariowise KKT-conditions

$$\begin{gathered} A^* y + c = p, \\ Ax - b \in K, \quad y \in K^*, \quad (Ax - b) \cdot y = 0, \end{gathered}$$

where A^* *is the scenariowise transpose of* A.

Indeed, this is a special case of (MP) *with* $f_0(x, \omega) = c(\omega) \cdot x$ *and* $f_j(x, \omega) = a_j(\omega) \cdot x - b_j(\omega)$ *for* $j = 1, \ldots, m$. *We get*

$$l(x, y, \omega) = x \cdot c(\omega) + y \cdot A(\omega)x - y \cdot b(\omega) - \delta_{K^*}(y)$$

and

$$\begin{aligned} f^*(p, y, \omega) &= \sup_{x \in \mathbb{R}^n} \{x \cdot v - l(x, y, \omega)\} \\ &= \begin{cases} y \cdot b(\omega) & \text{if } y \in K^* \text{ and } c(\omega) + A^*(\omega)y = p, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

This gives the dual problem while the KKT conditions follow directly from Theorem 33*.*

We will denote the *adapted projection* of an integrable process u by

$$^a u := (E_t u_t)_{t=0}^T.$$

Example 37 (Linear stochastic programming, reduced dual)**.** *In the setting of Example* 36 *assume that* $c \in \mathcal{V}$ *and* $A^*y \in \mathcal{V}$ *for all* $y \in \mathcal{Y}$*. Then, a pair* (p, y) *solves the dual if and only if* y *solves the* reduced dual problem

$$\begin{array}{ll} \text{minimize} & E[b \cdot y] \quad \text{over} \quad y \in \mathcal{Y}, \\ \text{subject to} & {}^a(c{+}A^*y) = 0, \; y \in K^* \quad a.s. \end{array}$$

and $p = c + A^*y - {}^a(c + A^*y)$*. If the elements of* c_t *and the columns* A_t *of* A *corresponding to* x_t *are* $\mathcal{F}_t$*-measurable, then the reduced dual can be written as*

$$\begin{array}{ll} \text{minimize} & E[b \cdot y] \quad \text{over} \quad y \in \mathcal{Y}, \\ \text{subject to} & c_t{+}A_t^* \cdot E_t y = 0 \;\; t = 0, \dots, T, \; y \in K^* \quad a.s. \end{array}$$

Proof. The first claim is clear and the second claim is a straightforward application of Lemma 63. □

6.2 Optimal stopping

Let R be a real-valued adapted stochastic process and consider the *optimal stopping problem*

$$\text{maximize} \quad ER_\tau \quad \text{over} \;\; \tau \in \mathcal{T}, \tag{OS}$$

where $\mathcal{T}$ is the set of *stopping times*, i.e. measurable functions $\tau : \Omega \to \{0, \dots, T{+}1\}$ such that $\{\omega \in \Omega \mid \tau(\omega) \le t\} \in \mathcal{F}_t$ for each $t = 0, \dots, T$. Choosing $\tau = T + 1$ is interpreted as not stopping at all. The problem

$$\begin{array}{ll} \text{maximize} & E\displaystyle\sum_{t=0}^{T} R_t x_t \quad \text{over} \;\; x \in \mathcal{N}, \\ \text{subject to} & x \ge 0, \; \displaystyle\sum_{t=0}^{T} x_t \le 1 \quad \text{a.s.} \end{array} \tag{ROS}$$

is a convex relaxation of (OS) in sense that their optimal values coincide and the extreme points of the feasible set of (ROS) can be identified with $\mathcal{T}$; see [19, Section 5.2].

Problem (ROS) fits the general duality framework with $n_t = 1$, $m = 1$,

$$f(x, u, \omega) = \begin{cases} -\sum_{t=0}^{T} x_t R_t(\omega) & \text{if } x \ge 0 \text{ and } \sum_{t=0}^{T} x_t + u \le 0, \\ +\infty & \text{otherwise} \end{cases}$$

and $\bar{u} = -1$. We get

$$
\begin{aligned}
l(x,y,\omega) &= \inf_{u\in\mathbb{R}^n}\{f(x,u,\omega) - uy\} \\
&= \inf_{u\in\mathbb{R}^n}\{-\sum_{t=0}^{T} x_t R_t(\omega) - uy \mid x \geq 0,\ \sum_{t=0}^{T} x_t + u \leq 0\} \\
&= \begin{cases} -\sum_{t=0}^{T} x_t R_t(\omega) + y\sum_{t=0}^{T} x_t + \delta_{\mathbb{R}^n_+}(x) & \text{if } y \geq 0, \\ -\infty & \text{otherwise} \end{cases} \\
&= \begin{cases} \sum_{t=0}^{T} x_t[y - R_t(\omega)] + \delta_{\mathbb{R}^n_+}(x) & \text{if } y \geq 0, \\ -\infty & \text{otherwise,} \end{cases}
\end{aligned}
$$

and

$$
\begin{aligned}
f^*(p,y,\omega) &= \sup_x\{x\cdot p - l(x,y,\omega)\} \\
&= \sup_{x\in\mathbb{R}^n_+} \sum_{t=0}^{T} x_t[p_t - y + R_t(\omega)] + \delta_{\mathbb{R}_+}(y) \\
&= \begin{cases} 0 & \text{if } y \geq 0 \text{ and } p_t + R_t(\omega) \leq y,\ t = 0,\dots,T, \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}
$$

Since $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$, Lemma 13 says that the dual of (ROS) can be written as

$$
\begin{array}{lll}
\text{minimize} & Ey & \text{over } (p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}_+ \\
\text{subject to} & p_t + R_t \leq y & t = 0,\dots,T \text{ a.s.}
\end{array} \tag{D_{OS}}
$$

It is clear that (ROS) is feasible, and (D_{OS}) is feasible when the pathwise maximum $\max_t R_t$ belongs $\mathcal{Y}$. Theorem 26 thus gives the following.

Theorem 38. *If $\max_t R_t \in \mathcal{Y}$, then the following are equivalent,*

1. *x solves (ROS), (p,y) solves (D_{OS}) and there is no duality gap.*
2. *$x \in \mathcal{X}_a$ and $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$ and*

$$
\begin{gathered}
x_t \geq 0,\ p_t + R_t \leq y,\ x_t(p_t + R_t - y) = 0 \quad t = 0,\dots,T, \\
y \geq 0,\ \sum_{t=0}^{T} x_t \leq 1,\ y(\sum_{t=0}^{T} x_t - 1) = 0
\end{gathered}
$$

almost surely.

In particular, a stopping time $\tau\in\mathcal{T}$ solves (OS), $(p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}$ solves (D_{OS}) and there is no duality gap if and only if $p_t + R_t \leq y$ for all t and $p_\tau + R_\tau = y$ almost surely.

Proof. The scenariowise KKT-condition in Theorem 26 can be written as

$$\begin{aligned} p_t + R_t - y &\in N_{\mathbb{R}_+}(x_t) \quad t = 0, \dots, T, \\ \sum_{t=0}^{T} x_t - 1 &\in N_{\mathbb{R}_+}(y), \end{aligned}$$

This is equivalent to the conditions given in the statement by (6). The second claim thus follows from Theorem 16 and Corollary 19. The last claim follows from the fact that a $\tau \in \mathcal{T}$ solves the optimal stopping problem (OS) if and only if the process $x \in \mathcal{X}_a$ given by

$$x_t = \begin{cases} 1 & \text{if } t = \tau, \\ 0 & \text{if } t \neq \tau \end{cases}$$

is optimal in (ROS); [19, Section 5.2]. □

Example 39 (Reduced dual). *Assume that $\max_t R_t \in \mathcal{Y}$ and $E_t\mathcal{Y} \subseteq \mathcal{Y} \subseteq \mathcal{V}_t$ for all t. The optimum value of (D_{OS}) equals that of the reduced dual*

$$\begin{aligned} &\text{minimize} \quad Ey_0 \quad \text{over } y \in \mathcal{M}_+^{\mathcal{Y}} \\ &\text{subject to} \quad R_t \le y_t \quad t = 0, \dots, T \ \textit{a.s.}, \end{aligned} \tag{rD_{OS}}$$

where $\mathcal{M}_+^{\mathcal{Y}}$ is the cone of nonnegative martingales y with $y_T \in \mathcal{Y}$. A pair $(p, y) \in \mathcal{V} \times \mathcal{Y}$ solves (D_{OS}) if and only if $p_t = y - E_t y$ and the process $y_t := E_t y$ solves (rD_{OS}). The following are equivalent

1. *$x \in \mathcal{X}_a$ solves (ROS), $y \in \mathcal{Y}$ solves (rD_{OS}) and there is no duality gap*
2. *$x \in \mathcal{X}_a$, $y \in \mathcal{M}^{\mathcal{Y}}$ and*

$$\begin{gathered} x_t \ge 0,\ R_t \le y_t,\ x_t(R_t - y_t) = 0 \quad t = 0, \dots, T, \\ y_T \ge 0,\ \sum_{t=0}^{T} x_t \le 1,\ y_T(\sum_{t=0}^{T} x_t - 1) = 0 \end{gathered}$$

almost surely.

In particular, a stopping time $\tau \in \mathcal{T}$ solves (OS), $y \in \mathcal{M}_+^{\mathcal{Y}}$ solves (rD_{OS}) and there is no duality gap if and only if $R_t \le y_t$ for all t and $R_\tau = y_\tau$ almost surely.

Proof. Let $y \in \mathcal{Y}$. The first two claims follow from the identity

$$\{y \in \mathcal{Y} \mid \exists p \in \mathcal{X}_a^\perp :\ p_t + R_t \le y \ \forall t\} = \{y \in \mathcal{Y} \mid R_t \le E_t[y] \ \forall t\}.$$

The last claim follows from the second one and Theorem 38. □

We end this section by applying the results of Section 5. Assumption 30 holds with $p = 0$ and $y = \max_t R_t$. Theorems 31 and 38 thus give the following.

Theorem 40. *If $\max_t R_t \in \mathcal{Y}$, then $\sup(OS) = \sup(ROS) = \inf(D_{OS})$, and (OS) and (ROS) have a solution. In this case, a dual feasible (p, y) solves (D_{OS}) if and only if there exists a stopping time $\tau \in \mathcal{T}$ with $p_\tau + R_\tau = y$ almost surely.*

6.3 Optimal control

Consider the optimal control problem

$$\begin{array}{ll} \text{minimize} & E\left[\sum_{t=0}^{T} L_t(X_t,U_t)\right] \quad \text{over} \quad (X,U)\in\mathcal{N}, \\ \text{subject to} & \Delta X_t = A_t X_{t-1} + B_t U_{t-1} + W_t \quad t=1,\dots,T \end{array} \tag{OC}$$

where the *state* X and the *control* U are processes with values in $\mathbb{R}^N$ and $\mathbb{R}^M$, respectively, A_t and B_t are $\mathcal{F}_t$-measurable random matrices, W_t is an $\mathcal{F}_t$-measurable random vector and the functions L_t are convex normal integrands. The linear constrains in (OC) are called the *system equations.*

The problem fits the general duality framework with $x=(X,U)$, $\bar{u} = (W_t)_{t=1}^T$ and

$$f(x,u,\omega) = \sum_{t=0}^{T} L_t(X_t,U_t,\omega) + \sum_{t=1}^{T} \delta_{\{0\}}(\Delta X_t - A_t(\omega)X_{t-1} - B_t(\omega)U_{t-1} - u_t).$$

We thus assume that $\mathcal{X}$ and $\mathcal{U}$ are solid decomposable spaces of $\mathbb{R}^{(T+1)(N+M)}$- and $\mathbb{R}^{TM}$-valued random variables, respectively, and that $(W_1,\dots,W_T)\in\mathcal{U}$. By solidity,

$$\mathcal{U} = \mathcal{U}_1\times\cdots\times\mathcal{U}_T, \quad \mathcal{U} = \mathcal{Y}_1\times\cdots\times\mathcal{Y}_T,$$

where $\mathcal{U}_t$ and $\mathcal{Y}_t$ are solid decomposable spaces of $\mathbb{R}^M$-valued random variables in separating duality under the bilinear form $(u_t,y_t)\mapsto E[u_t\cdot y_t]$. It follows that

$$\langle u,y\rangle = \sum_{t=1}^{T} E[u_t\cdot y_t].$$

For simplicity, we assume further that, for all t,

$$\begin{aligned} \mathcal{X}_t &= \mathcal{S}\times\mathcal{C}, \qquad & \mathcal{U}_t &= \mathcal{S} \\ \mathcal{V}_t &= \mathcal{S}'\times\mathcal{C}', \qquad & \mathcal{Y}_t &= \mathcal{S}' \end{aligned}$$

where $\mathcal{S}$ and $\mathcal{C}$ are solid decomposable spaces in separating duality with $\mathcal{S}'$ and $\mathcal{C}'$, respectively.

The Lagrangian integrand becomes

$$\begin{aligned} l(x,y,\omega) &= \inf_{u\in\mathbb{R}^m}\{f(x,u,\omega) - u\cdot y\} \\ &= \sum_{t=0}^{T} L_t(X_t,U_t,\omega) - \sum_{t=1}^{T}(\Delta X_t - A_t(\omega)X_{t-1} - B_t(\omega)U_{t-1})\cdot y_t \\ &= \sum_{t=0}^{T}[L_t(X_t,U_t,\omega) + X_t\cdot(\Delta y_{t+1} + A_{t+1}^*(\omega)y_{t+1}) + U_t\cdot B_{t+1}^*(\omega)y_{t+1}]. \end{aligned}$$

The conjugate integrand can be written as

$$
\begin{aligned}
f^*(v,y,\omega) &= \sup_{x\in\mathbb{R}^n}\{x\cdot v - l(x,y,\omega)\}\\
&= \sum_{t=0}^{T} L_t^*(v_t - (\Delta y_{t+1} + A_{t+1}^*(\omega)y_{t+1}, B_{t+1}^*(\omega)y_{t+1}),\omega),
\end{aligned}
$$

where $y_{T+1} := 0$, $A_{T+1} := 0$ and $B_{T+1} := 0$.

As soon as $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$, Lemma 13 says that the dual problem can be written as

$$
\begin{aligned}
&\text{maximize} \quad E\left[\sum_{t=1}^{T} W_t\cdot y_t - \sum_{t=0}^{T} L_t^*(p_t - (\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}))\right]\\
&\text{over} \quad (p,y)\in\mathcal{X}_a^\perp\times\mathcal{Y}.
\end{aligned}
\tag{D_{OC}}
$$

Theorem 16 and Corollary 19 give the following.

Theorem 41. *If $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$ and (OC) and (D_{OC}) are feasible, then the following are equivalent*

1. *(X,U) solves (OC), (p,y) solves (D_{OC}) and there is no duality gap,*
2. *(X,U) is feasible in (OC), (p,y) is feasible in (D_{OC}) and, for all t,*

$$
\begin{aligned}
p_t - (\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}) &\in \partial L_t(X_t,U_t),\\
\Delta X_t &= A_t X_{t-1} + B_t U_{t-1} + W_t
\end{aligned}
$$

almost surely.

The optimality conditions in Theorem 41 can be formulated also in the form of a stochastic maximum principle.

Remark 42 (Maximum principle)**.** *The scenariowise KKT-conditions in Theorem 41 mean that (X,U) satisfies the system equations and that*

$$
-(\Delta y_{t+1},0) \in \partial_{(X_t,U_t)} H_t(X_t,U_t,y_{t+1}) - p_t, \tag{7}
$$

where

$$
H_t(X_t,U_t,y_{t+1}) := L_t(X_t,U_t) + y_{t+1}\cdot(A_{t+1}X_t + B_{t+1}U_t).
$$

This can be written equivalently as

$$
\begin{aligned}
U_t &\in \operatorname*{argmin}_{U_t\in\mathbb{R}^M}\{H_t(X_t,U_t,y_{t+1}) - (X_t,U_t)\cdot p_t\},\\
-\Delta y_{t+1} &\in \partial_{X_t}\bar{H}_t(X_t,p_t,y_{t+1}),
\end{aligned}
$$

where

$$
\bar{H}_t(X_t,p_t,y_{t+1}) := \inf_{U_t\in\mathbb{R}^M}\{H_t(X_t,U_t,y_{t+1}) - (X_t,U_t)\cdot p_t\}.
$$

If, for all $(X_t, U_t, y_{t+1}) \in \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$,

$$\partial_{(X_t,U_t)} H_t(X_t, U_t, y_{t+1}) = \partial_{X_t} H_t(X_t, U_t, y_{t+1}) \times \partial_{U_t} H_t(X_t, U_t, y_{t+1}), \tag{8}$$

this can be written as

$$\begin{aligned} U_t &\in \underset{U_t \in \mathbb{R}^M}{\operatorname{argmin}} \{H_t(X_t, U_t, y_{t+1}) - (X_t, U_t) \cdot p_t\}, \\ -\Delta y_{t+1} &\in \partial_{X_t} \{H_t(X_t, U_t, y_{t+1}) - (X_t, U_t) \cdot p_t\} \end{aligned}$$

almost surely. Condition (8) *holds, in particular, if* L_t *is of the form*

$$L_t(X, U) = L_t^0(X, U) + L_t^1(X) + L_t^2(U),$$

where L_t^0 *is differentiable.*

Proof. The optimality conditions in Theorem 41 mean that

$$-(\Delta y_{t+1}, 0) \in \partial f_t(X_t, U_t), \tag{9}$$

where $f_t(X_t, U_t) := H_t(X_t, U_t, y_{t+1}) - (X_t, U_t) \cdot p_t$. The first claim thus follows from [26, Theorem 37.5]. Under (8), condition (9) can be written as

$$\begin{aligned} -\Delta y_{t+1} &\in \partial_{X_t} f_t(X_t, U_t), \\ 0 &\in \partial_{U_t} f_t(X_t, U_t), \end{aligned}$$

which is the second condition. □

Assumption 43. The spaces $\mathcal{S}'$ and $\mathcal{C}'$ are the Köthe duals of $\mathcal{S}$ and $\mathcal{C}$, respectively, and, for all t,

A $E_t \mathcal{S} \subseteq \mathcal{S}$ and $E_t \mathcal{C} \subseteq \mathcal{C}$,

B $A_t \mathcal{S} \subseteq \mathcal{S}$ and $B_t \mathcal{C} \subseteq \mathcal{S}$.

Except for condition B, Assumption 43 holds automatically e.g. in Lebesgue and Orlicz spaces. Part B holds e.g. if columns of A_t and B_t belong to L^∞ or, alternatively, if $\mathcal{C}$ and $\mathcal{S}$ are Cartesian products of spaces of finite moments (see [20, Section 6.1]) and the columns of A_t and B_t belong to $\mathcal{S}$. By Lemmas 3 and 4, Assumption 43 implies that, for all t,

A′ $E_t \mathcal{S}' \subseteq \mathcal{S}'$ and $E_t \mathcal{C}' \subseteq \mathcal{C}'$,

B′ $A_t \mathcal{S}' \subseteq \mathcal{S}'$ and $B_t \mathcal{S}' \subseteq \mathcal{C}'$.

Moreover, Assumption 43 implies the following.

Lemma 44. *Under Assumption* 43*,*

$$E_t[A_t^* y_t] = A_t^* E_t y_t \quad \textit{and} \quad E_t[B_t^* y_t] = B_t^* E_t y_t$$

for all $y_t \in \mathcal{S}'$.

Proof. B implies that $y_t \cdot A_t X_{t-1}$ is integrable for all $X_{t-1} \in \mathcal{S}$ and $y_t \in \mathcal{S}'$. Solidity of $\mathcal{S}$ implies that if we take $X_{t-1} \in \mathcal{S}$ and set all but one of its components to zero, the resulting vector is still in $\mathcal{S}$. Similarly for $\mathcal{S}'$. Condition B thus implies that $(A_t^*)_{i,j} y_t^j \in L^1$ for all i, j. The claim now follows from Lemma 63. □

Remark 45 (Reduced dual)**.** *Assume that each L_t is $\mathcal{F}_t$-measurable, each EL_t is proper on $\mathcal{S} \times \mathcal{C}$ and that Assumption 43 holds. Then the optimum value of the dual problem (D_{OC}) equals that of the* reduced dual problem

$$\text{maximize} \quad E\left[\sum_{t=1}^{T} W_t \cdot y_t - \sum_{t=0}^{T} [L_t^*(-E_t(\Delta y_{t+1} + A_{t+1}^* y_{t+1}, E_t B_{t+1}^* y_{t+1}))]\right] \quad \text{over} \quad y \in \mathcal{Y}_a \tag{rD_{OC}}$$

A pair $(p, y) \in \mathcal{X}_a^\perp \times \mathcal{Y}$ solves (D_{OC}), if and only if ${}^a y$ solves the reduced dual and

$$p_t = (\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}) - E_t(\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}).$$

The following are equivalent

1. *(X, U) solves (OC), y solves (rD_{OC}) and there is no duality gap,*
2. *(X, U) is feasible in (OC), y is feasible in (rD_{OC}) and, for all t,*

$$\begin{aligned} E_t[\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}] &\in \partial L_t(X_t, U_t), \\ \Delta X_t &= A_t X_{t-1} + B_t U_{t-1} + W_t \end{aligned}$$

   almost surely.

Proof. Given $y \in \mathcal{Y}$, the Jensen's inequality in Theorem 8 gives

$$\begin{aligned} &\inf_{p \in \mathcal{X}_a^\perp} E \sum_{t=0}^{T} L_t^*(p_t - (\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1})) \\ &\qquad\qquad = E \sum_{t=0}^{T} L_t^*(-E_t(\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1})), \end{aligned}$$

where the infimum is attained by the p given in the statement. The properties A′ and B′ stated after Assumption 43 imply that the p given in the statement belongs to $\mathcal{X}_a^\perp$ so it attains the infimum above. By Lemma 44,

$$-E_t(\Delta y_{t+1} + A_{t+1}^* y_{t+1}, B_{t+1}^* y_{t+1}) = -E_t(\Delta\, {}^a y_{t+1} + A_{t+1}^*\, {}^a y_{t+1}, B_{t+1}^*\, {}^a y_{t+1}),$$

so y can be chosen adapted without worsening the dual objective. The last claim follows from the second and Theorem 41. □

The optimality conditions Remark 45 are closely related to (5.20a)–(5.20e) in [3]. It should be noted however, that in [3], the functions L_t depend on W_{t+1}.

Remark 46 (Maximum principle in reduced form)**.** *The scenariowise optimality condition in Remark* 45 *can be written as*

$$-(E_t \Delta y_{t+1}, 0) \in \partial_{(X,U)} H_t(X_t, U_t, y_{t+1}),$$

where

$$H_t(X_t, U_t, y_{t+1}) := L_t(X_t, U_t) + E_t[A^*_{t+1} y_{t+1}] \cdot X_t + E_t[B^*_{t+1} y_{t+1}] \cdot U_t.$$

As in Remark 42*, this can be written also as*

$$\begin{aligned} U_t &\in \operatorname*{argmin}_{U_t \in \mathbb{R}^M} H_t(X_t, U_t, y_{t+1}), \\ -E_t \Delta y_{t+1} &\in \partial_X \bar{H}_t(X_t, y_{t+1}), \end{aligned}$$

where

$$\bar{H}_t(X_t, y_{t+1}) := \inf_{U_t \in \mathbb{R}^M} H_t(X_t, U_t, y_{t+1}).$$

We end this section by an application of Theorem 31. In optimal control, Assumption 30 holds under the following.

Assumption 47.

1. (OC) is feasible,
2. $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$,
3. $\{(X,U) \in \mathcal{N} \mid L^\infty_t(X_t, U_t) \leq 0 \; \Delta X_t = A_t X_{t-1} + B_t U_{t-1} \;\; t = 0, \dots T\}$ is a linear space,
4. there exists a $p \in \mathcal{X}_a^\perp$ and an $\epsilon > 0$ such that for all $\lambda \in (1-\epsilon, 1+\epsilon)$ there exist a $y \in \mathcal{Y}$ such that $(\lambda p, y)$ is feasible in (D_{OC}).

Theorems 31 and 41 give the following.

Theorem 48. *Under Assumption* 47*,* $\inf(OC) = \sup(D_{OC})$ *and* (OC) *has a solution. In this case, a dual feasible* (p, y) *solves* (D_{OC}) *if and only if there exists a primal feasible* x *such that, for all* t*,*

$$\begin{aligned} p_t - (\Delta y_{t+1} + A^*_{t+1} y_{t+1}, B^*_{t+1} y_{t+1}) &\in \partial L_t(X_t, U_t), \\ \Delta X_t &= A_t X_{t-1} + B_t U_{t-1} + W_t \end{aligned}$$

almost surely.

6.4 Problems of Lagrange

Consider the problem

$$\text{minimize} \quad E\sum_{t=0}^{T} K_t(x_t, \Delta x_t) \quad \text{over } x \in \mathcal{N}, \tag{L}$$

where x is a process of fixed dimension d, K_t are convex normal integrands and $x_{-1} := 0$. Problem (L) can be thought of as a discrete-time version of a problem studied in calculus of variations. Other problem formulations have $K_t(x_{t-1}, \Delta x_t)$ instead of $K_t(x_t, \Delta x_t)$ in the objective, or an additional term of the form $Ek(x_0, x_T)$, all of which fit the general format of stochastic optimization.

This fits the general duality framework with $\bar{u} = 0$ and

$$f(x, u, \omega) = \sum_{t=0}^{T} K_t(x_t, \Delta x_t + u_t, \omega).$$

We thus assume that both $\mathcal{X}$ and $\mathcal{U}$ are solid decomposable spaces of $\mathbb{R}^{(T+1)d}$-valued random variables. For simplicity, we assume that

$$\mathcal{X}_t = \mathcal{S}, \quad \mathcal{V}_t = \mathcal{S}', \quad \mathcal{U} = \mathcal{X}, \quad \mathcal{Y} = \mathcal{V},$$

where $\mathcal{S}$ and $\mathcal{S}'$ are solid decomposable spaces in separating duality.

The Lagrangian integrand becomes

$$\begin{aligned} l(x, y, \omega) &= \sum_{t=0}^{T} [\Delta x_t \cdot y_t + H_t(x_t, y_t, \omega))] \\ &= \sum_{t=0}^{T} [-x_t \cdot \Delta y_{t+1} + H_t(x_t, y_t, \omega)], \end{aligned}$$

where $y_{T+1} := 0$ and

$$H_t(x_t, y_t, \omega) := \inf_{u_t \in \mathbb{R}^d} \{K_t(x_t, u_t, \omega) - u_t \cdot y_t\}$$

is the associated *Hamiltonian*. The conjugate integrand can be written as

$$\begin{aligned} f^*(v, y, \omega) &= \sup\{x \cdot v - l(x, y, \omega)\} \\ &= \sum_{t=0}^{T} K_t^*(v_t + \Delta y_{t+1}, y_t, \omega). \end{aligned}$$

If (L) is feasible, Lemma 13 says that the dual problem can be written as

$$\text{maximize} \qquad E[-\sum_{t=0}^{T} K_t^*(p_t + \Delta y_{t+1}, y_t)] \quad \text{over } y \in \mathcal{Y}, p \in \mathcal{X}_a^{\perp} \tag{D_L}$$

where $y_{T+1} := 0$. Theorem 16 and Corollary 19 now give the following.

Theorem 49. *If* $\operatorname{dom} Ef \cap (\mathcal{X} \times \mathcal{U}) \neq \emptyset$ *and* (L) *and* (D_L) *are feasible, then the following are equivalent*

1. x *solves* (L), (p, y) *solves* (D_L) *and there is no duality gap,*
2. x *is feasible in* (L), (p, y) *is feasible in* (D_L) *and*

$$\begin{aligned} p_t + \Delta y_{t+1} &\in \partial_x H_t(x_t, y_t), \\ \Delta x_t &\in \partial_y[-H_t](x_t, y_t), \end{aligned}$$

   almost surely.

Note that, by [26, Theorem 37.5], the scenariowise KKT-conditions can be written equivalently as the discrete-time stochastic Euler–Lagrange equations

$$(p_t + \Delta y_{t+1}, y_t) \in \partial K_t(x_t, \Delta x_t) \tag{10}$$

or

$$(x_t, \Delta x_t) \in \partial K_t^*(p_t + \Delta y_{t+1}, y_t).$$

Assumption 50. The space $\mathcal{S}'$ is the Köthe dual of $\mathcal{S}$ and $E_t\mathcal{S} \subseteq \mathcal{S}$ for all t.

By Lemma 3, Assumption 50 implies that $E_t\mathcal{S}' \subset \mathcal{S}'$ for all t.

Remark 51 (Reduced dual)**.** *Consider Theorem* 49 *and assume that, for all* t*,* K_t *is* $\mathcal{F}_t$*-measurable and* EK_t *is proper on* $\mathcal{S} \times \mathcal{S}$ *and that Assumption* 50 *holds. Then the optimum value of the dual problem* (D_L) *equals that of the reduced dual problem*

$$\text{maximize} \qquad E[-\sum_{t=0}^{T} K_t^*(E_t\Delta y_{t+1}, y_t)] \quad \text{over } \ y \in \mathcal{Y}_a. \tag{rD_L}$$

A pair (p, y) *solves* (D_L) *if and only if* ${}^a y$ *solves* (rD_L) *and* $p_t = E_t\Delta y_{t+1} - \Delta y_{t+1}$*. The following are equivalent*

1. x *solves* (L), y *solves* (rD_L) *and there is no duality gap,*
2. x *is feasible in* (L), y *is feasible in* (rD_L) *and*

$$\begin{aligned} E_t\Delta y_{t+1} &\in \partial_x H_t(x_t, y_t), \\ \Delta x_t &\in \partial_y[-H_t](x_t, y_t), \end{aligned}$$

   almost surely.

Proof. Given $y \in \mathcal{Y}$, the Jensen's inequality in Theorem 8 gives

$$\inf_{p \in \mathcal{X}_a^\perp} E[\sum_{t=0}^{T} K_t^*(p_t + \Delta y_{t+1}, y_t)] = E[\sum_{t=0}^{T} K_t^*(E_t\Delta y_{t+1}, E_t y_t)],$$

where the infimum is attained by $p_t = E_t\Delta y_{t+1} - \Delta y_{t+1}$. This proves the first two claims. The last one follows from the second and Theorem 49. $\square$

Remark 52. *Any* $y \in \mathcal{Y}_a$ *has the* Doob decomposition

$$y_t = m_t + a_t,$$

where m *is a martingale and* a_t *is* $\mathcal{F}_{t-1}$*-measurable. Indeed, let* $\Delta a_t := E_{t-1}\Delta y_t$, $a_0 := 0$, $\Delta m_t := \Delta y_t - \Delta a_t$ *and* $m_0 := y_0$*. The reduced dual problem in* (rD_L) *can be written as*

$$\begin{aligned} &\text{maximize} \quad && E[-\sum_{t=0}^{T} K_t^*(\Delta a_{t+1}, m_t + a_t)] \quad \text{over} \quad (m,a) \in \mathcal{M}^{\mathcal{Y}} \times \mathcal{Y}_p \\ &\text{subject to} \quad && a_t \in \mathcal{F}_{t-1}, t = 1, \dots, T, a_0 = 0, \end{aligned}$$

where $\mathcal{M}^{\mathcal{Y}}$ *is the set of martingales and* $\mathcal{Y}_p$ *the predictable processes in* $\mathcal{Y}$*.*

Example 53 (Optimal stopping)**.** *The relaxed optimal stopping problem*

$$\underset{x \in \mathcal{N}_+}{\text{maximize}} \quad E\sum_{t=0}^{T} R_t \Delta x_t \quad \text{subject to} \quad \Delta x \geq 0,\ x_T \leq 1\ a.s. \tag{ROS}$$

from Section 6.2 can be written as a problem of Lagrange with $d = 1$ *and*

$$K_t(x_t, u_t) = -R_t u_t + \delta_{\mathbb{R}_-}(x_t - 1) + \delta_{\mathbb{R}_+}(u_t).$$

We get

$$\begin{aligned} K_t^*(v_t, y_t) &= \sup_{x_t, u_t \in \mathbb{R}} \{x_t \cdot v_t + u_t \cdot y_t - K_t(x_t, u_t)\} \\ &= \sup_{x_t, u_t \in \mathbb{R}} \{x_t \cdot v_t + u_t \cdot y_t + R_t u_t \mid x_t \leq 1,\ u_t \geq 0\} \\ &= \begin{cases} v_t & \text{if } v_t \geq 0 \text{ and } R_t + y_t \leq 0, \\ +\infty & \text{otherwise} \end{cases} \end{aligned}$$

so the reduced dual becomes

$$\begin{aligned} &\text{maximize} \quad && Ey_0 \ \text{ over } y \in \mathcal{Y}_a \\ &\text{subject to} \quad && E_t[\Delta y_{t+1}] \geq 0, \\ &&& R_t + y_t \leq 0, \end{aligned}$$

or with the change of variables $S := -y$,

$$\begin{aligned} &\text{minimize} \quad && ES_0 \ \text{ over } S \in \mathcal{Y}_a \\ &\text{subject to} \quad && E_t[\Delta S_{t+1}] \leq 0, \\ &&& R_t \leq S_t. \end{aligned}$$

Thus, feasible dual solutions are supermartingales that dominate the reward process R*.*

The Hamiltonian can be written as

$$\begin{aligned} H_t(x,y) &= \sup_{u\in\mathbb{R}}\{uy - K_t(x_t,u_t)\} \\ &= \begin{cases} +\infty & \text{if } x_t > 0, \\ 0 & \text{if } x_t \le 1 \text{ and } R_t + y_t \le 0, \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

so the optimality conditions become

$$\begin{aligned} E_t\Delta y_{t+1} &\in N_{\mathbb{R}_-}(x_t - 1) \\ \Delta x_t &\in N_{\mathbb{R}_-}(R_t + y_t). \end{aligned}$$

This implies that is Δx_t nonzero only when $R_t = -y_t$ and $E_t\Delta y_{t+1}$ is nonzero only when $x_t = 1$.

We end this section by an application of Theorem 31.

Assumption 54.

1. (L) is feasible,
2. $\operatorname{dom} Ef \cap (\mathcal{X}\times\mathcal{U}) \neq \emptyset$,
3. $\{x \in \mathcal{N} \mid \sum_{t=0}^T K_t^\infty(x_t, \Delta x_t) \le 0\}$ is a linear space,
4. there exists a $p \in \mathcal{X}_a^\perp$ and an $\epsilon > 0$ such that for all $\lambda \in (1-\epsilon, 1+\epsilon)$ there exist a $y \in \mathcal{Y}$ such that $(\lambda p, y)$ is feasible in (D_L).

Assumption 54 implies Assumption 30 so Theorems 31 and 49 give the following.

Theorem 55. *Under Assumption 54, $\inf(L) = \sup(D_L)$ and (OC) has a solution. In this case, a dual feasible (p,y) solves (D_L) if and only if there exists a primal feasible x such that, for all t,*

$$\begin{aligned} p_t + \Delta y_{t+1} &\in \partial_x H_t(x_t, y_t), \\ \Delta x_t &\in \partial_y[-H_t](x_t, y_t), \end{aligned}$$

almost surely.

6.5 Financial mathematics

Let $s = (s_t)_{t=0}^T$ be an adapted $\mathbb{R}^J$-valued stochastic process describing the unit prices of a finite set J of perfectly liquid tradeable assets. Assume also that there is a finite set K of derivative assets that can be bought or sold at time $t = 0$ and that provide random payments $C^j \in L^0$, $j \in K$ at time $t = T$. We denote $C = (C^j)_{j\in K}$. The cost of buying a derivative portfolio $x_{-1} \in \mathbb{R}^K$ at the best available market prices is denoted by $S_0(x_{-1})$. Such a function is convex and lsc with $S_0(0) = 0$. For example, if the buying and selling prices of the derivative assets are given by vectors $s^b \in \mathbb{R}^K$ and $s^a \in \mathbb{R}^K$ of bid- and ask-prices, respectively, and if we assume that one can buy and sell infinite quantities at these prices, then

$$S_0(x_{-1}) = \sup_{s \in [s^b, s^a]} x_{-1} \cdot s,$$

where $[s^b, s^a]$ denotes the K-dimensional box defined by s^b and s^a. If the bid and ask prices come with finite quantities given by vectors $q^b \in \mathbb{R}^K$ and $q^a \in \mathbb{R}^K$, respectively, then

$$S_0(x_{-1}) = \sup_{s \in [s^b, s^a]} x_{-1} \cdot s + \delta_{[-q^b, q^a]}(x_{-1}).$$

Similarly, one can express trading costs given by general limit order books by convex functions; see e.g. [12, 14].

Consider the problem of finding a dynamic trading strategy $x = (x_t)_{t=0}^T$ in the liquid assets J and a static portfolio x_{-1} in the derivatives K so that their combined revenue provides the "best hedge" against the financial liability of delivering a random amount $c \in L^0$ of cash at time T. If we assume that cash (or another numeraire asset) is a perfectly liquid asset that can be lent and borrowed at zero interest rate, the problem can be written as

$$\begin{array}{ll} \text{minimize} & EV\left(c - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} - C \cdot x_{-1} + S_0(x_{-1})\right) \quad \text{over} \quad x \in \mathcal{N}, x_{-1} \in \mathbb{R}^K, \\ \text{subject to} & x_t \in D_t \quad t = 0, \ldots, T-1 \ a.s., \end{array} \tag{SSH}$$

where V is a random "loss function" on $\mathbb{R}$ and D_t is a random $\mathcal{F}_t$-measurable set describing possible portfolio constraints. More precisely, the function V is a convex normal integrand such that $V(\cdot, \omega)$ nondecreasing and nonconstant for all ω. We will assume $D_T = \{0\}$, which means that all positions have to be closed at the terminal date. Note that nondecreasing convex loss functions V are in one-to-one correspondence with nondecreasing concave utility functions U via $V(c) = -U(-c)$; see e.g. [8, Section 8.2].

The special case where there are no statically held derivative assets, i.e. $K = \emptyset$, has been extensively studied in the literature of financial mathematics; see e.g. [24] its references. In the literature on "model-independent" mathematical finance, problems of finding both the dynamically updated portfolio x and the static part x_{-1} are often referred to as "semi-static hedging"; see e.g. [1]. One should note, however, that problem (SSH) is based on the assumption that one can buy and sell arbitrary quantities of the assets J at prices given by s. It also assumes that one can lend and borrow arbitrary amounts of cash at zero interest rate. Under these assumptions, the random variable c can be thought of as the difference of the claim to be hedged and the initial wealth and the sum in the objective can be interpreted as the proceeds from trading the assets

J over the period $[0,T]$. More realistic models for dynamic trading have been analyzed in [12, 14, 17].

As soon as $c \in \mathcal{U}$, problem (SSH) fits the general duality framework with the time index running from -1 to $T-1$, $\mathcal{F}_{-1} = \{\Omega, \emptyset\}$, $\bar{u} = c$ and f is given by

$$f(x,u,\omega) = V\left(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega) - C(\omega)\cdot x_{-1} + S_0(x_{-1}), \omega\right) + \sum_{t=0}^{T-1} \delta_{D_t(\omega)}(x_t,\omega)$$

if $x_{-1} \in \operatorname{dom} S_0$ and $f(x,u,\omega) := +\infty$ otherwise. By [17, Lemma 16], f is a normal integrand since our assumptions on V imply $V^\infty(c,\omega) > 0$ for $c > 0$. The Lagrangian integrand becomes

$$\begin{aligned} l(x,y,\omega) &= \inf_{u\in\mathbb{R}} \{f(x,u,\omega) - uy\} \\ &= y\left[S_0(x_{-1}) - C(\omega)\cdot x_{-1} - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega)\right] - V^*(y,\omega) + \sum_{t=0}^{T} \delta_{D_t(\omega)}(x_t). \end{aligned}$$

Here and in what follows, we define $yS_0 := \delta_{\operatorname{cl}\operatorname{dom} S_0}$ if $y = 0$. The conjugate of f becomes

$$\begin{aligned} f^*(v,y,\omega) &= \sup_{x\in\mathbb{R}^n} \{x\cdot v - l(x,y,\omega)\} \\ &= V^*(y,\omega) + \sum_{t=0}^{T-1} \sigma_{D_t(\omega)}(v_t + y\Delta s_{t+1}(\omega)) + \sup_{x_{-1}\in\mathbb{R}^K} \{x_{-1}\cdot(v_{-1} + yC(\omega)) - yS_0(x_{-1})\} \\ &= V^*(y,\omega) + \sum_{t=0}^{T-1} \sigma_{D_t(\omega)}(v_t + y\Delta s_{t+1}(\omega)) + (yS_0)^*(v_{-1} + yC(\omega)). \end{aligned}$$

It is natural to assume that $S_0(0) = 0$ and $0 \in D_t$ almost surely for all t. If EV is proper on $\mathcal{U}$, Lemma 13 then says that the dual problem can be written as

$$\underset{p\in\mathcal{X}_a^\perp, y\in\mathcal{Y}}{\text{maximize}} \qquad E\left[cy - V^*(y) - \sum_{t=0}^{T-1} \sigma_{D_t}(p_t + y\Delta s_{t+1}) - (yS_0)^*(p_{-1} + yC)\right]. \tag{D_{SSH}}$$

Theorem 56. *If (SSH) and (D_{SSH}) are feasible, then the following are equivalent*

1. *x solves (SSH), (p,y) solves (D_{SSH}) and there is no duality gap.*
2. *x is feasible in (SSH), (p,y) is feasible in (D_{SSH}) and*

$$y \in \partial V(u - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} - C\cdot x_{-1} + S_0(x_{-1})),$$

$$p_t + y\Delta s_{t+1} \in N_{D_t}(x_t) \quad t = 0,\dots,T,$$

$$p_{-1} + yC \in \partial(yS_0)(x_{-1})$$

almost surely.

Proof. By Theorem 16 and Corollary 19 it suffices to show that the scenariowise optimality conditions

$$p \in \partial_x l(x,y), \quad \bar{u} \in \partial_y[-l](x,y) \quad \text{a.s.}$$

can be written as the scenariowise conditions given above. By the sum-rule of subdifferentiation [26, Theorem 23.8], the first condition gives the last two in the statement while the second condition becomes

$$S_0(x_{-1}) - C(\omega)\cdot x_{-1} - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega) \in \partial V^*(y).$$

By [26, Corollary 23.5.1], $(\partial V^*)^{-1} = \partial V$ so this becomes the first condition above. □

Under the following assumption, the dual problem (D_{SSH}) can be written in a reduced form where the shadow price of information p has been optimized for a each given y.

Assumption 57. $\mathcal{X} = L^\infty$, $\mathcal{V} = L^1$, $\mathcal{Y}$ is the Köthe dual of $\mathcal{U}$ and, for all t,

A $E_t\mathcal{U} \subseteq \mathcal{U}$,

B $\Delta s_{t+1} \in \mathcal{U}$.

If part B holds, Assumption 57 holds, e.g., in Lebesgue and Orlicz spaces; see the examples in Section 2.1. By Lemmas 3 and 4, Assumption 57 implies that, for all t,

A′ $E_t\mathcal{Y} \subseteq \mathcal{Y}$,

B′ $y\Delta s_{t+1} \in L^1$ for all $y \in \mathcal{Y}$.

Remark 58 (Reduced dual)**.** *Under Assumption 57 the optimum value of (D_{SSH}) equals that of the reduced dual problem*

$$\underset{y\in\mathcal{Y}}{\text{maximize}} \qquad E\left[cy - V^*(y) - \sum_{t=0}^{T-1} \sigma_{D_t}(E_t[y\Delta s_{t+1}]) - (E[y]S_0)^*(E[yC])\right]. \tag{rD_{SSH}}$$

A pair (p,y) solves (D_{SSH}) if and only if y solves (rD_{SSH}) and

$$p_{-1} := \frac{E[yC]}{E[y]}y - yC \quad \text{and} \quad p_t = E_t[y\Delta s_{t+1}] - y\Delta s_{t+1} \quad t = 0,\dots,T-1,$$

where the fraction is interpreted as 0 if $E[y]=0$. The following are equivalent

1. *x solves (SSH), y solves (rD_{SSH}) and there is no duality gap,*

2. x is feasible in (SSH), y is feasible in (rD_{SSH}) and

$$
\begin{aligned}
y &\in \partial V(u - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} - C \cdot x_{-1} + S_0(x_{-1})),\\
E_t[y\Delta s_{t+1}] &\in N_{D_t}(x_t) \quad t = 0, \dots, T,\\
\frac{E[yC]}{E[y]} y &\in \partial(yS_0)(x_{-1})
\end{aligned}
$$

almost surely. Again, the fraction is interpreted as 0 *if* $E[y] = 0$

Proof. Given $y \in \mathcal{Y}_+$, the Jensen's inequality in Theorem 8 gives

$$
\inf_{p \in \mathcal{X}_a^\perp} E\left[\sum_{t=0}^{T-1} \sigma_{D_t}(p_t + y\Delta s_{t+1}) + (yS_0)^*(p_{-1} + yC)\right] = E\left[\sum_{t=0}^{T-1} \sigma_{D_t}(E_t[y\Delta s_{t+1}]) + (E[y]S_0)^*(E[yC])\right],
$$

where the infimum is attained by the p given in the statement. Indeed, if $E[y] \neq 0$, this choice gives, by sublinearity,

$$
\begin{aligned}
(yS_0)^*(p_{-1} + yC) &= (yS_0)^*(\frac{E[yC]}{E[y]} y)\\
&= (\frac{y}{E[y]} E[y]S_0)^*(\frac{y}{E[y]} E[yC])\\
&= \frac{y}{E[y]}(E[y]S_0)^*(E[yC]).
\end{aligned}
$$

If $E[y] = 0$ then $y = 0$ almost surely and $(p, y) = (0, 0)$ is feasible in (D_{SSH}) and gives the same objective value. This proves the first two claims. The third follows from the second and Theorem 56. □

If the optimality conditions were satisfied with $y = 0$, the first condition would imply

$$
u - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} - C \cdot x_{-1} + S_0(x_{-1}) \in \operatorname{argmin} V \quad a.s.,
$$

which would mean that x achieves a perfect hedge in terms of the loss function V. This is impossible e.g. if V is strictly increasing. In this case $y > 0$ almost surely and the scenariowise optimality conditions above can be written as

$$
\begin{aligned}
\lambda \frac{dQ}{dP} &\in \partial V(u - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} - C \cdot x_{-1} + S_0(x_{-1})),\\
E_t^Q[\Delta s_{t+1}] &\in N_{D_t}(x_t) \quad t = 0, \dots, T,\\
E^Q[C] &\in \partial S_0(x_{-1}),
\end{aligned}
$$

where $\lambda = E[y]$ and Q is the probability measure defined by $dQ/dP = y/E[y]$. If S_0 is sublinear, the reduced dual can be written as

$$\begin{array}{ll} \underset{y\in\mathcal{Y}}{\text{maximize}} & E\left[cy - V^*(y) - \sum_{t=0}^{T-1} \sigma_{D_t}(E_t[y\Delta s_{t+1}])\right] \\ \text{subject to} & E[yC] \in E[y]\operatorname{dom} S_0^*, \end{array}$$

where the set $E[y]\operatorname{dom} S_0^*$ is interpreted as the recession cone of $\operatorname{dom} S_0^*$ if $E[y] = 0$. If $E[y] \neq 0$, the constraint means that

$$E^Q C \in \operatorname{dom} S_0^*.$$

The constraint in the dual thus requires that the measure Q be "calibrated" to the observed market prices of the claims C. For example, if infinite quantities are available to buy and sell at prices $s^a \in \mathbb{R}^K$ and $s^b \in \mathbb{R}^K$, respectively, then

$$S_0(x_{-1}) = \sup_{s\in[s^b,s^a]} x_{-1} \cdot s$$

and the constraint becomes $E^Q C \in [s^b, s^a]$.

It turns out that, in the absence of portfolio constraints, the linearity condition in Assumption 30 becomes the classical *no-arbitrage* condition

$$x \in \mathcal{N},\ \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} \geq 0\ a.s. \implies \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} = 0\ a.s.; \tag{NA}$$

see [19, Section 5.5]. The second condition in Assumption 30 holds, in particular, if there exists a martingale measure $Q \ll P$ such that

$$dQ/dP \in \mathcal{Y} \cap \operatorname{dom} EV^*,$$

V is deterministic such that either

$$\limsup_{u\to-\infty} \frac{uV'(u)}{V(u)} < 1 \quad \text{or} \quad \liminf_{u\to+\infty} \frac{uV'(u)}{V(u)} > 1;$$

see [19, Remark 53]. More generally, Assumption 30 is implied by the following.

Assumption 59. There are no portfolio constraints and

1. the set

$$\mathcal{L} := \{x \in \mathcal{N} \mid S_0^\infty(x_{-1}) - x_{-1} \cdot C - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} \leq 0,\ x_t \in D_t^\infty\ \forall t\}$$

   is a linear space,

2. there exists y feasible in the reduced dual (rD_{SSH}) and ϵ such that $\lambda y \in \operatorname{dom} EV^*$ for all $\lambda \in (1-\epsilon, 1+\epsilon)$.

Theorem 60. *Under Assumption* 59, $\bar{\varphi}$ *is closed and the infimum in its definition is attained for every* $(z,u) \in \mathcal{X}\times\mathcal{U}$. *In particular,* $\inf(SSH) = \sup(D_{SSH})$ *and* (SSH) *has a solution. In this case, a dual feasible* (p,y) *solves* (D_{SSH}) *if and only if there is a primal feasible* x *such that*

$$
\begin{aligned}
y &\in \partial V(u - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} - C\cdot x_{-1} + S_0(x_{-1})),\\
p_t + y\Delta s_{t+1} &\in N_{D_t}(x_t) \quad t = 0,\dots,T,\\
p_{-1} + yC &\in \partial(yS_0)(x_{-1})
\end{aligned}
$$

almost surely.

Proof. By [26, Theorem 9.3] and [11, Theorem 7.3],

$$
f^\infty(x,u,\omega) = V^\infty\left(u - \sum_{t=0}^{T-1} x_t\cdot\Delta s_{t+1}(\omega) - C(\omega)\cdot x_{-1} + S_0^\infty(x_{-1}),\omega\right) + \sum_{t=0}^{T-1}\delta_{D_t^\infty(\omega)}(x_t,\omega).
$$

Since V is nonconstant and nondecreasing,

$$
\{c\in\mathbb{R} \mid V^\infty(c)\le 0\} = \mathbb{R}_-
$$

so the linearity condition in Assumption 59 implies the one in Assumption 30. For y from Assumption 59, p defined by

$$
p_{-1} := \frac{E[yC]}{E[y]}y - yC \quad \text{and} \quad p_t = E_t[y\Delta s_{t+1}] - y\Delta s_{t+1} \quad t = 0,\dots,T-1,
$$

satisfies Assumption 30. Thus the claims follow from Theorems 31 and 56. $\square$

Given a convex function g on $\mathbb{R}^n$, the set

$$
\operatorname{lin} g = \{x\in\mathbb{R}^n \mid g^\infty(x) = -g^\infty(-x)\}
$$

is called the *lineality space* of g.

Example 61 (Robust no-arbitrage condition)**.** *Assume that there are no portfolio constraints and that there exists a cost function* $\tilde{S}_0$ *such that*

$$
\begin{aligned}
\tilde{S}_0(x,\omega) &\le S_0^\infty(x,\omega) \quad \forall x\in\mathbb{R}^K,\\
\tilde{S}_0(x,\omega) &< S_0^\infty(x,\omega) \quad \forall x\notin \operatorname{lin} S_0(\cdot,\omega)
\end{aligned}
$$

and the market model described by $\tilde{S}_0$ *and* s *satisfies the no-arbitrage condition*

$$
\mathcal{C}\cap L^0_+ = \{0\}, \tag{11}
$$

where

$$
\mathcal{C} := \{c\in\mathcal{U} \mid \exists x)\in\mathcal{N} : \sum_{t=0}^{T-1} x_t\cdot\Delta s_{t+1} + x_{-1}\cdot c - \tilde{S}_0(x_{-1}) \ge c \quad a.s.\}.
$$

Then the linearity condition in Assumption 59 *holds. A violation of* (11) *would mean that there is a trading strategy* x *that superhedges a nonzero nonnegative claim* c.

Proof. If the linearity condition fails, there is a $x \in \mathcal{L}$ such that

$$S_0^\infty(-x_{-1}) + x_{-1} \cdot C + \sum_{t=0}^{T-1} x_t \Delta s_{t+1} > 0$$

on a set $A \in \mathcal{F}$ with $P(A) > 0$. It suffices to show that x is an arbitrage strategy for $\tilde{S}_0$. Since $x \in \mathcal{L}$, and $\tilde{S}_0 \leq S_0^\infty$, we have

$$\tilde{S}_0(x_{-1}) - x_{-1} \cdot C - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} \leq 0.$$

If $x_{-1} \notin \operatorname{lin} S_0$, then $\tilde{S}_0(x) < S_0^\infty(x)$ and the inequality is strict so x is an arbitrage strategy. If $x_{-1} \in \operatorname{lin} S_0$, then $\tilde{S}_0(x_{-1}) \leq S_0^\infty(x_{-1}) = -S_0^\infty(-x_{-1})$ so

$$\tilde{S}_0(x_{-1}) - x_{-1} \cdot C - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} \leq -S_0^\infty(-x_{-1}) - x_{-1} \cdot C - \sum_{t=0}^{T-1} x_t \Delta s_{t+1} < 0$$

on A so x is an arbitrage strategy in this case too. □

7 Appendix

Given extended real-valued random variables ξ_1 and ξ_2, their pointwise sum $\xi_2 + \xi_2$ is well-defined by the usual algebraic operations of the extended real-line $\overline{\mathbb{R}}$ except when one of them takes the value $+\infty$ and the other one $-\infty$. In this exceptional case, we define the sum as $+\infty$.

The proofs of the following two lemmas can be found in [19].

Lemma 62. *Given extended real-valued random variables ξ_1 and ξ_2, we have*

$$E[\xi_1 + \xi_2] = E[\xi_1] + E[\xi_2]$$

under any of the following:

1. *$\xi_1^+, \xi_2^+ \in L^1$ or $\xi_1^-, \xi_2^- \in L^1$.*
2. *$\xi_1 \in L^1$ or $\xi_2 \in L^1$,*
3. *ξ_1 or ξ_2 is $\{0, +\infty\}$-valued.*

Lemma 63. *Let ξ_1 and ξ_2 be extended real-valued random variables.*

1. *If ξ_1 and ξ_2 are quasi-integrable and satisfy any of the conditions in Lemma 62, then $\xi_1 + \xi_2$ is quasi-integrable and*

$$E^{\mathcal{G}}[\xi_1 + \xi_2] = E^{\mathcal{G}}[\xi_1] + E^{\mathcal{G}}[\xi_2].$$

2. *If ξ_2 and $(\xi_1 \xi_2)$ are quasi-integrable, and ξ_1 is $\mathcal{G}$-measurable, then*

$$E^{\mathcal{G}}[\xi_1 \xi_2] = \xi_1 E^{\mathcal{G}}[\xi_2].$$

References

[1] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.*, 17(3):477–501, 2013.

[2] S. Biagini, T. Pennanen, and A.-P. Perkkiö. Duality and optimality conditions in stochastic optimization and mathematical finance. *Journal of Convex Analysis*, 25, 2018.

[3] P. Carpentier, J.-P. Chancelier, G. Cohen, and M. De Lara. *Stochastic multi-stage optimization*, volume 75 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2015. At the crossroads between discrete time stochastic control and stochastic programming.

[4] M. H. A. Davis. Dynamic optimization: a grand unification. In *Proceedings of the 31st IEEE Conference on Decision and Control*, volume 2, pages 2035 – 2036, 1992.

[5] M. H. A. Davis and G. Burstein. A deterministic approach to stochastic optimal control with application to anticipative control. *Stochastics and Stochastics Reports*, 40(3&4):203–256, 1992.

[6] M. H. A. Davis and I. Karatzas. A deterministic approach to optimal stopping. In *Probability, statistics and optimisation*, Wiley Ser. Probab. Math. Statist. Probab. Math. Statist., pages 455–466. Wiley, Chichester, 1994.

[7] I. V. Evstigneev. Measurable selection and dynamic programming. *Math. Oper. Res.*, 1(3):267–272, 1976.

[8] H. Föllmer and A. Schied. *Stochastic finance*. De Gruyter Graduate. De Gruyter, Berlin, 2016. An introduction in discrete time, Fourth revised and extended edition of [MR1925197].

[9] Y. M. Kabanov. Hedging and liquidation under transaction costs in currency markets. *Finance and Stochastics*, 3(2):237–248, 1999.

[10] J. L. Kelley and I. Namioka. *Linear topological spaces*. Springer-Verlag, New York, 1976. With the collaboration of W. F. Donoghue, Jr., Kenneth R. Lucas, B. J. Pettis, E. T. Poulsen, G. B. Price, W. Robertson, W. R. Scott, and K. T. Smith, Second corrected printing, Graduate Texts in Mathematics, No. 36.

[11] T. Pennanen. Graph-convex mappings and K-convex functions. *J. Convex Anal.*, 6(2):235–266, 1999.

[12] T. Pennanen. Arbitrage and deflators in illiquid markets. *Finance and Stochastics*, 15(1):57–83, 2011.

[13] T. Pennanen. Convex duality in stochastic optimization and mathematical finance. *Mathematics of Operations Research*, 36(2):340–362, 2011.

[14] T. Pennanen. Superhedging in illiquid markets. *Mathematical Finance*, 21(3):519–540, 2011.

[15] T. Pennanen and I. Penner. Hedging of claims with physical delivery under convex transaction costs. *SIAM Journal on Financial Mathematics*, 1:158–178, 2010.

[16] T. Pennanen and A.-P. Perkkiö. Stochastic programs without duality gaps. *Mathematical Programming*, 136(1):91–110, 2012.

[17] T. Pennanen and A.-P. Perkkiö. Convex duality in optimal investment and contingent claim valuation in illiquid markets. *Finance and Stochastics*, 22(4):733–771, Oct 2018.

[18] T. Pennanen and A.-P. Perkkiö. Shadow price of information in discrete time stochastic optimization. *Math. Program.*, 168(1-2, Ser. B):347–367, 2018.

[19] T. Pennanen and A.-P. Perkkiö. Dynamic programming in convex stochastic optimization. *Manuscript*, 2022.

[20] T. Pennanen and A.-P. Perkkiö. Topological duals of locally convex function spaces. *Positivity*, 2022.

[21] T. Pennanen, A.-P. Perkkiö, and M. Rásonyi. Existence of solutions in non-convex dynamic programming and optimal investment. *Mathematics and Financial Economics*, pages 1–16, 2016.

[22] A.-P. Perkkiö. Stochastic programs without duality gaps for objectives without a lower bound. *manuscript*, 2016.

[23] S. R. Pliska. Duality theory for some stochastic control models. In *Stochastic differential systems (Bad Honnef, 1982)*, volume 43 of *Lect. Notes Control Inf. Sci.*, pages 329–337. Springer, Berlin, 1982.

[24] M. Rásonyi and L. Stettner. On utility maximization in discrete-time financial market models. *Ann. Appl. Probab.*, 15(2):1367–1395, 2005.

[25] R. T. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24:525–539, 1968.

[26] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[27] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.

[28] R. T. Rockafellar and R. J.-B. Wets. Nonanticipativity and L^1-martingales in stochastic optimization problems. *Math. Programming Stud.*, (6):170–187, 1976. Stochastic systems: modeling, identification and optimization, II (Proc. Sympos., Univ Kentucky, Lexington, Ky., 1975).

[29] R. T. Rockafellar and R. J.-B. Wets. The optimal recourse problem in discrete time: L^1-multipliers for inequality constraints. *SIAM J. Control Optimization*, 16(1):16–36, 1978.

[30] R. T. Rockafellar and R. J.-B. Wets. Deterministic and stochastic optimization problems of Bolza type in discrete time. *Stochastics*, 10(3-4):273–312, 1983.

[31] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[32] R. Tyrrell Rockafellar. Saddle-points and convex analysis. In *Differential Games and Related Topics (Proc. Internat. Summer School, Varenna, 1970)*, pages 109–127. North-Holland, Amsterdam, 1971.

[33] W. Schachermayer. A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. *Insurance Math. Econom.*, 11(4):249–257, 1992.

[34] W. Schachermayer. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Math. Finance*, 14(1):19–48, 2004.

[35] R. J-B Wets. On the relation between stochastic and deterministic optimization. In A. Bensoussan and J.L. Lions, editors, *Control Theory, Numerical Methods and Computer Systems Modelling*, volume 107 of *Lecture Notes in Economics and Mathematical Systems*, pages 350–361. Springer, 1975.