Interchangeability principle and dynamic equations in risk averse stochastic programming

Alexander Shapiro

School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, Georgia 30332-0205, USA
e-mail: ashapiro@isye.gatech.edu

Abstract. In this paper we consider interchangeability of the minimization operator with monotone risk functionals. In particular we discuss the role of strict monotonicity of the risk functionals. We also discuss implications to solutions of dynamic programming equations of risk averse multistage stochastic programming problems.

Key Words: Interchangeability principle, strict monotonicity, convex risk measures, two and multistage stochastic programming, dynamic equations, time consistency

1 Introduction

Interchangeability of the minimization and expectation operators is a basis for deriving dynamic programming equations in multistage stochastic programming. In a setting of functional spaces such interchangeability principle is derived, e.g., in Rockafellar and Wets [1, Theorem 14.60]. In a risk averse case interchangeability of the minimization and risk functionals was considered in [2, Theorem 7.1] and [3, Proposition 6.60]. We revisit the question of interchangeability with an emphasis on the role of *strict* monotonicity of considered risk functionals. Importance of such strict monotonicity was already pointed in relation to time consistency of optimal policies of risk averse stochastic programs in [3, Section 6.8.5] and [4]. We also discuss implications of strict monotonicity to solutions of dynamic programming equations.

2 Interchangeability principle

Let $(\Omega, \mathcal{F})$ be a sample space, i.e., $\mathcal{F}$ is a sigma algebra of subsets of Ω, X be an abstract set and $f : X \times \Omega \to \mathbb{R} \cup \{+\infty\}$. Consider

$$F(\omega) := \inf_{x \in X} f(x, \omega). \tag{2.1}$$

Let $\mathcal{Z}$ be a linear space of $\mathcal{F}$-measurable functions $Z : \Omega \to \mathbb{R}$. We deal with following cases.

($\aleph$1) The set $\Omega = \{\omega_1, ..., \omega_m\}$ is finite, $\mathcal{F}$ is sigma algebra of all subsets of Ω and $\mathcal{Z}$ is the space of all functions $Z : \Omega \to \mathbb{R}$. In this case the space $\mathcal{Z}$ is m-dimensional and can be identified with $\mathbb{R}^m$.

($\aleph$2) The sample space $(\Omega, \mathcal{F})$ is equipped with probability measure P and $\mathcal{Z} := L_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$. Equipped with the norm $\|Z\| = \left(\int |Z|^p dP\right)^{1/p}$ for $p \in [1, \infty)$, and $\|Z\| = \operatorname{ess\,sup}|Z(\omega)|$ for $p = \infty$, this becomes a Banach space.

($\aleph$3) The set Ω is a compact metric space, $\mathcal{F}$ is the Borel sigma algebra of Ω, and $\mathcal{Z} := C(\Omega)$ is the space of continuous functions $Z : \Omega \to \mathbb{R}$ equipped with the sup-norm $\|Z\| = \sup_{\omega \in \Omega} |Z(\omega)|$.

Of course the above case ($\aleph$1), of finite set Ω, can be considered as a particular case of setting ($\aleph$3), we write it separately since in that case the analysis simplifies considerably. In case ($\aleph$2) an element Z of the space $\mathcal{Z}$ is a class of p-integrable functions $Z : \Omega \to \mathbb{R}$ which coincide for all $\omega \in \Omega$ accept on a set of P-measure zero. By writing equalities like $F(\cdot) := \inf_{x \in X} f(x, \cdot)$ we mean that this equality holds for all $\omega \in \Omega$ in cases ($\aleph$1) and ($\aleph$3), and it holds for P-almost every (a.e.) $\omega \in \Omega$ in case ($\aleph$2).

In the above cases ($\aleph$1)-($\aleph$3) there is a naturally defined order relation between $Z, Z' \in \mathcal{Z}$. We write $Z \succeq Z'$ if $Z(\omega) \geq Z'(\omega)$ for all $\omega \in \Omega$ in cases ($\aleph$1) and ($\aleph$3), and $Z(\omega) \geq Z'(\omega)$ for a.e. $\omega \in \Omega$ in case ($\aleph$2). Consider a functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$. It is said that $\mathcal{R}$ is monotone if $Z, Z' \in \mathcal{Z}$ and $Z \succeq Z'$, then $\mathcal{R}(Z) \geq \mathcal{R}(Z')$. It is said that $\mathcal{R}$ is strictly monotone if $\mathcal{R}$ is monotone and $Z \succeq Z'$ and $Z \neq Z'$ imply that $\mathcal{R}(Z) > \mathcal{R}(Z')$. By saying that $\mathcal{R}$ is continuous we mean that it is continuous with respect to the norm topology of the space $\mathcal{Z}$. Let $\mathcal{X}$ be the set of mappings $\chi : \Omega \to X$ such that $f_\chi \in \mathcal{Z}$, where $f_\chi(\cdot) := f(\chi(\cdot), \cdot)$. We also write $\mathcal{R}(f(\chi(\omega), \omega)$ for $\mathcal{R}(f_\chi)$. Consider the following equation

$$\mathcal{R}(F) = \inf_{\chi \in \mathcal{X}} \mathcal{R}(f_\chi), \tag{2.2}$$

and the implications

$$\bar{\chi}(\cdot) \in \arg\min_{x \in X} f(x, \cdot) \Rightarrow \bar{\chi} \in \arg\min_{\chi \in \mathcal{X}} \mathcal{R}(f_\chi), \tag{2.3}$$

$$\bar{\chi} \in \arg\min_{\chi \in \mathcal{X}} \mathcal{R}(f_\chi) \Rightarrow \bar{\chi}(\cdot) \in \arg\min_{x \in X} f(x, \cdot). \tag{2.4}$$

Proposition 2.1 *Suppose that $F \in \mathcal{Z}$ and $\mathcal{R}$ is monotone. Then the following holds.* (i) *Suppose that the minimum of $f(x,\omega)$ over $x \in X$ is attained for all $\omega \in \Omega$. Then* (2.2) *and* (2.3) *follow; the implication* (2.4) *also follows if either the set $\arg\min_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi)$ is a singleton or $\mathcal{R}$ is strictly monotone.* (ii) *Suppose that $\mathcal{R}(\cdot)$ is continuous at F and there exists a sequence $\chi_k \in \mathcal{X}$ such that f_{χ_k} converges to F. Then* (2.2) *and* (2.3) *follow; the implication* (2.4) *also follows if $\mathcal{R}$ is strictly monotone.*

Proof. We have that $f_\chi \succeq F$ for any $\chi \in \mathcal{X}$. Hence by monotonicity of $\mathcal{R}$ it follows that $\mathcal{R}(f_\chi) \geq \mathcal{R}(F)$, and thus

$$\inf_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi) \geq \mathcal{R}(F).$$

Conversely, consider the setting of case (i), i.e., there exists

$$\bar{\chi}(\cdot) \in \arg\min_{x\in X} f(x,\cdot). \tag{2.5}$$

Then $F = f_{\bar{\chi}}$ and since $F \in \mathcal{Z}$ it follows that $\bar{\chi} \in \mathcal{X}$, and hence

$$\mathcal{R}(F) = \mathcal{R}(f_{\bar{\chi}}) \geq \inf_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi).$$

Thus (2.2) and the implication (2.3) follow. As it was shown above the minimizer $\bar{\chi}$ belongs to the set $\arg\min_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi)$. If this set is a singleton, then the implication (2.4) follows.

Suppose now that $\mathcal{R}$ is strictly monotone. Let $\hat{\chi} \in \arg\min_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi)$. We have that $\mathcal{R}(F) = \mathcal{R}(f_{\hat{\chi}})$. Also $f_{\hat{\chi}} \succeq F$ and hence by strict monotonicity of $\mathcal{R}$ it follows that $f_{\hat{\chi}} = F$, i.e., $f(\hat{\chi}(\cdot),\cdot) = \inf_{x\in X} f(x,\cdot)$. This proves the implication (2.4). This completes the proof of case (i).

Consider now case (ii). Let $\chi_k \in \mathcal{X}$ be a sequence such that f_{χ_k} converges to F. It follows by continuity of $\mathcal{R}$ that

$$\mathcal{R}(F) = \lim_{k\to\infty} \mathcal{R}(f_{\chi_k}) \geq \inf_{\chi\in\mathcal{X}} \mathcal{R}(f_\chi).$$

Hence (2.2) follows, and (2.3) follows as well. If moreover $\mathcal{R}$ is strictly monotone, then the implication (2.4) follows by the same arguments as in case (i). ∎

Let us discuss assumptions of the above proposition. In the setting of case (ℵ1) the function F belongs to the space $\mathcal{Z}$ if $F(\omega)$ is finite valued, i.e., for every $\omega \in \Omega$ it follows that $\inf_{x\in X} f(x,\omega) > -\infty$ and there is $\bar{x} \in X$ such that $f(\bar{x},\omega) < \infty$. Also in that case the space $\mathcal{Z}$ is finite dimensional. Consequently if the functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is convex, then it is continuous. Existence of the corresponding sequence χ_k holds automatically.

In case (ℵ3) suppose that the set X is a compact metric space and $f(x,\omega)$ is finite valued and continuous on $X \times \Omega$. Then $F(\omega)$ is finite valued and continuous, and hence F belongs to the space $C(\Omega)$. Also in that case $f(x,\omega)$ attains its minimal value for every $\omega \in \Omega$, and hence there is no need for the assumption (ii).

In case (ℵ2) we need to verify that $F(\omega)$ is measurable and p-integrable for $p \in [1,\infty)$, and essentially bounded for $p = \infty$. Suppose that $X = \mathbb{R}^n$. It is said that function $f(x,\omega)$

is random lower semicontinuous if its epigraphical mapping is closed valued and measurable, [1, Definition 14.28] (in some publications, in particular in [1], such functions are called normal integrands). If $f(x,\omega)$ is random lower semicontinuous, then $F(\omega)$ is measurable, [1, Theorem 14.37]. The condition of p-integrability can be verified by ad hoc methods. In particular this holds if $F(\omega)$ is essentially bounded. Also if $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is convex and monotone, then it is continuous in the norm topology of the space $\mathcal{Z} = L_p(\Omega, \mathcal{F}, P)$, $p \in [1,\infty)$ (cf., [2, Proposition 3.1]).

Proposition 2.2 *In the setting of case ($\aleph$2), suppose that $X = \mathbb{R}^n$, $f(x,\omega)$ is random lower semicontinuous, $F \in \mathcal{Z}$ and $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is monotone and continuous at F. Then* (2.2) *and* (2.3) *hold. If moreover $\mathcal{R}$ is strictly monotone, then* (2.4) *holds as well.*

Proof. By the second part of Proposition 2.1 we only need to verify existence of a sequence $\chi_k \in \mathcal{X}$ such that f_{χ_k} converges to F. Consider $\varepsilon > 0$. By the definition (2.1) of function F, for a.e. $\omega \in \Omega$ there is $\bar{\chi}(\omega) \in X$ such that $f(\bar{\chi}(\omega), \omega) < F(\omega) + \varepsilon$. Moreover $\bar{\chi}$ can be chosen in such a way that $f(\bar{\chi}(\cdot), \cdot)$ is measurable. Indeed, since $f(x,\omega)$ is random lower semicontinuous and hence its epigraphical mapping $\omega \mapsto \operatorname{epi} f(\cdot,\omega) \subset \mathbb{R}^n \times \mathbb{R}$ is closed valued and measurable, it follows by Castaing representation that there is a countable family of measurable mappings $(\chi^\nu, \alpha^\nu) : \Omega \to \mathbb{R}^n \times \mathbb{R}$, $\nu \in \mathbb{N}$, such that for every $\omega \in \Omega$ the set $\{(\chi^\nu(\omega), \alpha^\nu(\omega))\}$ is dense in $\operatorname{epi} f(\cdot,\omega)$, [1, Theorem 14.5]. Consider sets

$$A^\nu := \{\omega \in \Omega : f(\chi^\nu(\omega), \omega) < F(\omega) + \varepsilon\}.$$

It follows that the sets A^ν are measurable and $\cup_{\nu \in \mathbb{N}} A^\nu = \Omega$. Some of these sets can be empty. Define $\bar{\chi}(\omega)$ in the recursive way: $\bar{\chi}(\omega) := \chi^1(\omega)$ for $\omega \in A^1$, and $\bar{\chi}(\omega) := \chi^\nu(\omega)$ for $\omega \in A^\nu \setminus (\cup_{\iota=1}^{\nu-1} A_\iota)$ for $\nu = 2, \dots$.

Now let ε_k be a sequence of positive numbers converging to zero and $\chi_k(\omega)$ be measurable mappings such that

$$f(\chi_k(\omega), \omega) < F(\omega) + \varepsilon_k, \quad \omega \in \Omega. \tag{2.6}$$

By the definition of $F(\omega)$ we also have that $f(\chi_k(\omega), \omega) \geq F(\omega)$. Since $F \in \mathcal{Z}$ and hence is p-integrable, it follows from (2.6) that f_{χ_k} is also p-integrable and hence $f_{\chi_k} \in \mathcal{Z}$. It also follows from (2.6) that f_{χ_k} converges to F in the norm topology of $\mathcal{Z}$. ∎

As the following examples show the *strict* monotonicity condition is essential to ensure the implication (2.4).

Example 1 Consider the setting of case ($\aleph$1) and let $\mathcal{R}(Z) := \sum_{i=1}^m p_i Z(\omega_i)$, where p_i are nonnegative numbers such that $\sum_{i=1}^m p_i = 1$. The functional $\mathcal{R}$ can be viewed as the expectation operator $\mathcal{R} = \mathbb{E}$ associated with probabilities $p_i \geq 0$. This functional is monotone and continuous. The equation (2.2) takes here the form

$$\mathbb{E}\left[\inf_{x \in X} f(x,\omega)\right] = \inf_{\chi \in \mathcal{X}} \mathbb{E}[f(\chi(\omega), \omega)]. \tag{2.7}$$

If all $p_i > 0$, $i = 1, ..., m$, then $\mathcal{R} = \mathbb{E}$ is strictly monotone and both implications (2.3) and (2.4) follow.

Suppose now that one of the probabilities p_i is zero, say $p_1 = 0$. In that case $\mathbb{E}[f(\chi(\omega), \omega)]$ does not depend on $\chi(\omega_1)$ and hence $\bar{\chi}(\omega_1)$ can be any element of the set X in the left hand side of (2.4), provided that such minimizer $\bar{\chi}$ does exist. Hence there is no guarantee that $\bar{\chi}(\omega_1) \in \arg\min_{x \in X} f(x, \omega_1)$ and the implication (2.4) can be false. Of course here the probability of the event $\{\omega_1\}$ is zero, and the implication (2.4) becomes correct if the right hand side of (2.4) is understood to hold w.p.1. In the setting of case ($\aleph 2$) the expectation operator is strictly monotone. $\square$

Example 2 Consider the setting of case ($\aleph 3$) and let $\mathcal{R}(Z) := \max_{\omega \in \Omega} Z(\omega)$. This functional is monotone and continuous, but is not strictly monotone. The equality (2.2) takes here the form

$$\max_{\omega \in \Omega} \underbrace{\inf_{x \in X} f(x, \omega)}_{F(\omega)} = \inf_{\chi \in \mathcal{X}} \underbrace{\max_{\omega \in \Omega} f(\chi(\omega), \omega)}_{\mathcal{R}(f_\chi)} \tag{2.8}$$

and the implication (2.3) becomes

$$\bar{\chi}(\cdot) \in \arg\min_{x \in X} f(x, \cdot) \Rightarrow \bar{\chi} \in \arg\min_{\chi \in \mathcal{X}} \left\{ \max_{\omega \in \Omega} f(\chi(\omega), \omega) \right\}. \tag{2.9}$$

The converse of the implication (2.9) does not need to hold here. Similar conclusion follows in the setting of case ($\aleph 2$) with $\mathcal{Z} = L_\infty(\Omega, \mathcal{F}, P)$ and $\mathcal{R}(Z) := \operatorname{ess\,sup}_{\omega \in \Omega} Z(\omega)$. $\square$

Recall that in the setting of case ($\aleph 2$) the dual of space $\mathcal{Z} = L_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty)$, is the space $\mathcal{Z}^* = L_q(\Omega, \mathcal{F}, P)$, $q \in (1, \infty]$, $1/p + 1/q = 1$. If functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is convex continuous, then its subdifferential $\partial \mathcal{R}(Z) \subset \mathcal{Z}^*$ is nonempty for every $Z \in \mathcal{Z}$. The space $L_\infty(\Omega, \mathcal{F}, P)$ is paired with $L_1(\Omega, \mathcal{F}, P)$. For convex functional $\mathcal{R}$ we have the following characterization of monotonicity (cf., [4]).

Proposition 2.3 *In the setting of case ($\aleph 2$), a convex continuous functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is monotone iff for every $Z \in \mathcal{Z}$ and $\gamma \in \partial \mathcal{R}(Z)$ it follows that $\gamma \succeq 0$. Moreover, $\mathcal{R}$ is strictly monotone iff for every $Z \in \mathcal{Z}$ and $\gamma \in \partial \mathcal{R}(Z)$ it follows that $\gamma(\omega) > 0$ for a.e. $\omega \in \Omega$.*

In the setting of case ($\aleph 3$) the dual of space $\mathcal{Z} = C(\Omega)$ is the space $\mathcal{Z}^*$ of finite signed measures with the corresponding scalar product

$$\langle \mu, Z \rangle = \int_\Omega Z(\omega) d\mu(\omega), \; Z \in \mathcal{Z}, \mu \in \mathcal{Z}^*.$$

In that case we have the following characterization of monotonicity.

Proposition 2.4 *In the setting of case ($\aleph 3$), a convex continuous functional $\mathcal{R} : \mathcal{Z} \to \mathbb{R}$ is monotone iff for any $Z \in \mathcal{Z}$ and $\mu \in \partial \mathcal{R}(Z)$ it follows that the measure μ is nonnegative, i.e., $\mu(A) \geq 0$ for every $A \in \mathcal{F}$. Moreover, $\mathcal{R}$ is strictly monotone iff for every $\mu \in \partial \mathcal{R}(Z)$ it follows that $\mu(A) > 0$ for every open set $A \subset \Omega$.*

Proof. Consider the conjugate of $\mathcal{R}$,

$$\mathcal{R}^*(\zeta) = \sup_{Z \in \mathcal{Z}} \langle \zeta, Z \rangle - \mathcal{R}(Z), \; \zeta \in \mathcal{Z}^*,$$

and its domain $\mathrm{dom}(\mathcal{R}^*) = \{\zeta \in \mathcal{Z}^* : \mathcal{R}^*(\zeta) < \infty\}$. We have that $\mathcal{R}$ is monotone iff every $\zeta \in \mathrm{dom}(\mathcal{R}^*)$ is nonnegative (cf., [3, Theorem 6.5]). Since $\mathcal{R}$ is convex continuous its subdifferential $\partial \mathcal{R}(Z)$ is nonempty for all $Z \in \mathcal{Z}$ and

$$\partial \mathcal{R}(Z) = \arg\max_{\zeta \in \mathbb{R}^m} \{\langle \zeta, Z \rangle - \mathcal{R}^*(\zeta)\},$$

and hence $\partial \mathcal{R}(Z) \subset \mathrm{dom}(\mathcal{R}^*)$. It follows that if $\mathcal{R}$ is monotone, then every $\mu \in \partial \mathcal{R}(Z)$ is nonnegative.

Consider a subgradient $\mu \in \partial \mathcal{R}(Z)$. Then for any $Z, Z' \in \mathcal{Z}$,

$$\mathcal{R}(Z') \geq \mathcal{R}(Z) + \langle \mu, Z' - Z \rangle. \tag{2.10}$$

It follows that if $Z' \succeq Z$ and μ is nonnegative, then $\langle \mu, Z' - Z \rangle \geq 0$, and hence $\mathcal{R}(Z') \geq \mathcal{R}(Z)$. Moreover, if $Z \neq Z'$, then since $Z, Z' : \Omega \to \mathbb{R}$ are continuous, there exist an open set $A \subset \Omega$ and $\varepsilon > 0$ such that $Z'(\omega) \geq Z(\omega) + \varepsilon$ for all $\omega \in A$. Hence if $\mu(A) > 0$, then the strict inequality $\mathcal{R}(Z') > \mathcal{R}(Z)$ follows.

Conversely suppose that $\mathcal{R}$ is strictly monotone. We argue by a contradiction. Suppose that there is a nonnegative measure $\mu \in \partial \mathcal{R}(Z)$ and open set $A \subset \Omega$ such that $\mu(A) = 0$. Let $\bar{Z} \in \mathcal{Z}$ be such that $\bar{Z}(\omega) > 0$ for all $\omega \in A$ and $\bar{Z}(\omega) = 0$ for all $\omega \in \Omega \setminus A$. Consider $Z' := Z - \bar{Z}$. We have by (2.10) and since $\langle \mu, Z' - Z \rangle = -\langle \mu, \bar{Z} \rangle = 0$, that $\mathcal{R}(Z') \geq \mathcal{R}(Z)$. On the other hand, $Z \succeq Z'$ and hence $\mathcal{R}(Z) \geq \mathcal{R}(Z')$. It follows that $\mathcal{R}(Z') = \mathcal{R}(Z)$. Since $Z' \neq Z$ this contradicts strict monotonicity of $\mathcal{R}$. This competes the proof. ∎

Example 3 In the setting of case ($\aleph$3), let $\mathfrak{P}$ be the set of probability measures on the sample space $(\Omega, \mathcal{F})$ and consider a set $\mathfrak{M} \subset \mathfrak{P}$ defined by the moment constraints

$$\mathfrak{M} := \{\mu \in \mathfrak{P} : \langle \mu, \psi_i \rangle = b_i, \; i = 1, ..., q\},$$

where $\psi_i : \Omega \to \mathbb{R}$, $i = 1, ..., q$, are measurable functions. Suppose that the set $\mathfrak{M}$ is nonempty and define functional

$$\mathcal{R}(Z) := \sup_{\mu \in \mathfrak{M}} \langle \mu, Z \rangle. \tag{2.11}$$

Note that for any $\mu \in \mathfrak{P}$ and $Z \in \mathcal{Z}$ it follows that $\langle \mu, Z \rangle \leq \|Z\|$, and hence the functional $\mathcal{R}$ is finite valued. By Richter-Rogosinski Theorem the maximum in the right hand side of (2.11) is attained at a probability measure supported on a finite set of no more than $q+1$ points. Since $\partial \mathcal{R}(Z) = \arg\max_{\mu \in \mathfrak{M}} \langle \mu, Z \rangle$, the functional $\mathcal{R}$ has a subgradient $\mu \in \partial \mathcal{R}(Z)$ supported on a finite set of no more than $q+1$ points. It follows by Proposition 2.4 that $\mathcal{R}$ is monotone, but is not strictly monotone if the set Ω has more than $q+1$ points, in particular if the set Ω is not finite. □

3 Dynamic equations

3.1 Two-stage problems

Consider the following two stage stochastic programming problem

$$\min_{x\in X} \big\{ f(x) := \mathcal{R}(F_x) \big\}, \tag{3.1}$$

where $F_x(\omega) = F(x,\omega)$ is the optimal value of the second stage problem

$$\min_{y\in \mathcal{G}(x,\omega)} g(x,y,\omega), \tag{3.2}$$

with $X \subset \mathbb{R}^n$, $g : \mathbb{R}^n \times \mathbb{R}^k \times \Omega \to \mathbb{R}$ and $\mathcal{G} : \mathbb{R}^n \times \Omega \rightrightarrows \mathbb{R}^k$ being a multifunction. We assume that $F_x \in \mathcal{Z}$ for all $x \in X$, and hence $\mathcal{R}(F_x)$ is well defined. In particular this implies that $\mathcal{G}(x,\cdot)$ is nonempty for all $x \in X$, i.e., that the problem has relatively complete recourse. For a thorough discussion of measurability of the optimal value function $F(x,\omega)$ we can refer to [1, chapter 14(D)].

Together with (3.1) consider the following formulation

$$\min_{x\in X,\, \eta(\cdot)\in \mathcal{G}(x,\cdot)} \mathcal{R}[g(x,\eta(\omega),\omega)]. \tag{3.3}$$

We can write F_x in the form

$$F(x,\omega) = \inf_{y\in\mathbb{R}^k} \bar{g}(x,y,\omega), \tag{3.4}$$

where

$$\bar{g}(x,y,\omega) := \begin{cases} g(x,y,\omega), & \text{if } y \in \mathcal{G}(x,\omega), \\ +\infty, & \text{otherwise.} \end{cases}$$

Suppose that $\mathcal{R}$ is monotone and either the corresponding assumptions (i) or (ii) of Proposition 2.1 hold. Then

$$\mathcal{R}(F_x) = \inf_{\eta(\cdot)\in \mathcal{G}(x,\cdot)} \mathcal{R}[g(x,\eta(\omega),\omega)], \quad x \in X, \tag{3.5}$$

and hence the optimal values of problems (3.1) and (3.3) are equal to each other. Also it follows by (2.3) that if $\bar{x}$ is an optimal solution of the first stage problem (3.1) and

$$\bar{\eta}(\omega) \in \arg\min_{y\in \mathcal{G}(\bar{x},\omega)} g(\bar{x},y,\omega), \quad \omega \in \Omega, \tag{3.6}$$

then $(\bar{x},\bar{\eta}(\cdot))$ is an optimal solution of problem (3.3). Moreover, if $\mathcal{R}$ is strictly monotone, then by (2.4) we have that $(\bar{x},\bar{\eta}(\cdot))$ is an optimal solution of (3.3) iff $\bar{x}$ is a solution of the first stage problem (3.1) and condition (3.6) holds.

Remark 3.1 Without *strict* monotonicity it could happen that (3.3) has an optimal solution $(\bar{x},\bar{\eta}(\cdot))$ such that $\bar{\eta}$ does not satisfy (3.6). Such solution is not time consistent in the sense that $\bar{\eta}$ is not optimal for the second stage problem (3.2) conditional on $x = \bar{x}$ (cf., [4]). We demonstrate this in Example 4 below.

Example 4 Consider the setting of case (ℵ3) with Ω being a compact subset of $\mathbb{R}$ such that $\sup(\Omega) = 1$. As in Example 2 define $\mathcal{R}(Z) := \max_{\omega\in\Omega} Z(\omega)$. Recall that this functional is monotone and continuous, but is not strictly monotone. Suppose further that $X := \mathbb{R}_+$ and

$$F(x, \omega) = \inf\{y \in \mathbb{R} : y \geq x + \omega\}.$$

That is, $F(x, \omega) = x + \omega$ and $\mathcal{R}(F_x) = \max_{\omega\in\Omega}(x + \omega) = x + 1$. Hence problem (3.1) has optimal value 1 and optimal solution $\bar{x} = 0$. Here the corresponding minimum is attained and

$$\mathcal{R}(F_x) = \inf_{\eta\in C(\Omega)} \left\{ \max_{\omega\in\Omega} \eta(\omega) : \eta(\omega) \geq x + \omega \right\},$$

and the optimal values of the respective problems (3.1) and (3.3) are equal to each other. The set of optimal solutions of problem (3.3) is given by $\bar{x} = 0$ and any continuous function $\hat{\eta} : \Omega \to \mathbb{R}$ such that $\hat{\eta}(\omega) \geq \omega$ for all $\omega \in \Omega$ and $\max_{\omega\in\Omega} \hat{\eta}(\omega) = 1$. On the other hand, for $\bar{x} = 0$ the solution $\bar{\eta}(\cdot)$ given by (3.6), is the unique function $\bar{\eta}(\omega) = \omega$, $\omega \in \Omega$.

This can happen even if the set Ω is finite. For example, suppose that $\Omega = \{\omega_1, \omega_2\}$ with $\omega_1 = 0$, $\omega_2 = 1$. We have that $\bar{x} = 0$, $\bar{y}_1 = 1/2$, $\bar{y}_2 = 1$ is an optimal solution of the problem

$$\min_{x,y_1,y_2} x + \max\{y_1, y_2\} \text{ s.t. } x \geq 0,\ y_1 \geq x,\ y_2 \geq x + 1. \tag{3.7}$$

However, $\bar{y}_1$ is not optimal for the second stage problem $\min\limits_{y\geq x+\omega} y$, for $x = \bar{x}$ and $\omega = \omega_1$. That is, problem (3.7) possesses optimal solution $(\bar{x}, \bar{y}_1, \bar{y}_2)$ which does not satisfy the dynamic equations and is not time consistent. $\square$

3.2 Multistage problems

Let us discuss now risk averse multistage stochastic programs. We follow here the notation of [3, Section 6.8.1], where more detailed descriptions and results can be found. Consider a sequence of nested sigma algebras $\mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T$ with $\mathcal{F}_T = \mathcal{F}$ and $\mathcal{F}_1 = \{\emptyset, \Omega\}$ being trivial. Let $\mathcal{Z}_t$ be a linear space of $\mathcal{F}_t$-measurable functions $Z : \Omega \to \mathbb{R}$ and $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$, $t = 2, ..., T$. Since $\mathcal{F}_1$ is trivial, the space $\mathcal{Z}_1$ can be identified with $\mathbb{R}$. For mappings ρ_t the concepts of monotonicity and strict monotonicity are defined in the same way as for real valued functionals considered in Section 2. We assume that mappings ρ_t, $t = 2, ..., T$, satisfy the following condition of *translation equivariance*:

$$\rho_t(Z + Y) = \rho(Z) + Y, \quad Z \in \mathcal{Z}_t, Y \in \mathcal{Z}_{t-1}. \tag{3.8}$$

Consider the composite function $\bar{\rho} := \rho_2 \circ \cdots \circ \rho_T : \mathcal{Z}_T \to \mathbb{R}$. By translation equivariance we have that for $Z_t \in \mathcal{Z}_t$, $t = 1, ..., T$,

$$\bar{\rho}(Z_1 + ... + Z_T) = Z_1 + \rho_2\Big[Z_2 + ... + \rho_{T-1}\big[Z_{T-1} + \rho_T[Z_T]\big]\Big]. \tag{3.9}$$

Consider now the following multistage problem

$$\min_{\pi\in\Pi} \bar{\rho}\left[f_1(x_1) + f_2(x_2(\omega), \omega) + ... + f_T(x_T(\omega), \omega)\right]. \tag{3.10}$$

The minimization in (3.10) is performed over the set Π of policies $\pi = (x_1, x_2(\omega), ..., x_T(\omega))$ adapted to filtration $\mathfrak{F} = (\mathcal{F}_1, ..., \mathcal{F}_T)$ and satisfying the feasibility constraints

$$x_1 \in \mathcal{X}_1,\ x_t(\cdot) \in \mathcal{X}_t(x_{t-1}(\cdot), \cdot),\ t = 2, ..., T.$$

With problem (3.10) are associated dynamic programming equations

$$Q_t(x_{t-1}, \omega) = \inf_{x_t \in \mathcal{X}_t(x_{t-1}, \omega)} \Big\{ f_t(x_t, \omega) + \rho_{t+1}\left[Q_{t+1}(x_t, \omega)\right] \Big\}, \tag{3.11}$$

$t = 2, ..., T$, with the term $\rho_{T+1}\left[Q_{t+1}(x_t, \omega)\right]$ at the last stage omitted, and the first stage problem

$$\min_{x_1 \in \mathcal{X}_1} f_1(x_1) + \rho_2[Q_2(x_1)]. \tag{3.12}$$

For $T = 2$ problem (3.10) takes the form[1]

$$\min_{x_1 \in \mathcal{X}_1, x_2(\cdot) \in \mathcal{X}_2(x_1, \cdot)} \rho_2\left[f_1(x_1) + f_2(x_2(\omega), \omega)\right]. \tag{3.13}$$

This problem is of the form (3.3) with $g(x, y, \omega) = f_1(x) + f_2(y, \omega)$. Equations (3.1) and (3.2) can be considered as the dynamic programming equations for this two stage problem.

For the sake of simplicity, in order to avoid delicate questions of measurability etc, we discuss below the case of finite set Ω with sigma algebra $\mathcal{F}$ of all its subsets, and spaces $\mathcal{Z}_t$ consisting of all $\mathcal{F}_t$-measurable functions $Z : \Omega \to \mathbb{R}$. In that case the filtration $\mathfrak{F}$ can be represented by a finite scenario tree (cf., [3, Section 6.8.1]). We make the following assumption.

(A) The cost-to-go functions $Q_t(x_{t-1}, \cdot)$, $t = 2, ..., T$, are finite valued for all x_{t-1} satisfying the feasibility constraints, so that the corresponding functionals in (3.11) are well defined.

Proposition 3.1 *Suppose that assumption* (A) *is fulfilled, the mappings ρ_t, $t = 2, ..., T$, are monotone, continuous and translation equivariant. Then:* (i) *the optimal values of problems* (3.10) *and* (3.12) *are equal to each other,* (ii) *if*

$$\bar{x}_t(\cdot) \in \arg \min_{x_t \in \mathcal{X}_t(\bar{x}_{t-1}, \cdot)} \Big\{ f_t(x_t, \cdot) + \rho_{t+1}\left[Q_{t+1}(x_t, \cdot)\right] \Big\},\ t = 1, ..., T, \tag{3.14}$$

then $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T(\omega))$ is an optimal solution of problem (3.10), (iii) *if moreover mappings ρ_t, $t = 2, ..., T$, are strictly monotone and $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T(\omega))$ is an optimal solution of problem* (3.10)*, then condition* (3.14) *follows.*

[1] By definition $\mathcal{X}_1(x_0, \cdot) \equiv \mathcal{X}_1$.

Proof. For fixed $x_1, ..., x_{T-1}(\cdot)$ consider minimization in (3.10) with respect to the last component $x_T(\cdot)$:

$$\begin{array}{ll} \min\limits_{x_T(\cdot)} & f_1(x_1) + \rho_2\Big[f_2(x_2(\omega),\omega) + \cdots + \rho_{T-1}\big[f_{T-1}(x_{T-1}(\omega),\omega) + \rho_T[f_T(x_T(\omega),\omega)]\big]\Big] \\ \text{s.t.} & x_T(\cdot) \in \mathcal{X}_T(x_{T-1}(\cdot),\cdot). \end{array} \tag{3.15}$$

Consider the problem

$$\min_{x_T(\cdot)} \rho_T[f_T(x_T(\omega),\omega)] \text{ s.t. } x_T(\cdot) \in \mathcal{X}_T(x_{T-1}(\cdot),\cdot), \tag{3.16}$$

conditional on $\mathcal{F}_{T-1}$. By "conditional on $\mathcal{F}_{T-1}$" we mean that we consider (3.16) pointwise for elementary events of the sigma algebra $\mathcal{F}_{T-1}$. That is, let Ω be partitioned into union of disjoint sets A_i, $i = 1, ..., r$, such that $Z : \Omega \to \mathbb{R}$ is $\mathcal{F}_{T-1}$-measurable iff $Z(\omega)$ is constant on each A_i. Conditional on $\omega \in A_i$ we can view $\rho_T : \mathcal{Z}_T \to \mathcal{Z}_{T-1}$ as real valued. We can apply now Proposition 2.1 to conclude that conditional on $\omega \in A_i$, $i = 1, ..., r$, the optimal value of problem (3.16) is equal to $\rho_T[Q_T(x_{T-1}(\omega),\omega)]$. Note that since $x_{T-1}(\cdot)$ is $\mathcal{F}_{T-1}$-measurable, it is constant on every set A_i. It follows that the optimal value of problem (3.10) is equal to the optimal value of

$$\min f_1(x_1) + \rho_2\Big[f_2(x_2(\omega),\omega) + ... + \rho_{T-1}\big[f_{T-1}(x_{T-1}(\omega),\omega) + \rho_T[Q_T(x_{T-1}(\omega),\omega)]\big]\Big], \tag{3.17}$$

where the minimization is performed over $x_1 \in \mathcal{X}_1, ..., x_{T-1}(\cdot) \in \mathcal{X}_{T-1}(x_{T-2}(\cdot),\cdot)$. We can apply then the same interchange procedure to $f_{T-1}(x_{T-1}(\omega),\omega) + \rho_T[Q_T(x_{T-1}(\omega),\omega)$ with respect to ρ_{T-1} and $x_{T-1}(\omega)$, and so on going backward in time. Eventually we obtain that the optimal value of problem (3.10) is equal to the optimal value of the first stage problem (3.12).

Also by (2.3) we can conclude that if (3.14) holds, then $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T(\omega))$ is an optimal solution of problem (3.10). If moreover mappings ρ_t, $t = 2, ..., T$, are strictly monotone, then by (2.4) the converse implication holds as well. This completes the proof. ∎

Remark 3.2 Without the strict monotonicity condition the converse implication (iii) of the above proposition does not necessarily holds if problem (3.10) has more than one optimal solution. This was already demonstrated in Example 4 for a two stage problem (see also [4] for such simple example). That is, in order to ensure that an optimal policy satisfies the dynamic programming equations, the strict monotonicity condition is essential. In the setting of case ($\aleph$2) the conditional expectation mappings $\rho_t := \mathbb{E}_{|\mathcal{F}_t}$ are strictly monotone. Therefore in the risk neutral case a policy π is optimal for problem (3.10), with $\bar{\rho} = \mathbb{E}$, iff it satisfies the respective dynamic programming equations.

Acknowledgements
This research was partly supported by NSF grant 1633196 and DARPA EQUiPS program, grant SNL 014150709.

References

[1] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[2] A. Ruszczyński and A. Shapiro, Optimization of convex risk functions, *Mathematics of Operations Research*, 31 (2006), 433–452.

[3] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, second edition, SIAM, Philadelphia, 2014.

[4] A. Shapiro and A. Pichler, Time and Dynamic Consistency of Risk Averse Stochastic Programs, Published electronically in: Optimization Online, 2016.