TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

Diplomarbeit

Distribution-Constrained Optimal Stopping Problems in Discrete Time

Ausgeführt am Institut für
Stochastik und Wirtschaftsmathematik
der Technischen Universität Wien

unter der Anleitung von
Univ. Prof. Dipl.-Ing. Dr.techn. Mathias BEIGLBÖCK

durch

Gudmund PAMMER
Blumauergasse 20/3, 1020 Wien

September 10, 2017

Abstract

This thesis is focused on a more general type of optimal stopping problems in discrete time. Varying approaches of viewing this problem are discussed and introduced, e.g. using a space of couplings under linear constraints or so-called adapted random probability measures. A connection between these views is made and existence of an optimal solution is shown. Further, a modified version of Monge-Kantorovich duality is established. The final sections show a monotonicity principle with examples. For a special class of cost functions, optimality (and uniqueness) of a "greedy strategy" is established. In particular, the proof resembles the main idea behind a monotonicity principle for discrete time, which in turn is based on a monotonicity principle for continuous time. Finally, optimality of the "greedy strategy" is shown using monotonicity.

Zusammenfassung

Die vorliegende Arbeit hat eine verallgemeinerte Version eines "optimal stopping"-Problems in diskreter Zeit als Hauptfokus. Unterschiedliche Herangehensweisen an dieses Problem werden vorgezeigt und besprochen, wie das Verwenden von einem Raum von "couplings", welche zusätzlich lineare Nebenbedingungen erfüllen, oder jenes von sogenannten "adapted random probability measures". Weiters wird eine Verbindung dieser Sichtweisen aufgezeigt und die Existenz einer optimalen Lösung bewiesen. In den abschließenden Kapiteln wird ein Monotonie-Prinzip anhand eines Beispiels vorgeführt. Für eine spezielle Klasse an Kostenfunktionen wird Optimalität (und Eindeutigkeit) einer "greedy"-Strategy gezeigt. Der Beweis basiert stark auf jener Idee, die hinter einem Monotonie-Prinzip in diskreter Zeit steckt, welche wiederum von einem Monotonie-Prinzip in stetiger Zeit abgeleitet wurde. Zuletzt wird Optimalität auch unter Verwendung dieses Monotonie-Prinzips gezeigt.

Acknowledgments

I would like to thank my supervisor, Prof. Mathias Beiglböck, for guiding and encouraging me - and, especially, for giving me the opportunity to work on this topic.

I would like to thank all my colleagues and friends at TU Wien for making studying mathematics an unforgettable experience, the countless hours spent in Resselpark and strolling through Vienna having highly enjoyable conversations. These experiences did not only help me to improve scientifically but also on a personal level.

Last but not least I want to thank my girlfriend Rebecca and my family for the constant support and encouragement throughout the years.

Danksagung

Zuerst will ich mich bei meinem Betreuer, Prof. Mathias Beiglböck, für die Unterstützung und Leitung bedanken - besonders für die Möglichkeit, überhaupt an diesem Thema arbeiten zu können.

Ich will mich bei all meinen Kollegen und jahrelangen Freunden an der TU Wien dafür bedanken, dass sie das Mathematik-Studium für mich zu einem unbeschreiblichen Erlebnis gemacht haben. Die unzähligen Stunden, welche wir, im Resselpark und beim Spazieren durch Wien, mit spannenden Gesprächen verbracht haben, waren eine unvergessliche Erfahrung, welche mir nicht nur half, mich im wissenschaftlichen Bereich zu verbessern, sondern auch auf einer persönlichen Ebene.

Zu guter Letzt bedanke ich mich bei meiner Freundin Rebecca und meiner ganzen Familie für die kontinuierliche Unterstützung und Ermutigungen.

Contents

CHAPTER 1

Introduction

In unit-linked life insurances - e.g. the life insurance of a married couple - the physical and emotional health can strongly depend on one of the partners. As a consequence, health can drastically deteriorate when one's partner dies. Therefore, it is not reasonable to assume independence of the times of death of either partner. An attempt of modeling this dependency can be made by using measures related to a stochastic process which mimick ordinary stopping times and are, in fact, a generalization of the latter. An informal way of posing this problem is the following:

Given a payoff function c, which may depend on the values of the stochastic process up to a point in time t, we seek to maximize

$$\tau \mapsto \mathbb{E}[c((Z_t)_{t\leq\tau}, \tau)]$$

where τ is not an ordinary stopping time in the filtration generated by Z, i.e., it does not stop the process at one time $\tau(\omega)$, but rather $\tau(\omega)$ is a sub-probability measure on the time domain by itself. Essentially this can be formalized in three different ways:

1. As an optimal stopping problem where adapted random probability measures are used instead of ordinary stopping times.
2. As an optimal transport problem by reformulating it by means of randomized stopping times.
3. As an ordinary optimal stopping problem on a larger probability space, cf. [7, Lemma 3.11].

Outline of Thesis

In Chapter 2 the maximization problems ($\textsc{OptStop}^{\gamma}$) and ($\textsc{OptStop}^{\pi}$) are formally introduced. Section 2.2 introduces the notions of couplings (Cpl) and randomized stopping times (RST), and explores their relation. The subsequent Section 2.3 draws the connection between ($\textsc{OptStop}^\gamma$) and ($\textsc{OptStop}^\pi$). Existence of a maximizer of ($\textsc{OptStop}^\pi$) is shown in Section 3.1 utilizing Prokhorov's Theorem and [7, Lemma 2.3]. Based on the theory of optimal transport and recent results [12] in this area, duality in the sense of Kantorovich is deduced. In Chapter 4 examples are investigated and maxmizers are determined. Finally, Chapter 5 shortly sketches different monotonicity principles and uses one to show optimality of the maximizer introduced in the previous chapter.

Notational Conventions

- In this paper, we consider a discrete time domain. Its index set is denoted by I. Typical examples are $I := \mathbb{N}$ for an infinite time horizon and $I := \{1, \dots, T\}$, $T \in \mathbb{N}$ for a finite index set. If the reader is interested in the time continuous case, they may be referred to [2] and [6]. For $t \in I$ we define the set $I_{<t} := \{s \in I | s < t\}$ of all times before t, the set $I_{\leq t} := \{s \in I | s \leq t\}$ of all times up to t, the set $I_{\geq t} := \{s \in I | s \geq t\}$ of all times from t on, and the set $I_{>t} := \{s \in I | s > t\}$ of all times after t.

- Given a topological space $(X, \mathcal{T})$, we denote its Borel-σ-algebra with $\mathcal{B}(X) = \sigma(\mathcal{T})$, the interior of a set $A \subseteq X$ with $\operatorname{int}(A)$ and its boundary with $\partial(A)$. The space of all Borel-measurable functions from X into $\mathbb{R}$ are denoted by $B(X)$ and its subspace of all bounded, Borel-measurable functions by $B_b(X)$.

- Typically, we will work with (sub-)probability measures on the Polish space $\mathbb{R}^I$. To facilitate the notation of projections onto particular subspace of $\mathbb{R}^I$, which we may define as
$$\mathbb{R}^I =: \prod_{i \in I} X_i,$$
and for instance call the projection of a measure μ on $\mathbb{R}^I$ onto the first component $\operatorname{proj}_{X_1}(\mu)$.

- Several different notations will be used to refer to elements of $\mathbb{R}^I$. For any vector $\omega \in \mathbb{R}^I$, its entries are denoted with
$$\omega = (\omega_t)_{t \in I} = (\omega_1, \omega_2, \dots).$$
Parts of the vector (path) ω will be referred to by
$$(\omega_t)_{t \in I_{\leq s}} = \omega_{\upharpoonright [0,s]}, \quad (\omega_t)_{t \in I_{>s}} = \omega_{\upharpoonright (s,T]}, \quad s \in I,$$
where $\omega_{\upharpoonright J}$ with $J \subseteq I$ stands for the restriction of ω onto $\mathbb{R}^J$.

- If $\omega \in \mathbb{R}^I$, $s \in I$ and $\theta \in \mathbb{R}^{I_{>s}}$, we may use $\oplus$ to indicate the concatenation of the paths $\omega_{\upharpoonright [0,s]}$ and θ, such that
$$\omega_{\upharpoonright [0,s]} \oplus \theta := (\omega_1, \dots, \omega_s, \omega_s + \theta_1, \omega_s + \theta_2, \dots) \in \mathbb{R}^I.$$

- In the following $Z = (Z_t)_{t \in I}$ will denote a distinguished stochastic process. If Z is assumed to have independent increments, i.e., for any $t_1, \dots, t_n \in I$ with $t_1 < \dots < t_n$ the increments $Z_{t_1}, Z_{t_2} - Z_{t_1}, \dots, Z_{t_n} - Z_{t_{n-1}}$ are independent, it is convenient to define $(p_i)_{i \in I}$ via

$$Z_t = Z_0 + \sum_{i \le t} p_i,$$

  where Z_0 is the initial distribution of the stochastic process Z. The measure induced by the process starting in 0, $\tilde{Z}_t := \sum_{i \le t} p_i$, on $\mathbb{R}^I$ is denote by $\mathbb{P}$.

- For signed measures ξ there exists a Hahn-Jordan decomposition,

$$\xi = \xi^+ - \xi^-,$$

  where ξ^+ and ξ^- are the positive and negative parts of ξ, respectively.

CHAPTER 2

The Maximization Problem

Let $(\Omega, \mathcal{G}, \mathbb{G} := (\mathcal{G}_t)_{t\in I}, \mathbb{P})$ be an abstract filtered probability space and $Z := (Z_t)_{t\in I}$ be the stochastic, real-valued and $\mathbb{G}$-adapted process of interest. Further, let ν denote a (discrete) probability measure on $(I, \mathcal{B}(I))$. We assume that the process Z is uniformly integrable, i.e.,

$$\forall \epsilon > 0,\ \exists \delta > 0 : \quad \int_E |Z_t| \mathrm{d}\mathbb{P} < \epsilon \text{ whenever } Z_t \in L^1(\mathbb{P}) \text{ for all } t \in I \text{ and } \mathbb{P}(E) < \delta.$$

Furthermore, by μ we denote the probability measure induced onto the measurable space $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$ by the stochastic process Z via

$$\mu(B) := Z_\#\mathbb{P}(B) \quad \forall B \in \mathcal{B}(\mathbb{R}^I),$$

and call the probability triplet $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I), \mu)$ the path space of Z. The payoff function c is assumed to be real-valued and Borel-measurable on S with

$$S := \{(x, t) \mid x \in \mathbb{R}^I,\ t \in I\}.$$

The space S is adequate for our purposes since for a given time t and path $(Z_s((\omega))_{s\le t} =: x$ up to the time t, the function c returns the payoff $c(x, t)$. Note that the space is Polish as it is the direct sum of Polish spaces. For example, the topology induced on S by the metric $d : S \times S \to \mathbb{R}$ defined as

$$((x_i)_{i\le s}, s), (y_i)_{i\le t}, t)) \mapsto \max\left(|t-s|, \max_{i\le \min(s,t)}(|x_i - y_i|)\right),$$

causes (S, d) to be Polish. Further, there exists a surjective, open, continuous map r with

$$\begin{aligned} r\ :\ (\mathbb{R}^I \times I, \mathcal{B}(\mathbb{R}^I \times I)) &\to (S, d), \\ ((x_s)_{s\in I}, t) &\mapsto ((x_s)_{s\le t}, t), \end{aligned} \tag{2.1}$$

such that the topology on S is the final topology on S with respect to the map r. Note that the map r is Borel-measurable.

Adapted Random Probability Measures

Instead of restricting ourselves to $\mathbb{G}$-stopping times on $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, we introduce a generalization of the notion of $\mathbb{G}$-stopping times. Assume τ to be a $\mathbb{G}$-stopping time, then it can be naturally identified with a $\mathbb{G}$-adapted stochastic process $\gamma := (\gamma_t)_{t \in I}$ such that

$$\gamma_t(\omega) := \mathbb{1}_{\{t\}}(\tau(\omega)), \quad \omega \in \Omega,\ t \in I.$$

Thus, for a.e. ω the stochastic process γ defines a probability measure on I, which in turn tells us the probability of having already stopped at time t. This leads us to the notion of *adapted random probability measures*.

Definition 2.1 (Adapted Random Probability Measure)**.**
We call a real-valued, stochastic process $\gamma := (\gamma_t)_{t \in I}$ on $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ an adapted random probability measure, if

1. $\gamma_t \geq 0$ a.s. for all $t \in I$,
2. $\sum_{t \in I} \gamma_t = 1$ a.s.,
3. γ is $\mathbb{F}$-adapted.

The space of all these adapted random probability measures is denoted by $\mathcal{M}_I$.
Given a probability measure ν on I, we might be interested in all adapted random probability measures γ such that

4. $\mathbb{E}[\gamma_t] = \nu(t)$ for all $t \in I$.

The restriction of $\mathcal{M}_I$ to all adapted random probability measures which additionally satisfy (4) is denoted by $\mathcal{M}_I^\nu$.

As explained in the introduction, we want to maximize the expected payoff given a cost function $c \colon S \to \mathbb{R}$ and a stochastic process Z where the maximization is now taken over all adapted random probability measures, which continue along a given probability measure.

Problem (OptStop$^\gamma$)**.** Given a Borel-measurable payoff function $c : S \to \mathbb{R}$ and a probability measure ν on I, we seek to find a maximizer of

$$\gamma \mapsto \mathbb{E}\left[\sum_{t \in I} c((Z_s)_{s \leq t}, t)\gamma_t\right], \quad \gamma \in \mathcal{M}_I^\nu.$$

Remark 2.2. Note that this is an enlargement of the standard optimal stopping problem

$$\tau \mapsto \mathbb{E}[c((Z_s)_{s \leq \tau}, \tau)]$$

where τ is a $\mathbb{G}$-stopping time and $\mathbb{P}([\tau = t]) = \nu(t)$ for all $t \in I$. We denote the space of all $\mathbb{G}$-stopping times by $\mathcal{T}_I$ and its restriction to all stopping times τ such that $\mathcal{L}(\tau) = \nu$ with $\mathcal{T}_I^\nu$. Obviously it holds that

$$\sup_{\gamma \in \mathcal{M}_I^\nu} \mathbb{E}\left[\sum_{t \in I} c((Z_s)_{s \leq t}, t)\gamma_t\right] \geq \sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[c((Z_s)_{s \leq \tau}, \tau)],$$

because $\mathcal{T}_I^\nu$ is embedded in $\mathcal{M}_I^\nu$.

Randomized Stopping Times and Couplings

Rather than working with an abstract filtered probability space, it is possible to work on the path space of the stochastic process Z, $(\mathbb{R}^I, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t\in I}, \mu)$, where $\mathbb{F}$ is the natural filtration of Z and $\mathcal{F}$ the Borel-σ-algebra on $\mathbb{R}^I$. If $T < \infty$ the space $\mathbb{R}^I$, equipped with the product topology, is a complete metric space, where the metric can be chosen as

$$\max_{t\in I} |x_t - y_t|.$$

If we admit $T = \infty$, i.e., an infinite time horizon, the path space with the product topology remains Polish as a countable product of Polish spaces. Furthermore, a possible metric which induces the product topology is

$$\begin{aligned}\rho(x,y) \ : \ \mathbb{R}^I \times \mathbb{R}^I &\to \mathbb{R}\\ (x,y) &\mapsto \sum_{t\in I} \frac{1}{2^t} \frac{|x_t - y_t|}{1 + |x_t - y_t|}.\end{aligned}$$

As stated in the introduction, instead using stopping times which stop at one point in time, it is possible to generalize this with so-called randomized stopping times. A randomized stopping time π tries to mimic stopping times by assigning almost every path ω a probability measure π_ω on I which again tells us the probability with which we stop at time t.

Definition 2.3 (Randomized Stopping Time)**.**
A probability measure π on $\mathbb{R}^I \times I$ is called randomized stopping time, if

1. $\text{proj}_{\mathbb{R}^I}(\pi) = \mu$,
2. the mapping $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$, where $(\pi_\omega)_{\omega\in\mathbb{R}^I}$ is a disintegration of π. Or equivalently, the with π associated process $A := (A_t)_{t\in I}$, where $A_t(\omega) := \sum_{s\le t} \pi_\omega(s)$ is $\mathcal{F}_t$-measurable.

Again, the space of all randomized stopping times on $\mathbb{R}^I \times I$ which satisfy (1) and (2) are denoted by $\text{RST}(\mu)$. Given a probability measure ν on I, we are interested in random stopping times π such that

3. $\text{proj}_I(\pi) = \nu$.

The restriction of $\text{RST}(\mu)$ to all probability measures which in addition satisfy (3) is denoted by $\text{RST}(\mu, \nu)$.

Remark 2.4. The marginal of the random stopping time π is assumed have the distribution of μ. This can be understood as that the probabilities of the paths are preserved. Since we are working on Polish spaces, the (unique) disintegration $(\pi_\omega)_{\omega\in\mathbb{R}^I}$ exists and assigns μ-almost every path ω a probability measure on I.

In the setting of optimal transport it is more convenient to work with so-called couplings, which are product probability measures such that the marginals satisfy a certain law. For our case it is reasonable, to consider all couplings on $\mathbb{R}^I \times I$ between μ and ν.

Definition 2.5 (Couplings)**.**
A coupling on $\mathbb{R}^I \times I$ with marginals μ and ν is a product probability measure π on $\mathbb{R}^I \times I$ such that

1. $\mathrm{proj}_{\mathbb{R}^I}(\pi) = \mu$,
2. $\mathrm{proj}_I(\pi) = \nu$.

The space of all product measures on $\mathbb{R}^I \times I$ satisfying (1) and (2) is denoted by $\mathrm{Cpl}(\mu, \nu)$.

3. $\int \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)\mathrm{d}\pi(\omega, s) = 0 \quad \forall g \in B_b(\mathbb{R}^I),\ t \in I.$

The restriction of $\mathrm{Cpl}(\mu, \nu)$ to all couplings satisfying (3) is denoted by $\mathrm{Cpl}^{ad}(\mu, \nu)$.

For a coupling π property (3) corresponds to property (2) in the Definition 2.3 of randomized stopping times. In fact, $\mathrm{Cpl}^{ad}(\mu, \nu)$ coincides with $\mathrm{RST}(\mu, \nu)$. This follows from Lemma 2.6 which is an adaptation of [2, Theorem 3.8], where also a proof for the more complex time continuous case can be found.

Lemma 2.6. *Let $\pi \in \mathrm{Cpl}(\mu, \nu)$. Then the following are equivalent:*

1. $\pi \in \mathrm{Cpl}^{ad}(\mu, \nu)$,
2. *Given a disintegration $(\pi_\omega)_{\omega \in \mathbb{R}^I}$ of π, the random variable $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$.*

Proof. To show the equivalence we use a different characterization of measurability of integrable random variables, see e.g. in [9]:

An integrable random variable X on $\mathbb{R}^I$ is $\mathcal{F}_t$-measurable iff

$$\mathbb{E}[X(Y - \mathbb{E}[Y|\mathcal{F}_t])] = 0 \quad \forall Y \text{ integrable and Borel-measurable.}$$

Instead of working with all integrable random variables, we can restrict us to bounded, Borel-measurable random variables. Thus, by a monotone class argument and setting $X := \pi_\omega(t)$, this is equivalent to

$$\mathbb{E}[\mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])] = \int_{\mathbb{R}^I} \pi_\omega(t)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)\mu(\mathrm{d}\omega) = 0 \quad \forall g \in B_b(\mathbb{R}^I).$$

□

As explained in the introduction, we want to maximize the expected payoff given a cost function $c\colon S \to \mathbb{R}$ and the paths ω of a stochastic process Z where now the maximization is taken over all randomized stopping times which are in $\mathrm{RST}(\mu, \nu)$.

Problem ($\textsc{OptStop}^\pi$)**.** Let $\tilde{c}: S \to \mathbb{R}$ be Borel-measurable, then we can define the Borel-measurable function $c := \tilde{c} \circ r$, with r given by (2.1), by

$$\begin{aligned} c : \ \mathbb{R}^I \times I &\to \mathbb{R} \\ (\omega, t) &\mapsto \tilde{c}((\omega)_{s \leq t}, t). \end{aligned}$$

We want to find a maximizer of

$$\pi \mapsto \int_{\mathbb{R}^I \times I} c(\omega, t) \mathrm{d}\pi(\omega, t), \quad \pi \in \mathrm{RST}(\mu, \nu).$$

Remark 2.7. By Lemma 2.6 the problem 2.2 is equivalent to

$$\pi \mapsto \int_{\mathbb{R}^I \times I} c(\omega, t) \mathrm{d}\pi(\omega, t), \quad \pi \in \mathrm{Cpl}^{ad}(\mu, \nu).$$

Connection Between The Views

The next theorem gives us the connection between the different ways of formalizing our considered problem.

Theorem 2.8. *If the filtration $\mathbb{G}$ of the abstract probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ coincides with the natural filtration of the stochastic process Z, then there is a bijection between $\mathcal{M}_I$ and $\mathrm{RST}(\mu)$.*
Given a probability measure ν on I, then there is a bijection from $\mathcal{M}_I^\nu$ into $\mathrm{RST}(\mu, \nu)$.

Proof. Since $\mathbb{G}$ coincides with the natural filtration of the stochastic process Z, there is a Borel-measurable functions h_γ for every $\gamma \in \mathcal{M}_I$ such that

$$\begin{aligned} h_\gamma : \ S &\to \mathbb{R}, \\ \gamma_t(\bar{\omega}) &= h_\gamma(Z_s(\bar{\omega})_{s \leq t}, t) \quad \mathbb{P}\text{-a.e.}, \quad t \in I. \end{aligned}$$

We already know that the mapping $r : \ \mathbb{R}^I \times I \to S$ is Borel-measurable. Thus,

$$\Phi_\gamma := h_\gamma \circ r : \mathbb{R}^I \times I \to \mathbb{R}$$

is Borel-measurable, $\Phi_\gamma(\cdot, t)$ is $\mathcal{F}_t$-measurable and

$$\gamma_t(\bar{\omega}) = \Phi_\gamma(Z(\bar{\omega}), t) \quad \mathbb{P}\text{-a.e.}$$

Therefore, we deduce that for any $C \in \mathcal{G}_t$

$$\mathbb{E}[\mathbb{1}_C \gamma_t] = \mathbb{E}[\mathbb{1}_C \Phi_\gamma(Z, t)] = \mathbb{E}[\mathbb{1}_{\tilde{C}}(\omega) \Phi_\gamma(\omega, t)]$$

where $\tilde{C}$ is the with C associated set in $\mathcal{F}_t$. We define π such that $\pi(\mathrm{d}\omega, t) := \Phi_\gamma(\omega, t)\mu(\mathrm{d}\omega)$ which indeed defines a probability measure on $\mathbb{R}^I \times I$, and $\pi \in \mathrm{RST}(\mu)$. As a result, the map Ψ

$$\Psi : \ \mathcal{M}_I \to \mathrm{RST}(\mu) : \ \gamma \mapsto \pi$$

is well-defined and one-to-one.

For any $\pi \in \mathrm{RST}(\mu)$, the map $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-measurable. Hence, $\bar{\omega} \mapsto \pi_{Z(\bar{\omega})}$ is $\mathcal{G}_t$-measurable. Therefore, we may define $\tilde{\gamma}_t(\bar{\omega}) = \pi_{Z(\bar{\omega})}(t)$ and conclude

$$\Psi((\tilde{\gamma}_t)_{t\in I}) = \pi(Z(\bar{\omega}), t) \quad \mathbb{P}\text{-a.e, } t \in I,$$

which proofs the first part of the assertion. Using Lemma 2.6 the second part follows analogously. □

Remark 2.9. If the probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ coincides with $(\mathbb{R}^I, \mathcal{F}, \mathbb{F}, \mu)$ then the bijection of Theorem 2.8 follows due to the relation of disintegration and product measure on Polish spaces.

Corollary 2.10. *Under the assumptions of Theorem 2.8, we find that the optimization problems* ($\textsc{OptStop}^\gamma$) *and* ($\textsc{OptStop}^\pi$) *are equivalent.*

Proof. For any $\gamma \in \mathcal{M}_I^\nu$, we define a product measure on $\Omega \times I$ via $\tilde{\gamma}(d\bar{\omega}, t) = \gamma_t(\bar{\omega})\mathbb{P}(\mathrm{d}\bar{\omega})$. Following the proof of Theorem 2.8 we see that $(Z, id)_\# \tilde{\gamma} = \pi$, where π is the associated product measure on $\mathbb{R}^I \times I$.

$$\int_\Omega \sum_{t\in I} \tilde{c}((Z_s(\bar{\omega})_{s\le t}, t)\mathrm{d}\mathbb{P}(\bar{\omega}) = \int_{\Omega\times I} \tilde{c}((Z_s(\bar{\omega})_{s\le t}, t)\mathrm{d}\tilde{\gamma}(\bar{\omega}, t)$$
$$\int_{\mathbb{R}^I\times I} \tilde{c}((Z_s(\bar{\omega})_{s\le t}, t)\mathrm{d}\pi(\bar{\omega}, t) = \int_{\mathbb{R}^I\times I} c(\omega, t)\mathrm{d}\pi(\omega, t)$$

Following the second part of the proof of Theorem 2.8, we may define for any $\pi \in \mathrm{Cpl}^{ad}(\mu, \nu)$ an adapted random probability measure $\gamma \in \mathcal{M}_I^\nu$ via

$$\gamma_t(\bar{\omega}) := \pi_{Z(\bar{\omega})}(t),$$

where π_ω is the disintegration of π. Then

$$\int_{\mathbb{R}^I\times I} c(\omega, t)\mathrm{d}\pi(\omega, t) = \int_{\mathbb{R}^I} \sum_{t\in I} \tilde{c}((\omega_s)_{s\le t}, t)\mathrm{d}\pi_\omega(t)\mathrm{d}\mu(\omega) = \int_\Omega \sum_{t\in I} \tilde{c}((Z_s(\bar{\omega})_{s\le t}, t)\gamma_t(\bar{\omega})\mathbb{P}(d\bar{\omega}).$$

□

Hence, we have shown the connection between the approaches of Section 2.1 and 2.2. However, from now on we will restrict ourselves to the view described in Section 2.2, namely to reformulate it as an optimal transport problem.

CHAPTER 3

Existence and Duality

Existence of a Maximizer

Proposition 3.1. *For every* $b : \mathbb{R}^I \times I \to \mathbb{R}$*, bounded and Borel-measurable, the functional*

$$F : \operatorname{Cpl}(\mu,\nu) \to \mathbb{R} : \pi \mapsto \int b(\omega,t)\mathrm{d}\pi(\omega,t) =: \pi(b)$$

is continuous wrt. the weak topology on $\operatorname{Cpl}(\mu,\nu)$*.*

Proof. Given a sequence $\pi_n \in \operatorname{Cpl}(\mu,\nu),\ n \in \mathbb{N}$ such that $\pi_n \rightharpoonup \pi$, we know by [7, Lemma 2.3] that for all $A \in \mathcal{B}(\mathbb{R}^I)$ and $t \in I$

$$\pi_n(A \times \{t\}) \to \pi(A \times \{t\}).$$

Since b is bounded and measurable, there exists a sequence of simple functions b_m such that $|b - b_m| < \frac{1}{m}$ μ-a.e. Therefore

$$\forall \epsilon > 0\ \forall n \in \mathbb{N} \quad \exists m_\epsilon \colon |\pi(b - b_m)| < \epsilon \text{ and } |\pi_n(b - b_m)| < \epsilon, \quad m \geq m_\epsilon,$$

and

$$\forall \epsilon > 0\ \forall m \in \mathbb{N} \quad \exists n_\epsilon \colon |\pi_n(b_m) - \pi(b_m)| < \epsilon, \quad n \geq n_\epsilon.$$

We conclude

$$|\pi(b) - \pi_n(b)| \leq |\pi(b - b_m)| + |\pi(b_m) - \pi_n(b_m)| + |\pi_n(b - b_m)| < 3\epsilon, \quad n \geq n_\epsilon,\ m \geq m_\epsilon.$$

□

Corollary 3.2. *For every $h\colon \mathbb{R}^I \times I \to \mathbb{R} \cup \{-\infty\}$, bounded from above and Borel-measurable, the functional*

$$H \ : \ \operatorname{Cpl}(\mu, \nu) \to \mathbb{R} \ : \ \pi \mapsto \int h(\omega, t) \mathrm{d}\pi(\omega, t)$$

is upper semi-continuous wrt. the weak topology on $\operatorname{Cpl}(\mu, \nu)$.

Proof. We define $h_n := \max(h, -n),\ n \in \mathbb{N}$, which are bounded, measurable and $h_n \searrow h$ pointwise. By Proposition 3.1, we can define a sequence of continuous functionals

$$H_n \ : \ \operatorname{Cpl}(\mu, \nu) \to \mathbb{R} \ : \ \pi \mapsto \pi(h_n), \quad \text{where } \inf_n H_n(\pi) = H(\pi).$$

Let $\pi_m \rightharpoonup \pi$ in $\operatorname{Cpl}(\mu, \nu)$, then

$$H(\pi) = \inf_n H(\pi) = \inf_n \limsup_m H_n(\pi_m) \geq \limsup_m \inf_n H_n(\pi_m) = \limsup_m H(\pi_m).$$

□

For the sake of completeness, we want to state the notable *Prokhorov's* Theorem, see [11, Lemma 4.4].

Theorem 3.3 (Prokhorov)**.** *If X is a Polish space, then a set $P \subset \mathcal{P}(X)$ is precompact for the weak topology if and only if it is tight, i.e., for any $\epsilon > 0$ there is a compact set K_ϵ such that $\pi(X \backslash K_\epsilon) \leq \epsilon$ for all $\pi \in P$.*

The strategy behind showing existence is the following:

1. Show that the set, over which the supremum is taken, is compact.
2. Show that the functional is upper semi-continuous.

Note that $\operatorname{Cpl}^{ad}(\mu, \nu)$ is non-empty, since the product measure $\mu \otimes \nu$ satisfies the marginal properties of Definition 2.5 and it holds that

$$\begin{aligned}
&\int_{\mathbb{R}^I \times I} \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega) \, \mathrm{d}\mu \otimes \nu(\omega, s) \\
&= \int_I \mathbb{1}_{\{t\}}(s) \, \mathrm{d}\nu(s) \int_{\mathbb{R}^I} (g - \mathbb{E}[g|\mathcal{F}_t])(\omega) \, \mathrm{d}\mu(\omega) = \nu(\{t\}) \mathbb{E}_\mu[g - \mathbb{E}[g|\mathcal{F}_t]] = 0.
\end{aligned}$$

Thus, a maximizing sequence exists and the compactness provides a maximizer.

Theorem 3.4 (Existence of a Maximizer)**.** *Let $c \ : \ S \to \mathbb{R} \cup \{-\infty\}$ be bounded from above and Borel-measurable, then there exists a solution to* (OptStop^π).

Proof. As a direct consequence of Prokhorov's Theorem, see Theorem 3.3, we get that $\operatorname{Cpl}(\mu, \nu)$ is relatively compact. By [7, Lemma 2.3], we obtain in addition that $\operatorname{Cpl}(\mu, \nu)$ is closed, hence compact in the weak topology. To see the compactness of $\operatorname{Cpl}^{ad}(\mu, \nu)$, we consider

$$b(\omega, s) := \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)$$

for a bounded and Borel-measurable function g on $\mathbb{R}^I$ and $t \in I$. Let $\pi_n \rightharpoonup \pi$ in $\mathrm{Cpl}(\mu,\nu)$ and $\pi_n \in \mathrm{Cpl}^{ad}(\mu,\nu),\ n \in \mathbb{N}$, by applying Proposition 3.1, we obtain

$$0 = \pi_n(b) \to \pi(b).$$

Since s and g were arbitrary, we conclude $\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)$ signifying the compactness of $\mathrm{Cpl}^{ad}(\mu,\nu)$.

Now, choose a sequence $\pi_n \in \mathrm{Cpl}^{ad}(\mu,\nu)$ such that

$$\lim_n \pi_n(c) = \sup_{\tilde{\pi}\in\mathrm{Cpl}^{ad}(\mu,\nu)} \int_{\mathbb{R}^I\times I} c(\omega,s)\mathrm{d}\tilde{\pi}(\omega,s) =: C.$$

Due to the compactness of $\mathrm{Cpl}^{ad}(\mu,\nu)$, we can extract a convergent subsequence $\pi_{n_k} \rightharpoonup \pi \in \mathrm{Cpl}^{ad}(\mu,\nu)$ such that π possesses the desired property by Corollary 3.2

$$C = \limsup_{n_k} \pi_{n_k}(c) \leq \pi(c) \leq C.$$

□

Duality

The classical Monge-Kantorovich problem deals with the topic of minimizing the expected loss given a cost function $c:\ X\times Y \to \mathbb{R}\cup\{\infty\}$ when iterating over all couplings $\pi \in \mathrm{Cpl}(\mu,\nu)$, where μ and ν are probability measures on X and Y, respectively,

$$\mathrm{Cpl}(\mu,\nu) \ni \pi \mapsto \int c\mathrm{d}\pi.$$

The function c can be interpreted as the cost of moving mass from X which is distributed according to μ to Y, which shall be distributed according to ν. The couplings $\pi \in \mathrm{Cpl}(\mu,\nu)$ are called transport plans; and the coupling, which minimizes the expected loss, is called an optimal transport plan. We want to state the usual Kantorovich duality Theorem, see [11, Theorem 5.10].

Theorem 3.5 (Kantorovich Duality)**.** *Let (X,μ) and (Y,ν) be two Polish probability spaces and let $c:\ X\times Y \to \mathbb{R}\cup\{+\infty\}$ be a lower semi-continuous cost function, such that*

$$\forall (x,y) \in X\times Y,\quad c(x,y) \geq a(x) + b(y)$$

for some real-valued upper semi-continuous functions $a \in L^1(\mu)$ and $b \in L^1(\nu)$. Then there is duality

$$\inf_{\pi\in\mathrm{Cpl}(\mu,\nu)} \pi(c) = \sup_{\substack{(f,g)\in C_b(X)\times C_b(Y)\\ f_1+f_2\leq c}} \int_X f_1\mathrm{d}\mu + \int_Y f_2\mathrm{d}\nu.$$

We are interested in maximizing (OptStop$^\pi$), which is an maximization problem over the space $\mathrm{Cpl}^{ad}(\mu,\nu)$. In fact, this space is a restriction of $\mathrm{Cpl}(\mu,\nu)$. We may define a space W as the linear span of

$$\{w(\omega,s) := \mathbb{1}_{\{t\}}(s)(g-\mathbb{E}[g|\mathcal{F}_t])(\omega) \mid g \in C(\mathbb{R}^I)\cap L^1(\mu),\ t\in I\}. \tag{3.1}$$

Evidently $\mathrm{Cpl}^{ad}(\mu,\nu)$ coincides with the restriction of $\mathrm{Cpl}(\mu,\nu)$ to all couplings π satisfying the additional linear constraints

$$\pi(w) = 0 \quad \forall w \in W. \tag{3.2}$$

Obviously, we can switch from a maximization problem to a minimization problem, simply by multiplying the payoff function c with -1, and hence call it cost function. Thus the Kantorovich Duality Theorem has to be extended to the case, where linear constraints are posed to $\mathrm{Cpl}(\mu,\nu)$. A generalized version of this problem was shown by Zaev in [12]. Following the proof of [12, Theorem 2.1] and using Theorem 3.5 yields the following Theorem:

Theorem 3.6 (Duality)**.** *If the cost function c satisfies the assumptions of Theorem 3.5 with $X := \mathbb{R}^I$ and $Y := I$. Then there is duality*

$$\inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) = \sup_{\substack{(f_1,f_2,w)\in C_b(\mathbb{R}^I)\times C_b(I)\times W\\ f_1+f_2+w\le c}} \int_{\mathbb{R}^I} f_1 \mathrm{d}\mu + \sum_{t\in I} f_2(t)\nu(t).$$

Proof. The inequality

$$\inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) \ge \sup_{f_1+f_2+w\le c} \mu(f_1)+\nu(f_2)$$

follows immediately

$$\inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) \ge \inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \sup_{f_1+f_2+w\le c} \mu(f_1)+\nu(f_2) = \sup_{f_1+f_2+w\le c} \mu(f_1)+\nu(f_2).$$

For the reverse inequality we consider

$$\sup_{f_1+f_2+w\le c} \mu(f_1)+\nu(f_2) = \sup_{w\in W} \sup_{f_1+f_2\le c-w} \mu(f_1)+\nu(f_2).$$

Note that $W \subseteq C_b(\mathbb{R}^I\times I)$, thus, $c-w$ is again lower semi-continuous on $\mathbb{R}^I\times I$. We may choose $a_w(x) := a(x) - \|w\|_\infty$ and $b_w(x) := b(x)-\|w\|_\infty$ which in turn satisfy the assumption of Theorem 3.5. Thereby, we obtain that

$$\sup_{f_1+f_2+w\le c} \mu(f_1)+\nu(f_2) = \sup_{w\in W} \inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c-w).$$

For any $\pi \notin \mathrm{Cpl}^{ad}(\mu,\nu)$ there exists a $w\in W$ such that $\pi(w)<0$, and

$$\sup_{\alpha>0} \pi(c-\alpha w) = +\infty.$$

Since $\mathrm{Cpl}^{ad}(\mu,\nu)$ is not empty, we conclude that

$$\inf_{\pi\in\mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) = \sup_{\substack{(f_1,f_2,w)\in C_b(\mathbb{R}^I)\times C_b(I)\times W\\ f_1+f_2+w\le c}} \mu(f_1)+\nu(f_2).$$

□

CHAPTER 4

Examples

First, we want to consider a stochastic process $(Z_t)_{t\in I}$ with independent increments on the time domain $I := \{1, 2\}$ and maximize the following functional on $\mathrm{RST}(\mu, \nu)$

$$\pi \mapsto \int_{\mathbb{R}^2\times\{1,2\}} c(\omega, t)\,\mathrm{d}\pi(\omega, t), \tag{4.1}$$

as in ($\textsc{OptStop}^{\pi}$). To facilitate notations, we will denote by $\xi(\cdot, t)$ for $t \in I$ and $\xi \in \mathrm{RST}(\mu, \nu)$ the (sub-)probability measure on $\mathbb{R}^I$ induced by

$$A \mapsto \pi(B, t) \quad B \in \mathcal{B}(\mathbb{R}^I).$$

Given a randomized stopping time π, we may consider the measure m on $X_1 = \mathbb{R}$ such that

$$m(A) := \pi(A \times \mathbb{R} \times \{1\}) = \mathrm{proj}_{X_1}(\pi(\cdot, 1))(A) \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

Remember that $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-adapted, in particular for $t = 1$, we find a Z_1-measurable function $h : \mathbb{R}^I \to \mathbb{R}$ such that

$$\pi_\omega(1) = h(\omega) \quad \mu\text{-a.e. } \omega \in \mathbb{R}^I, \tag{4.2}$$

$$h(\omega) = h(\tilde{\omega}) =: h(\omega_1) \quad \omega,\ \tilde{\omega} \in \mathbb{R}^I,\ \omega_1 = \tilde{\omega}_1. \tag{4.3}$$

Therefore, we deduce that

$$m(A) = \int \mathbb{1}_{A\times\mathbb{R}}(\omega)\,\pi_\omega(1)\mathrm{d}\mu(\omega) = \int \mathbb{1}_A(\omega_1)h(\omega_1)\mathrm{d}\,\mathrm{proj}_{X_1}(\mu)(\omega), \tag{4.4}$$

which implies $m \leq \mathrm{proj}_{X_1}(\mu)$ and let us define a measure n on $\mathbb{R}$ satisfying

$$n := \mathrm{proj}_{X_1}(\mu) - m.$$

Using the measures m and n in the maximization problem (4.1) yields to

$$
\begin{aligned}
\int_{\mathbb{R}^I \times I} c(\omega,t)\mathrm{d}\pi(\omega,t) &= \int_{\mathbb{R}^I} c(\omega,1)\mathrm{d}\pi(\omega,1) + \int_{\mathbb{R}^I} c(\omega,2)\mathrm{d}\pi(\omega,2) \\
&= \int_{X_1} \tilde{c}((\omega_1),1)\mathrm{d}m(\omega_1) + \int_{\mathbb{R}^I} c(\omega,2)(1-h(\omega))\mu(\mathrm{d}\omega) \\
&= \int_{X_1} \tilde{c}((\omega_1),1)\mathrm{d}m(\omega_1) + \int_{X_1} (1-h(\omega_1)) \int_{\mathbb{R}} c((\omega_1,\omega_1+z),2)\mathrm{d}p(z)\, \mathrm{proj}_{X_1}(\mu)(\mathrm{d}\omega_1) \\
&= \int_{X_1} \tilde{c}((\omega_1),1)\mathrm{d}m(\omega_1) + \int_{X_1}\int_{\mathbb{R}} c((\omega_1,\omega_1+z),2)\mathrm{d}p(z)\mathrm{d}n(\omega_1).
\end{aligned} \tag{4.5}
$$

This equality can be interpreted in the following way: The measures m and n describe with which probability a path is stopped at time 1 or continues to time 2, respectively. Therefore the first integral in Line (4.5) describes the expected payoff if the process is stopped at time 1, whereas the second integral describes the expected payoff at time 2. Note that

$$\int_{\mathbb{R}} c((x,x+z),2)\mathrm{d}p(z)$$

is the expected payoff when we stop at time 2, conditioned on $\omega_1 = x$.

For any path $\omega := (\omega_1,\omega_2) \in \mathbb{R}^I$ and the corresponding probability measure π_ω on I, we see that

$$(c(\omega,1)\pi_\omega(1) + c(\omega,2)\pi_\omega(t)) - c(\omega,1) = (c(\omega,2) - c(\omega,1))\pi_\omega(2).$$

Without loss of generality we may assume that

$$c(\omega,1) = 0. \tag{4.6}$$

If $c(\omega,1) \neq 0$, than we might define $\bar{c}(\omega,t) := c(\omega,t) - c(\omega,1)$. Clearly, it hold that $\bar{c}(\omega,1) = 0$. Because subtracting a constant from the functional (4.1) doesn't change the property of being a maximizer of it, we will instead maximize

$$\pi \mapsto \int_{\mathbb{R}^I \times I} \bar{c}(\omega,t)\, \mathrm{d}\pi(\omega,t) = \int_{\mathbb{R}^I \times I} c(\omega,t)\mathrm{d}\pi(\omega,t) - \int_{\mathbb{R}^I} c(\omega,1)\mathrm{d}\mu(\omega),$$

where the last equality holds due to $\mathrm{proj}_{\mathbb{R}^I}(\pi) = \mu$. Note that both maximization problems are equivalent, but $\bar{c}(\omega,1) = 0$.

Using the reformulation (4.5) and assumption (4.6) our problem given in (4.1) is reduced to the following maximization problem:

Among all measures n on $X_1 = \mathbb{R}$ satisfying

$$n \leq \mathrm{proj}_{X_1}(\mu) \quad \text{and} \quad n(\mathbb{R}) = n(X_1) = \nu(2)$$

find the maximizer of

$$n \mapsto \int_{X_1}\int_{\mathbb{R}} c((x,x+z),2)\mathrm{d}p(z)\mathrm{d}n(x).$$

These considerations lead to the following Theorem:

Theorem 4.1. *Let* $I = \{1, 2\}$, $\mathbb{R}^I = X_1 \times X_2$.
Given an optimal $\pi^* \in \mathrm{RST}(\mu, \nu)$ *such that*

$$\pi^*(c) = \sup_{\pi \in \mathrm{RST}(\mu,\nu)} \pi(c). \tag{4.7}$$

Then,

$$n_{\pi^*} := \mathrm{proj}_{X_1}(\mu - \pi^*(\cdot, 1))$$

is a measure on $X_1 = \mathbb{R}$ *satisfying*

$$n_{\pi^*} \leq \mathrm{proj}_{X_1}(\mu), \quad n_{\pi^*}(\mathbb{R}) = n_{\pi^*}(X_1) = \nu(2), \tag{4.8}$$

and maximizes under all measures satisfying (4.8)

$$n_{\pi^*}(k) = \sup_n n(k), \quad k(x) := \int_{\mathbb{R}} c((x, x + z), 2)\mathrm{d}\mathbb{P}(z), \tag{4.9}$$

where $Z_2 - Z_1 \sim \mathbb{P}$. *Vice versa, let* n^* *be a measure on* $\mathbb{R}$ *satisfying* (4.8) *and maximizing* (4.9).

Then,

$$\pi_{n^*}((\mathrm{d}x, \mathrm{d}y), t) = \begin{cases} (\mathrm{proj}_{X_1}(\mu) - n^*)(\mathrm{d}x)\mathbb{P}(\mathrm{d}y) & \textit{for } t = 1, \\ n^*(\mathrm{d}x)\mathbb{P}(\mathrm{d}y) & \textit{for } t = 2, \end{cases}$$

defines a RST which maximizes (4.7).

Remark 4.2. Note the overlaps of the definitions with regard to the arguments c and (ω, t) or ω_1.

Proof. First, we take a closer look at n_{π^*} and note that it satisfies (4.8) and, due to Lemma 2.6 and (4.2),

$$n_{\pi^*}(\mathrm{d}x) = \mathrm{proj}_{X_1}(\mu - \pi^*(\cdot, 1))(\mathrm{d}x)(1 - h(x))\,\mathrm{proj}_{X_1}(\mu)(\mathrm{d}x),$$

implying

$$\pi^*(c) = \int_{X_1 \times \mathbb{R}} c((\omega_1, \omega_1 + z), 2)\pi^*_{(\omega_1, \omega_1+z)}(2)\mathrm{d}\mu((\omega_1, \omega_1 + z)) = \int_{\mathbb{R}} k(\omega_1)\mathrm{d}n_{\pi^*}(\omega_1). \tag{4.10}$$

On the other hand, π_{n^*} defines a measure on $\mathbb{R}^I \times I$ such that $\mathrm{proj}_I(\pi_{n^*}) = \nu$ and

$$\pi_{n^*}(c) = \int_{\mathbb{R}} k(\omega_1)\mathrm{d}n^*(\omega_1). \tag{4.11}$$

For any bounded, Borel-measurable function g on $\mathbb{R}^I$, the function

$$f(x) := \int_{\mathbb{R}} g(x, x + z)\mathrm{d}\mathbb{P}(z)$$

is Borel-measurable and

$$f(Z_1) = \mathbb{E}[f(Z_1)|\mathcal{F}_1] = \int_{\mathbb{R}} \mathbb{E}[g(Z_1, Z_1 + z)|\mathcal{F}_1]\mathrm{d}\mathbb{P}(z) \quad \text{a.e.},$$

which particularly yields for $t \in I$

$$\int_{X_1} \int_{\mathbb{R}} (g - \mathbb{E}[g|\mathcal{F}_t])(\omega_1, \omega_1 + z)\mathrm{d}\mathbb{P}(z)\mathrm{d}\,\mathrm{proj}_{X_1}(\mu)(\omega_1) = 0,$$

implying that $\pi_{n^*} \in \mathrm{RST}(\mu, \nu)$. Due to (4.10) and (4.11) the optimality of n_{π^*} and π_{n^*}, respectively, follow, and thus the assertion. $\square$

Remark 4.3. By Theorem 4.1, it becomes obvious that for two time steps, it is sufficient to consider a quantile q of k such that q is maximal in $\mathbb{R}$ with the property $\nu(2) \leq \mathrm{proj}_{X_1}(U_q)$, where $U_q := \{x \in X_1 \ : \ k(x) \geq q\}$.

$$n(\mathrm{d}x) := \begin{cases} \mathrm{proj}_{X_1}(\mu)(\mathrm{d}x) & k(x) > q, \\ \alpha \cdot \mathrm{proj}_{X_1}(\mu)(\mathrm{d}x) & k(x) = q, \\ 0 & \text{else}, \end{cases} \tag{4.12}$$

where $\alpha \in [0, 1]$ is chosen such that $n(\mathbb{R}) = \nu(2)$. It is apparent that n maximizes (4.7), since for any other measure $\tilde{n}$ satisfying (4.8), it holds

$$n(k) - \tilde{n}(k) = \int_{\mathbb{R}} k(x)(n - \tilde{n})(\mathrm{d}x) \geq \int_{\mathbb{R}} q - k(x)(\tilde{n} - n)^+(\mathrm{d}x) \geq 0.$$

The second part of Theorem 4.1 yields the maximizer π_n. As already mentioned, this method can be interpreted as not-stopping those paths, which have a higher expected payoff in the next turn. Instead of interpreting it this way, we can view it as stopping all paths with a lower expected payoff – indeed, this is would result in the same quantiles.

For a special class of cost functions, we will be able to extend the idea of using quantiles of the corresponding k for each time step inductively, to construct an optimizer for an arbitrary amount of time steps.

Example: Symmetric Random Walk

For $I := \{1, \ldots, T\}$, $T \in \mathbb{N}$, we consider a symmetric random walk $(Z_t)_{t \in I}$ on $\mathbb{Z}$ starting at 0, where the increments $Z_s - Z_{s-1}$ are independent and uniformly distributed on $\{-1, 1\}$. Let the payoff function c be

$$c \ : \ \mathbb{R}^I \times I \to \mathbb{R} \ : \ (\omega, t) \mapsto t \cdot \omega_t$$

which is indeed $\mathcal{F}_t$-adapted and bounded, if the time horizon is finite.

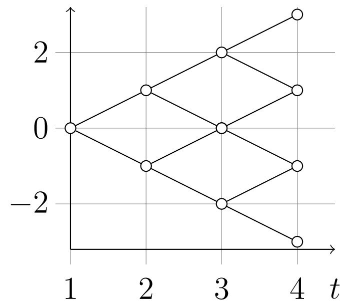

Figure 4.1: Paths of the process $(Z_t)_{t\in I}$ starting at 0 for $I := \{1, 2, 3, 4\}$.

We want to find an optimal $\pi \in \mathrm{RST}$. Note that for $\omega, \eta \in \mathbb{R}^I$ and $t \in I$

$$\begin{aligned} \int \tilde{c}\Big((\omega_0, \ldots, \omega_{t-1}, \omega_{t-1} + z), t\Big)\, \mathrm{d}p_t(z) &\geq \int \tilde{c}\Big((\eta_0, \ldots, \eta_{t-1}, \eta_{t-1} + z), t\Big)\, \mathrm{d}p_t(z) \\ \iff \int t \cdot (\omega_{t-1} + z)\, \mathrm{d}p_t(z) &\geq \int t \cdot (\eta_{t-1} + z)\, \mathrm{d}p_t(z) \\ \iff \omega_{t-1} &\geq \eta_{t-1}. \end{aligned} \tag{4.13}$$

By using Theorem 4.1 we are able to solve the problem restricted to two time steps. By skillful projection of the path space, we may consider a case with exactly two time steps. For the marginal of π to satisfy the constraint ν, we have to define $\pi(\omega, 1) = \nu(1)$. For every step of our recursion $i \geq 2$ we consider the two dimensional space $X_{t_i} \times X_{t_{i+1}}$. Note that for $I := \{1, \ldots, T\}$, $T \in \mathbb{N}$, we consider $X_t \times X_{t+1}$, and the recursion step i corresponds to the time step t. Therefore we define measures $\mu_i := \mathrm{proj}_{X_t \times X_{t+1}}(\mu - \sum_{s<t} \pi(\cdot, s))$ and ν_i such that

$$\nu_i(1) = \nu(t), \quad \nu_i(2) = \sum_{s \in I_{>t}} \nu(s).$$

Then we have to solve the problem to find the maximizer of

$$n \mapsto \int_{X_t} \int_{\mathbb{R}} c((x, x+z), 2) \mathrm{d}p_t(z) \mathrm{d}n(x),$$

among all measures n_i on $X_t = \mathbb{R}$ satisfying

$$n_i \leq \mathrm{proj}_{X_t}(\mu_i) \quad \text{and} \quad n_i(\mathbb{R}) = \nu_i(2),$$

which can be done according to Remark 4.3, particularly (4.12). Thus, by applying Theorem 4.1 recursively, we obtain a randomized stopping time $\pi^i \in \mathrm{RST}(\mu_i, \nu_i),\ i \in I$ - which can be naturally merged into $\pi \in \mathrm{RST}(\mu, \nu)$ via $\pi(\omega, t) := (\mu(\omega) - \sum_{s<t} \pi(\omega, s)) \cdot \pi^t_{(\omega_t, \omega_{t+1})}(t)$. Optimality can be shown in the following way: Starting with another arbitrary $\xi \in \mathrm{RST}(\mu, \nu)$ we assume that there exists a minimal t such that $\xi(\cdot, t) \neq \pi(\cdot, t)$. Therefore, there exist $\zeta, \eta \in \mathbb{R}^I$ such that

$$\alpha_1 := \xi_\eta(t) - \pi_\eta(t) > 0, \quad \alpha_2 := \pi_\zeta(t) - \xi_\zeta(t) > 0.$$

The overall mass of all paths $\omega \in \mathbb{R}^I$ such that the initial segments $\omega_{\restriction[0,t]}$ coincide with that of η or ζ, is 2^{-t}. Hence, it is possible to swap $2^{-t}\alpha$ mass from $\{\omega \in \mathbb{R}^I : \omega_{\restriction[0,t]} =$

$\eta_{\restriction[0,t]}\}$ to $\{\omega \in \mathbb{R}^I : \omega_{\restriction[0,t]} = \zeta_{\restriction[0,t]}\}$, where $\alpha := \min(\alpha_1, \alpha_2)$. A new measure is gradually defined by

$$\tilde{\xi}(\omega, s) = \xi(\omega, s) \quad \omega \in \mathbb{R}^I,\, s < t$$

and

$$\tilde{\xi}_\omega(t) := \begin{cases} \xi_\omega(t) - \alpha & \forall \omega \in \mathbb{R}^I \text{ such that } \omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}, \\ \xi_\omega(t) + \alpha & \forall \omega \in \mathbb{R}^I \text{ such that } \omega_{\restriction[0,t]} = \zeta_{\restriction[0,t]}, \\ \xi_\omega(t) & \text{otherwise} \end{cases}$$

So far, $\nu(s) = \mathrm{proj}_I(\tilde{\xi})$ still holds as long as $s \leq t$. To fully preserve the marginals, i.e., $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$, mass is carefully added to the remaining paths $\omega \in \mathbb{R}^I$ where $\omega_{\restriction[0,t]} \in \{\eta_{\restriction[0,t]}, \zeta_{\restriction[0,t]}\}$. Let $\omega \in \mathbb{R}^I$ such that $\omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}$, then there exists $\theta \in \mathbb{R}^I$ such that $\theta_{\restriction[0,t]} = \zeta_{\restriction[0,t]}$ and $(\theta - \omega)(r) = 0$ for all $r > t$. We may set

$$\tilde{\xi}_\omega(s) := \xi_\omega(s) + \frac{\alpha}{1 - \sum_{r \leq t} \xi_\theta(r)} \cdot \xi_\theta(s), \quad \tilde{\xi}_\theta(s) := \xi_\theta(s) - \frac{\alpha}{1 - \sum_{r \leq t} \xi_\theta(r)} \cdot \xi_\theta(s),$$

which yields the correct marginals for $\tilde{\xi}$. Clearly, $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$ and due to Theorem 4.1 together with equation (4.13) it holds that $\eta_t \geq \zeta_t$, yielding

$$\tilde{\xi}(c) - \xi(c) = \int_{\{\omega \in \mathbb{R}^I \,:\, \omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}\} \times I_{>t}} (\eta_t - \zeta_t)(s - t)(\tilde{\xi} - \xi)^+(\mathrm{d}\omega, s) \geq 0. \tag{4.14}$$

Hence, continuing this construction recursively leads in a finite amount of steps to

$$\tilde{\xi}(\omega, s) = \pi(\omega, s) \quad s \leq t,\ \omega \in \mathbb{R}^I.$$

By equation (4.14), $\pi(c)$ is an upper bound for the payoff of any constructed $\tilde{\xi}$, and as a consequence an upper bound for $\xi(c)$.

Remark 4.4. The method shown in Example 4.1 can be used to show optimality for the introduced "greedy" strategy π for a larger class of optimization problems. In the setting of Example 4.1, we cannot expect uniqueness of the optimizer π, since for any $\eta, \zeta \in \mathbb{R}^I,\ \eta_t = \zeta_t$ and $t \in I$ such that

$$\pi_\eta(t) > 0, \quad \pi_\zeta(t) < 1,$$

it is possible to swap some mass analogously as described above, creating a new randomized stopping time, but preserving marginals and payoff.

Example: Generalized Setting

Let the stochastic process $(Z_t)_{t \in I}$ have independent increments and the payoff function c be of the form

$$c(\omega, t) = \omega_t f(t),$$

where $f : I \to \mathbb{R}^+$ is monotonously increasing. Motivated by Theorem 4.1 and Example 4.1, we want to show that a greedy algorithm is optimal here, cf. Theorem 4.8. Analogous to (4.13), we obtain for the expected payoff

$$\begin{aligned} \int \tilde{c}((\omega_1, \dots, \omega_t, \omega_t + z), t+1) \mathrm{d}p_{t+1}(z) &= \int f(t+1) z \mathrm{d}p_{t+1}(z) + f(t+1)\omega_t \\ &\underset{(=)}{>} \int \tilde{c}((\eta_1, \dots, \eta_t, \eta_t + z), t+1) \mathrm{d}p_{t+1}(z) \qquad (4.15) \\ &\iff \omega_t \underset{(=)}{>} \eta_t. \end{aligned}$$

We construct a randomized stopping time π by defining a quantile q_t for any $t \in I$ such that

$$q_t := \inf\{q \in \mathbb{R} : \ \operatorname{proj}_{X_t}(\mu - \sum_{s<t} \pi(\cdot, s))((-\infty, q]) \geq \nu(t)\},$$

where $\pi(\omega, t)$ can be defined as

$$\pi(\omega, t) := \operatorname{proj}_{X_t}(\mu - \sum_{s<t} \pi(\cdot, s))_{\restriction(-\infty, q_t]}(\omega_t), \tag{4.16}$$

if the quantile q_t is exact, else it can be defined similar to Remark 4.3.

As in Example 4.1, we will show optimality of this strategy by transforming any randomized stopping time iteratively into the proposed one, without lowering the payoff. But, before we can show optimality we need some preparations to conduct the swapping of mass.

Lemma 4.5. *Let m, n be finite measures on $[0,1]$ such that $m([0,1]) = n([0,1])$. Then there exists a Borel-measurable maps $U = (U_1, U_2)$ from $[0,1] \times [0,1]$ onto $[0,1] \times [0,1]$ such that*

$$m([0,x)) + u \cdot m(\{x\}) = n([0, U_1(x,u))) + U_2(x,u) \cdot n(\{U_1(x,u)\}).$$

In addition, the map U_1 is surjective onto $\operatorname{supp}(n)$.

Proof. We can easily extend the measures m and n to measures M and N on $[0,1] \times [0,1]$ by defining them via

$$M(\mathrm{d}x, \mathrm{d}u) = m(\mathrm{d}x), \quad N(\mathrm{d}y, \mathrm{d}v) = n(\mathrm{d}y).$$

For a given pair $(x,u) \in [0,1] \times [0,1]$ we may define the first component of U as

$$\begin{aligned} U_1(x,u) := \inf\{y \in [0,1] \mid \\ M([0,x) \times [0,1]) + M(\{x\} \times [0,u]) \leq n([0,y])\} \in \operatorname{supp}(n). \end{aligned} \tag{4.17}$$

The corresponding second component of U can be defined as follows

$$\begin{aligned} U_2(x,u) := \inf\{v \in [0,1] : \ M([0,x) \times [0,1]) + M(\{x\} \times [0,u]) \\ = N([0, U_1(x,u)) \times [0,1]) + N(\{U_1(x,u)\} \times [0,v])\}. \end{aligned} \tag{4.18}$$

By construction, $U(x,u) = (U_1(x,u), U_2(x,u))$ is well-defined and Borel-measurable. Further,

$$\begin{aligned} m([0,x)) + u \cdot m(\{x\}) &= M([0,x) \times [0,1]) + M(\{x\} \times [0,u]) \\ &= N([0,U_1(x,u)) \times [0,1]) + N(\{U_1(x)\} \times [0,U_2(x,u)]) \\ &= n([0,U_1(x,u))) + U_2(x,u) \cdot n(\{U_1(x,u)\}), \end{aligned}$$

where the second equality follows from (4.18).

Assume that there exist $(x,u) \in [0,1] \times [0,1]$ with $U_1(x,u) =: y \notin \mathrm{supp}(n)$, then there exists a $\delta > 0$ such that $n([y-\delta, y+\delta]) = 0$, and hence

$$M([0,x) \times [0,1]) + M(\{x\} \times [0,u]) \leq n([0,y-\delta]),$$

which contradicts the definition of $U_1(x,u)$, see (4.17). Furthermore, for any $y \in \mathrm{supp}(n)$ there exist $(x,u) \in [0,1] \times [0,1]$ with

$$m([0,x)) \leq n([0,y]) \leq m([0,x]),$$

$$M([0,x) \times [0,1]) + M(\{x\} \times [0,u]) = n([0,y]).$$

$z \in \mathrm{supp}(n),\ z < y$ implies $n([0,z]) < n([0,y])$, which yields $U_1(x,u) = y$ and surjectivity. $\square$

Lemma 4.6. *Under the assumptions of Lemma 4.5, we may define a map $V_y : [0,1] \to [0,1]$ for fixed $y \in \mathrm{supp}(n)$ by*

$$V_y(x) := \begin{cases} 1 & x \in \mathrm{int}(\gamma_y),\ n(\{y\}) \neq 0, \\ \sup_{(x,u) \in U_1^{-1}(\{y\})} u - \inf_{(x,v) \in U_1^{-1}(\{y\})} v & x \in \partial\gamma_y,\ n(\{y\}) \neq 0, \\ 0 & \textit{else}. \end{cases}$$

The map V_y is well-defined, where $\gamma_y := \{x : \exists u \in [0,1] \text{ s.t. } U_1(x,u) = y\}$. The set γ_y is a closed interval and the maps $y \mapsto \inf_{(x,t) \in U_1^{-1}(\{y\})} x =: x_l(y)$ and $(x,y) \mapsto V_y(x)$ are Borel-measurable.

Proof. Let $y \in \mathrm{supp}(n)$. Note that for any $(x_1,u_1),(x_2,u_2) \in U_1^{-1}(\{y\})$ holds

$$(x_1,u_1) \leq (x,u) \leq (x_2,u_2) \implies (x,t) \in U_1^{-1}(\{y\}),$$

where $\leq$ refers to the lexicographical order. Hence,

$$\gamma_y = \{x \in [0,1] : \exists u \in [0,1] \text{ s.t. } (x,u) \in U_1^{-1}(\{y\})\}.$$

Especially, γ_y is an interval with left and right boundary points x_l and x_r. When $n(\{y\}) \neq 0$, for any point $x \in (x_l, x_r)$ follows $V_y(x) = 1$. Further, γ_y contains its boundary points, since $(x_l,1),(x_r,0) \in U_1^{-1}(\{y\})$. If $y_1, y_2 \in \mathrm{supp}(n),\ y_1 < y_2$ with

$$\gamma_{y_1} = [a_l, a_r],\ \gamma_{y_2} = [b_l, b_r],$$

implies that $a_r \leq b_l$. Therefore, the map $y \mapsto x_l(y)$ is monotonously increasing, and hence Borel-measurable. As a matter of fact there can only be a countable amount of point masses of n, which shows measurability of $(x,y) \mapsto V_y(x)$. $\square$

To simplify notation, for any starting point $x \in \mathbb{R}$ we set

$$\vec{x} := [x, 0, \dots, 0] \in \mathbb{R}^I.$$

Given the starting distribution $\mathrm{proj}_{X_1}(\mu) = \frac{1}{2}(\delta_x + \delta_y)$ for $x \neq y$ and $\xi \in \mathrm{RST}(\mu, \nu)$. By virtue of Z's independent increments, we can construct another probability measure $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$ such that

$$\tilde{\xi}_{\omega+\vec{x}} = \xi_{\omega+\vec{y}}, \quad \tilde{\xi}_{\omega+\vec{y}} = \xi_{\omega+\vec{x}}, \quad \omega \in \mathbb{R}^I.$$

In the following Lemma, we construct another randomized stopping time by following the idea of "swapping branches" when the starting distributions are arbitrary (sub-)probability measures.

Lemma 4.7. *Under the assumptions of Lemma 4.5, let $\xi \in \mathrm{RST}(\mu, \nu)$ where m and n satisfy*

$$m = \mathrm{proj}_{X_1}(\mu), \quad n = \mathrm{proj}_{X_1}(\tilde{\mu}),$$

and are the starting distributions of the (sub-)probability measures μ and $\tilde{\mu}$ associated with Z. Then, for fixed $y \in \mathrm{supp}(n)$ the measure

$$m_y(\mathrm{d}x) = \begin{cases} \delta_{x_l(y)}(\mathrm{d}x) & n(\{y\}) = 0, \\ \frac{V_y(x) m(\mathrm{d}x)}{n(\{y\})} & \textit{else}, \end{cases}$$

is well-defined. Furthermore, there exists $\tilde{\xi} \in \mathrm{RST}(\tilde{\mu}, \nu)$ with disintegration $(\tilde{\xi}_\omega)_{\omega \in \mathbb{R}^I}$ such that for $t \in I$ and $\omega \in \mathbb{R}^I$

$$\tilde{\xi}_\omega(t) := \begin{cases} \int \xi_{\omega+\vec{x}-\vec{\omega_1}}(t) m_{\omega_1}(\mathrm{d}x) & \omega_1 \in \mathrm{supp}(n), \\ \mathbb{1}_{\{1\}}(t) & \textit{else}. \end{cases} \tag{4.19}$$

Proof. According to Lemma 4.6 the measure m_y is well-defined and $y \mapsto m_y$ measurable. Following the proof of Lemma 4.5, there exist $u_l, u_r \in [0, 1]$ such that

$$m([0, x_l)) + u_l \cdot m(\{x_l\}) = n([0, y)),$$
$$m([0, x_r)) + u_r \cdot m(\{x_r\}) = n([0, y]),$$

where $[x_l, x_r] = \gamma_y$. We will only discuss the case that $x_l < x_r$, since the other case can be dealt with in similar fashion. If $n(\{y\}) > 0$, it holds that $u_l = 1 - V_y(x_l)$ and $u_r = V_y(x_r)$.

$$m(V_y) = m([x_l, x_r)) + u_r \cdot m(\{x_r\}) - (1 - u_l) \cdot m(\{x_l\}) = n(\{y\})$$

Thus m_y is a probability measure on $[0, 1]$. As a composition of measurable functions, $\omega \mapsto \tilde{\xi}_\omega$ is measurable. By construction, $\tilde{\xi}$ is $\mathbb{F}$-adapted, therefore it remains to establish the marginal properties. Let $\omega \in \mathbb{R}^I$ and $\omega_1 \in \mathrm{supp}(n)$.

$$\sum_{t \in I} \tilde{\xi}_\omega(t) \overset{(4.19)}{=} \int \sum_{t \in I} \tilde{\xi}_{\omega+\vec{x}-\vec{\omega}_1} m_{\omega_1}(\mathrm{d}x) = 1,$$

which implies

$$\mathrm{proj}_{\mathbb{R}^I}(\tilde{\xi})(\mathrm{d}\omega) = \tilde{\mu}(\mathrm{d}\omega).$$

$$\begin{aligned}\mathrm{proj}_I(\tilde{\xi})(t) &= \int_{\mathbb{R}^I} \tilde{\xi}(\omega, t)\tilde{\mu}(\mathrm{d}\omega) = \int \tilde{\xi}_{\theta+\vec{y}}(t)\mathbb{P}(\mathrm{d}\theta)n(\mathrm{d}y) \\ &= \int\int \xi_{\theta+\vec{x}} m_y(\mathrm{d}x)n(\mathrm{d}y)\mathbb{P}(\mathrm{d}\theta)\end{aligned}$$

To prove $\mathrm{proj}_I(\tilde{\xi}) = \nu(t)$ it is sufficient to show that for any interval $A := [a, b] \subseteq [0, 1]$

$$m(A) = \int_{[0,1]^2} \mathbb{1}_A(x)m_y(\mathrm{d}x)n(\mathrm{d}y). \tag{4.20}$$

since the assertion follows then by the Monotone Class Theorem. With Lemma 4.5 we obtain for $U(a, 0) =: (y_1, v_1)$ and $U(b, 1) =: (y_2, v_2)$. As above, we assume that $y_1 < y_2$ which yields

$$\begin{aligned}m([0, a)) &= n([0, y_1)) + v_1 \cdot n(\{y_1\}), \\ m([0, b]) &= n([0, y_2)) + v_2 \cdot n(\{y_2\}),\end{aligned}$$

cf. (4.17) and (4.18). And hence

$$\begin{aligned}m([a, b]) &= (1 - v_1) \cdot n(\{y_1\}) + n((y_1, y_2)) + v_2 \cdot n(\{y_2\}), \\ m([0, a)) &\le n([0, y]) \le m([0, b]) \quad \forall y \in (y_1, y_2) \cap \mathrm{supp}(n).\end{aligned}$$

For any $y \in (y_1, y_2) \cap \mathrm{supp}(n)$, $x \in \gamma_y$ and $v \in [0, 1]$, we note that

$$m([0, a]) \le n([0, y_1]) \le m([0, x)) + v \cdot m(\{x\}) \le n([0, y_2)) \le m([0, b]),$$

which implies that $\gamma_y \subseteq A$ and $m_y(\gamma_y \cap A) = 1$, and thus

$$n((y_1, y_2)) = \int_{(y_1,y_2)} \int_{[0,1]} \mathbb{1}_A(x)m_y(\mathrm{d}x)n(\mathrm{d}y).$$

In the case that $n(\{y_1, y_2\}) = 0$, the assertion follows. If y_1 or y_2 are point masses, we have to show that $(1 - v_1) = m_{y_1}(A)$ and $v_2 = m_{y_2}(A)$, respectively. We only consider the case that y_l is a point mass, since the other cases (y_3 or $y_1 = y_2$ is a point mass) follow analogously. We know that $\gamma_{y_1} \cap A = [a, c] \subseteq [a, b]$ and

$$\begin{aligned}m([0, a)) &= n([0, y_1)) + v_1 \cdot n(\{y_1\}), \\ n([0, y_1]) &= m([0, c)) + u \cdot m(\{c\}).\end{aligned}$$

Based on these two equations, the assertion follows.

$$1 - v_1 = \frac{1}{n(\{y_1\})}(m([a, c]) + u \cdot m(\{c\})) = m_{y_1}(A)$$

□

Theorem 4.8. *The greedy strategy π is optimal, i.e.,*

$$\pi(c) \geq \xi(c), \quad \forall \xi \in \mathrm{RST}(\mu, \nu).$$

If f is strictly increasing π is the unique optimizer in the following sense: Any $\tilde{\pi} \in \mathrm{RST}(\mu, \nu)$ with $\tilde{\pi}(c) = \pi(c)$ satisfies for all $t \in I,\ t < T$

$$\mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot, t) - \tilde{\pi}(\cdot, t)) = 0 \quad \mathrm{proj}_{X_t \times X_{t+1}}(\mu)\text{-}a.e.$$

Proof. Let $t \in I$ be fixed and assume that $\pi(\cdot, s) = \xi(\cdot, s)$ μ-a.s. for all $s < t$. Hence, there exists a $\mathcal{F}_t$-measurable set $B \subseteq \mathbb{R}^I$ with full measure such that $\pi_\omega(s) = \xi_\omega(s)$ holds pointwise for all $\omega \in B$, $s < t$.

$$A^+ := \{\omega \,:\, (\pi - \xi)_\omega(t) > 0\} \cap B, \quad A^- := \{\omega \,:\, (\xi - \pi)_\omega(t) > 0\} \cap B.$$

$$A := A^+ \cup A^-$$

In view of the (quantile) structure of π, cf. (4.16), it follows that

$$\omega_t \leq \eta_t \quad \forall \omega \in A^+,\ \forall \eta \in A^-. \tag{4.21}$$

Let us consider the finite measures ϕ^+ and ϕ^- on $A \times I$

$$\phi^+ := (\pi - \xi)^+_{\restriction A \times I}, \quad \phi^- := (\xi - \pi)^+_{\restriction A \times I}.$$

Note that $\mathrm{proj}_{\mathbb{R}^I}(\phi^+) = \mathrm{proj}_{\mathbb{R}^I}(\phi^-)$, which implies

$$0 < \phi^+(A^+, t) =: \alpha \leq \phi^+(A^+ \times I_{\geq t}) = \phi^-(A^+ \times I_{>t}) =: \beta.$$

The second marginal property of π and ξ, i.e., $\mathrm{proj}_I(\pi)(t) = \nu(t) = \mathrm{proj}_I(\xi(t))$, yields

$$\alpha = \phi^+(A^+, t) = \phi^-(A^-, t).$$

We may define two $\mathbb{F}$-adapted measures ψ and χ via

$$\psi(\mathrm{d}\omega, s) := \frac{\alpha}{\beta} \cdot \begin{cases} \phi^-(\mathrm{d}\omega, s) & (\omega, s) \in A^+ \times I, \\ 0 & \text{else.} \end{cases},$$

$$\chi(\mathrm{d}\omega, s) := \begin{cases} \phi^-(\mathrm{d}\omega, s) & (\omega, s) \in A^- \times \{t\}, \\ 0 & \text{else.} \end{cases}$$

Due the scaling factor of $\frac{\alpha}{\beta}$, we obtain $\psi(A^+ \times I) = \chi(A^- \times I)$. Therefore, we can define m and n as the starting distributions of

$$\bar{\psi} := \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\psi), \quad \bar{\chi} := \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\chi),$$

$$m := \mathrm{proj}_{X_t}(\bar{\psi}), \quad n := \mathrm{proj}_{X_t}(\bar{\chi}).$$

Let $C \subseteq A$ be a $\mathcal{F}_s$-measurable set for $s \geq t$ and C' its projection onto $\mathbb{R}^{I_{\geq t}}$, then

$$\psi(C) = \bar{\psi}(C'), \quad \chi(C) = \bar{\chi}(C').$$

Again by a monotone class argument it follows that

$$\psi(c) = \bar{\psi}(c'), \quad \chi(c) = \bar{\chi}(c'),$$

where c' is the natural restriction of c onto $\mathbb{R}^{I_{\geq t}} \times I_{\geq t}$. Applying Lemma 4.7 to m, n, $\bar{\psi}$ and $\bar{\chi}$ results in two measures $\tilde{\psi}$ and $\tilde{\chi}$, preserving the marginals of $\bar{\psi}$ and $\bar{\chi}$, respectively. In the final step, we want to extend $\tilde{\chi}$ and $\tilde{\psi}$ to measures $\hat{\psi}$ and $\hat{\chi}$ on $\mathbb{R}^I \times I$.

$$\hat{\psi}(\mathrm{d}\omega, s) := \begin{cases} \tilde{\psi}_{\omega_{\upharpoonright[t,T]}}(s)\,\mathrm{proj}_{\mathbb{R}^I}(\psi)(\mathrm{d}\omega) & \omega \in A^+,\ s \geq t, \\ 0 & \text{else.} \end{cases}$$

$$\hat{\chi}(\mathrm{d}\omega, s) := \begin{cases} \tilde{\chi}_{\omega_{\upharpoonright[t,T]}}(s)\,\mathrm{proj}_{\mathbb{R}^I}(\chi)(\mathrm{d}\omega) & \omega \in A^-,\ s \geq t, \\ 0 & \text{else.} \end{cases}$$

For $s \geq t$ we obtain

$$\mathrm{proj}_I(\psi + \chi)(s) = \mathrm{proj}_{I_{\geq t}}(\bar{\psi} + \bar{\chi})(s) = \mathrm{proj}_{I_{\geq t}}(\hat{\psi} + \hat{\chi})(s) = \mathrm{proj}_I(\tilde{\psi} + \tilde{\chi})(s).$$

$$\mathrm{proj}_{\mathbb{R}^I}(\psi + \chi) = \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\tilde{\psi} + \tilde{\chi}) = \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\tilde{\psi} + \tilde{\chi}) = \mathrm{proj}_{\mathbb{R}^I}(\hat{\psi} + \hat{\chi}),$$

holds μ-a.e. Hence, we are able to define $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$ via

$$\tilde{\xi} := \xi - \psi - \chi + \hat{\psi} + \hat{\chi}.$$

$$\begin{aligned}
\psi(c) + \chi(c) = \bar{\psi}(c') + \bar{\chi}(c') = &\\
\int c'(\theta + \vec{y}, 1)\bar{\psi}_{\theta+\vec{y}}(1)\mathbb{P}(\mathrm{d}\theta)n(\mathrm{d}y) &\\
+ \int c'(\theta + \vec{x}, s))\bar{\chi}_{\theta+\vec{x}}(s)\mathbb{P}(\mathrm{d}\theta)m(\mathrm{d}x) &\\
\leq \int c'(\theta + \vec{y}, s)\bar{\chi}_{\theta+\vec{x}}(s)\mathbb{P}(\mathrm{d}\theta)m_y(\mathrm{d}x)n(\mathrm{d}y) &\\
+ \int c'(\theta + \vec{x}, 1)\bar{\psi}_{\theta+\vec{y}}(1)\mathbb{P}(\mathrm{d}\theta)n_x(\mathrm{d}y)m(\mathrm{d}x) &\\
= \tilde{\psi}(c') + \tilde{\chi}(c') = \hat{\psi}(c) + \hat{\chi}(c), &
\end{aligned}$$

where the inequality holds due to (4.21), which implies

$$x \geq y \quad \forall x \in \mathrm{supp}(m),\ y \in \mathrm{supp}(n),$$

together with the explicit form of c' and (4.20). We see that we can exchange ξ with $\tilde{\xi}$ without lowering the payoff. Further, $\tilde{\xi}(\omega, s) = \pi(\omega, s)$ ω-a.e. for all $s \leq t$. By continuing this procedure iteratively over t, we can transform the initial ξ into π without lowering the payoff. Hence, π is optimal. If f is monotonously strictly increasing, this inequality holds strictly iff

$$\mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot, t) - \xi(\cdot, t)) \neq 0 \quad \mathrm{proj}_{X_t \times X_{t+1}}(\mu) - \text{a.e.}$$

Thus, π is the unique optimizer in the sense described above. $\square$

CHAPTER 5

Monotonicity Principles in an Example

To test if a randomized stopping time is a possible candidate for optimality in problem (OptStop$^\pi$), different monotonicity criterions were developed. In this context, the so-called c-cyclical monotonicity as in [11] deserves a special mention, which is in fact a geometric property of the support of an optimal transport plan. In the initial form the monotonicity was shown only for couplings which do not have to satisfy additional adaptivity constraints. Zaev introduced (c, W)-cyclical monotonicity in [12, Theorem 3.6], which enhances the notion with constraints, denoted by W. In our considerations, randomized stopping times are couplings satisfying additional linear constraints given through (3.2) and (3.1). Thus, Zaev's monotonicity principle can be applied naturally. Contrary to the classical c-monotonicity, the (c, W)-monotonicity of a support of a randomized stopping time is a necessary optimality condition, but in general not sufficient. In independent work, Beiglböck and Griessler found a closely related monotonicity principle which includes the result [12, Theorem 3.6] as a special case, see [3, Theorem 1.4].

Inspired by the classical c-monotonicity which shows that optimality is an attribute of the support of a coupling, other different monotonicity principles have been developed in the area of martingale optimal transport problems, cf. [2, 6]. The approach of the previous section for showing optimality is strongly inspired by the latter article which deals with time-continuous distribution-constrained optimal stopping problems where the underlying stochastic process is a Brownian motion. Analogously, it is possible to find a monotonicity principle for the time-discrete case. Again, we assume that $(Z_t)_{t\in I}$ is a stochastic process in discrete time with independent increments.

Definition 5.1. The set RST^t_κ of randomized stopping times (of a stochastic process Z with initial distribution κ) is defined as the set of all Z-adapted probability measures π on $R^{I_{\geq t}} \times I_{\geq t}$ such that $Z \sim \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\pi)$.

Definition 5.2 (Conditional randomized stopping times)**.** For $\pi \in \mathrm{RST}(\mu, \nu)$ and $(\omega, t) \in S$, we define $\pi^{(\omega,t)} \in \mathrm{RST}^t$ by defining a disintegration $(\pi^{(\omega,t)}_\theta)_{\theta\in\mathbb{R}^{I_{\geq t}}}$ with respect to $\tilde{Z}$ as

$$\pi_\theta^{(\omega,t)} := \begin{cases} \frac{1}{1-\pi_{(\omega,t)}(I_{\le t})}(\pi_{(\omega,t)\circ\theta})_{\restriction I_{\ge t}} & \text{for } \pi_{(\omega,t)}(I_{\le t}) < 1, \\ \delta_t & \text{for } \pi_{(\omega,t)}(I_{\le t}) = 1, \end{cases}$$

where δ_t is the Dirac measure concentrated at t and $\theta \in \mathbb{R}^{I_{\ge t}}$ with $\theta_1 = 0$.

Definition 5.3 (Relative Stop-Go pairs)**.** For $\xi \in \mathrm{RST}(\mu,\nu)$ define $\mathrm{SG}^\xi \subseteq (\mathbb{R}^I \times I) \times (\mathbb{R}^I \times I)$ as the set of all pairs $(\omega,t),(\eta,t) \in \mathbb{R}^I \times I$ such that there exist $\tilde{\xi}_1 \in \mathrm{RST}^t_{\delta_{\omega(t)}}$ and $\tilde{\xi}_2 \in \mathrm{RST}^t_{\delta_{\eta(t)}}$ such that

- $\mathrm{proj}_{\{t,\dots,T\}}(\xi^{(\omega,t)} + \xi^{(\eta,t)}) = \mathrm{proj}_{\{t,\dots,T\}}(\tilde{\xi}_1 + \tilde{\xi}_2)$,
- $\xi^{(\omega,t)}(c) + \xi^{(\eta,t)}(c) < \xi_1(c) + \xi_2(c)$.

Theorem 5.4 (Monotonicity Principle)**.** *Assume that π is a solution of* ($\textsc{OptStop}^\pi$), *then there is a measurable, $\mathbb{F}$-adapted set $\Gamma \subseteq \mathbb{R}^I \times I$ such that*

$$\pi(\Gamma) = 1$$

and

$$SG \cap (\Gamma^< \times \Gamma) = \emptyset,$$

where $\Gamma^< := \{(\omega,s) \in \mathbb{R}^I \times I \ : \ (\omega,t) \in \Gamma \text{ for some } t > s\}$.

Equipped with this general result, we can easily show optimality of the greedy strategy introduced in the last section. For this class of payoff functions in particular, it can be shown that monotonicity is already a sufficient condition for being an optimizer.

Corollary 5.5. *Let the payoff function c be given as*

$$c(\omega,t) = f(t)\omega_t \quad \omega \in \mathbb{R}^I,\ t \in I,$$

where $f : I \to \mathbb{R}^+$ is monotonously increasing. Then the greedy strategy $\pi \in \mathrm{RST}(\mu,\nu)$ is a maximizer of ($\textsc{OptStop}^\pi$)*. If $\xi \in \mathrm{RST}(\mu,\nu)$ satisfies the assertions of 5.4, then*

$$\mathrm{proj}_{X_t \times X_{t+1}}(\xi(\cdot,t)) = \mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot,t)) \quad \mathrm{proj}_{X_t\times X_{t+1}}(\mu)\text{-}a.e.,\ t \in I.$$

Proof. From the construction of π via quantiles, see Example 4.2, there exist $A_t := (-\infty, a_t] \subseteq \mathbb{R},\ t \in I$ such that $M_t = \prod_{s<t}[a_s,\infty) \times A_t \times \prod_{r>t}\mathbb{R}$, each a_t is minimal with

$$\mu\left(\sum_{s\le t} M_s\right) \ge \sum_{s \le t} \nu(s).$$

Let $\xi \in \mathrm{RST}(\mu,\nu)$ be optimal, then we denote the set of Theorem 5.4 with Γ. Assume that

$$\mathrm{proj}_{X_s\times X_{s+1}}(\xi(\cdot,s) - \pi(\cdot,s)) = 0 \quad \mathrm{proj}_{X_s \times X_{s+1}}(\mu)\text{-a.s.},\ s < t < T,$$

$$\mathrm{proj}_{X_t\times X_{t+1}}(\xi(\cdot,t)-\pi(\cdot,t))\neq 0 \quad \mathrm{proj}_{X_t\times X_{t+1}}(\mu)\text{-a.e..}$$

Then there exists $\omega\in\Gamma$ such that

$$\omega_s\in(a_s,\infty) \quad s<t, \quad \omega(t)\in[a_t,\infty) \text{ and } \xi_\omega(t)>0.$$

Since ξ has to preserve the marginals, there exists $\eta\in M_t\cap\Gamma$ such that $\xi_\eta(s)>0$ for an $s>t$, which yields $\eta_t<\omega_t$. We want to show that $((\eta,t),(\omega,t))\in\mathrm{SG}^\xi$ which would lead to a contradiction. Therefore, we construct $\xi^1\in\mathrm{RST}^t_{\delta_{\omega_t}}$ and $\xi^2\in\mathrm{RST}^t_{\delta_{\eta_t}}$ by defining two disintegrations

$$\xi^1_\theta(s):=\xi^{(\omega,t)}_\theta(t)\cdot\xi^{(\eta,t)}_\theta(s)+\begin{cases}0 & s=t,\\ \xi^{(\omega,t)}_\theta(s) & s>t.\end{cases}$$

$$\xi^2_\theta(s):=(1-\xi^{(\omega,t)}_\theta(t))\cdot\xi^{(\eta,t)}_\theta(s)+\begin{cases}\xi^{(\omega,t)}_\theta(t) & s=t,\\ 0 & s>t.\end{cases}$$

Computing the payoff yields

$$\begin{aligned}&\xi^1(c)-\xi^{(\omega,t)}(c)+\xi^2(c)-\xi^{(\eta,t)}(c)=\\ &-\int f(t)\omega_t\xi^{(\omega,t)}_\theta(t)\cdot(1-\xi^{(\eta,t)}_\theta(t))\mathbb{P}(\mathrm{d}\theta)+\sum_{t>s}\int f(s)(\omega_t+\theta_s)\xi^{(\omega,t)}_\theta(t)\cdot\xi^{(\eta,t)}_\theta(s)\mathbb{P}(\mathrm{d}\theta)\\ &+\int f(t)\eta_t\xi^{(\omega,t)}_\theta(t)\cdot(1-\xi^{(\eta,t)}_\theta(t))\mathbb{P}(\mathrm{d}\theta)-\sum_{t>s}\int f(s)(\eta_t+\theta_s)\xi^{(\omega,t)}_\theta(t)\cdot\xi^{(\eta,t)}_\theta(s)\mathbb{P}(\mathrm{d}\theta)\\ &=\sum_{s>t}\int(f(s)-f(t))(\omega_t-\eta_t)\xi^{(\omega,t)}_\theta(t)\xi^{(\eta,t)}(\mathrm{d}\omega,s)>0.\end{aligned}$$

Therefore, $((\eta,t),(\omega,t))\in\mathrm{SG}^\xi\cap\Gamma^<\times\Gamma$ which is a contradiction. □

Corollary 5.6. *Under the assumption of Corollary 5.5, the support of the greedy strategy $\pi\in\mathrm{RST}(\mu,\nu)$ introduced in Example 4.2 satisfies the assertions of Theorem 5.4.*

Proof. Let $((\eta,t),(\omega,t))\in\Gamma^<\times\Gamma$ and $\kappa:=\frac{1}{2}(\delta_{\eta_t}+\delta_{\omega_t})$, then $\frac{1}{2}(\pi^{(\eta,t)}+\pi^{(\omega,t)})=:\tilde\pi\in\mathrm{RST}^t_\kappa$ can be viewed as a greedy strategy to the auxiliary problem:

$$\text{Maximize } \xi\mapsto\xi(c) \text{ under}$$

$$\xi\in\mathrm{RST}^t_\kappa \text{ s.t. } \mathrm{proj}_{I_{\geq t}}(\xi-\tilde\pi)=0.$$

By applying Corollary 5.5 we obtain optimality of $\tilde\pi$, and hence

$$\mathrm{SG}^\pi\cap\Gamma^<\times\Gamma=\emptyset.$$

□

CHAPTER 6

Conclusions

In this thesis, a more general class of optimal stopping problems, see ($\textsc{OptStop}^{\gamma}$) and ($\textsc{OptStop}^{\pi}$), was examined. This type of problems naturally arises from ordinary optimal stopping problems, e.g., in financial and actuarial mathematics, where additional dependencies have to be modeled. Even though scientific literature concerning distribution-constrained optimal stopping problems is scarce, recent research proves an increasing interest in this topic as in [1, 6, 5, 4, 10].

In Chapter 2, the optimization problems ($\textsc{OptStop}^{\gamma}$) and ($\textsc{OptStop}^{\pi}$) were formally introduced and a link between them was made. Based on the theory of optimal transport, existence of optimizers was shown in Chapter 3. In addition, a Kantorovich-type duality theorem was developed, inspired by recent work of Zaev [12]. Chapter 4 deals with the optimization of a class of payoff functions. For this, an optimal strategy was found in Theorem 4.8. Finally, in Chapter 5 different geometric optimality criteria – so-called monotonicity principles – are formulated, which have their roots in the theory of optimal transport. It is shown how they can be adapted for ($\textsc{OptStop}^{\pi}$) and are applied to show optimality of the strategy introduced in Section 4.2.

We close with an outlook on future research possibilities. Efforts were made by Cox and Källblad [8, 10] in the area of time-continuous distribution-constrained optimal stopping problems to reformulate the problem using so-called measure-valued martingales. Hence, they were able to view it as a stochastic control problem, establishing a dynamic programming principle and deducing Hamilton-Jacobi-Bellman equations for it. In the time-discrete case, it would be interesting to further address the measure-valued martingale approach, since it is apriori not clear if it is adequate and/or if there is a more suitable alternative.

Bibliography

[1] E. Bayraktar and C. W. Miller. Distribution-Constrained Optimal Stopping. *ArXiv e-prints*, April 2016.

[2] M. Beiglböck, A. Cox, and M. Huesmann. Optimal transport and Skorokhod embedding. *Invent. Math., to appear*, 2016.

[3] M. Beiglböck and C. Griessler. A land of monotone plenty. *ArXiv e-prints*, April 2014.

[4] M. Beiglböck and N. Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016.

[5] Mathias Beiglböck, Alexander M. G. Cox, and Martin Huesmann. Optimal transport and skorokhod embedding. *Inventiones mathematicae*, pages 1–74, 2016.

[6] Mathias Beiglböck, Manu Eder, Christiane Elgert, and Uwe Schmock. Geometry of distribution-constrained optimal stopping problems, 2016.

[7] Mathias Beiglböck and Aldo Pratelli. Duality for rectified cost functions. *Calculus of Variations and Partial Differential Equations*, 45(1):27–41, 2012.

[8] Alexander MG Cox and Sigrid Källblad. Model-independent bounds for asian options: a dynamic programming approach. *arXiv preprint arXiv:1507.02651*, 2015.

[9] Claude Dellacherie and Paul-Andre Meyer. Probabilities and potentials a, 1978.

[10] Sigrid Källblad. A dynamic programming principle for distribution-constrained optimal stopping. *arXiv preprint arXiv:1703.08534*, 2017.

[11] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[12] D. Zaev. On the Monge-Kantorovich problem with additional linear constraints. *ArXiv e-prints*, 2014.