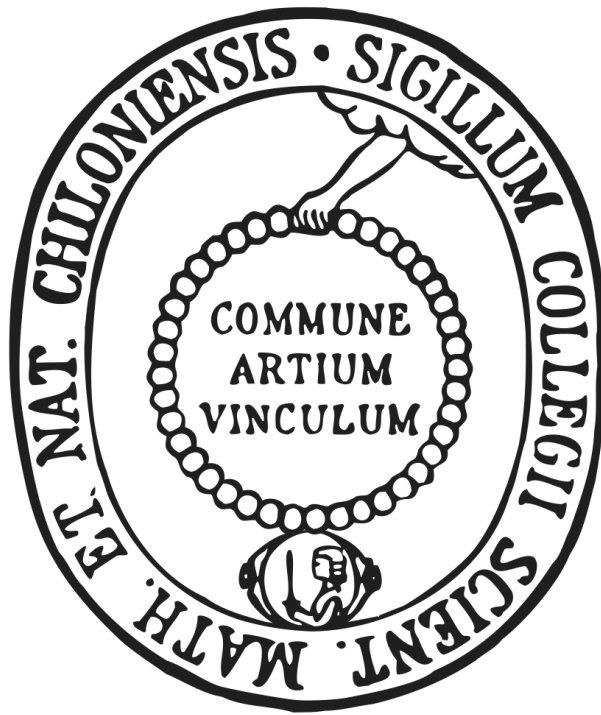


Pricing Bermudan Options by Forward Improvement Iteration



Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel
vorgelegt von

Julian Peter Lemburg

Kiel
Dezember 2010

Referent: Prof. Dr. Albrecht Irle

Korreferent: Priv.-Doz. Dr. Volkert Paulsen

Tag der mündlichen Prüfung: 31.01.2011

Zum Druck genehmigt: Kiel, 31.01.2011

Der Dekan, gez. Prof. Dr. Lutz Kipp

Ich danke Herrn Professor Doktor Albrecht Irl
herzlich für die gute Betreuung dieser Dissertation.

Mathematical Subject Classification MSC2010

62L15

Statistics

Sequential methods

Optimal stopping

60G40

Probability theory and stochastic processes

Stochastic processes

Stopping times; optimal stopping problems

60J05

Probability theory and stochastic processes

Markov processes

with discrete parameter

91G60

Game theory, economics, social and behavioral sciences

Mathematical finance

Numerical methods

Zusammenfassung

Gegenstand dieser Arbeit ist die Vorstellung, Analyse und Erweiterung des 1980 von Irle in [Irl80] veröffentlichten Forward Improvement Iteration (FII) Algorithmus und ähnlicher Algorithmen aus [BKS08], [BS06], [KS06] und [Pre10], sowie die Angabe von passenden Anwendungen.

Der FII Algorithmus löst Probleme des Optimalen Stoppens in diskreter Zeit, indem er in jedem Iterationsschritt eine untere Schranke der Snell Envelope und eine untere Approximation einer optimalen Stoppzeit ausgibt, wobei die Ausgaben gegen die Snell Envelope bzw. eine optimale Stoppzeit konvergieren.

Die oben genannten, verwandten Algorithmen führen teilweise zusätzliche Parameter für den jeweils behandelten Spezialfall ein. Diese Arbeit zeigt auf, dass sich alle diese Algorithmen zu einem einzigen erweiterten FII Algorithmus zusammenfassen lassen, der die zusätzlichen Parameter übernimmt und bei dem die in den jeweils vorgestellten Varianten aufgezeigten Resultate ebenfalls gelten.

Ferner wird der FII Algorithmus in dieser Arbeit (einschließlich sinnvoll übertragbarer Zusatzparameter) auf den Markov-Fall übertragen. Abgerundet wird die Arbeit von einigen numerischen Beispielen und der Anwendung auf den Monotonie-Fall sowie auf ein Sekretärinnenproblem, wobei die Lösungen der Spezialfälle aus [Irl80] zu einer allgemeinen Lösung erweitert werden.

Abstract

The object of this thesis is the presentation, analysis and extension of the Forward Improvement Iteration Algorithm (FII) published 1980 by Irle in [Irl80] and similar algorithms in [BKS08], [BS06], [KS06] und [Pre10], as well as giving adequate applications.

The FII algorithm solves problems of optimal stopping in discrete time by giving a lower approximation of the optimal stopping time and a lower bound of the Snell envelope in each iteration step, which converge to the Snell envelope resp. optimal stopping time.

The similar algorithms already mentioned above introduced additional parameters for the special cases they treated. This thesis shows that these algorithms can be put together into a single extended FII algorithm, inheriting the additional parameters and results.

In this thesis the algorithm is transferred to the Markov case, also inheriting reasonable additional parameters. We finish the work with some numerical applications and an investigation of the monotone case and a secretary problem, in which the solutions of the special cases in [Irl80] are extended to a general solution.

Contents

Zusammenfassung	5
Abstract	5
Thesis Outline	10
Preface	10
1 Setting and Definitions	14
1.1 Conventions	14
1.2 Setting	14
1.3 Definition and Remark: Stopping Rule and Snell Envelope	14
1.4 Lemma: Integrability	15
1.5 Theorem: Snell Envelope	15
1.6 Theorem: Convergence of the Snell Envelope	15
1.7 Theorem: Backward Dynamic Programming of the Snell Envelope	16
1.8 Convention	16
1.9 Convention: Fixed N	16
1.10 Definition: (Consistent) N -Suitable Family of Stopping Rules	16
1.11 Definition: Optimal	16
1.12 Lemma: Consistent, Optimal	17
1.13 Definition and Theorem: First and Last Optimal N -Suitable Family of Stopping Rules	17
2 Forward Improvement Iteration Algorithm	18
2.1 Definitions	18
2.2 Definition: N -suitable Adapted Random Set and Generated Family	19
2.3 Definition and Theorem: Corresponding	19
2.4 Remark: Corresponding	20
2.5 Definition: Corresponding To	20
2.6 Definition and Remark: Essential	21
2.7 Lemma: Elements of \mathcal{T} compared	21
2.8 The Window Parameter	22
2.9 Input for FII Algorithm	22
2.10 Remark: Input for FII Algorithm	22
2.11 Output of FII Algorithm	23
2.12 Remarks: Output of FII Algorithm	23
2.13 Performing FII Algorithm	24
2.14 Remarks: Performing FII Algorithm	25
2.15 Comparison with the Literature	25
2.16 Lemma: Elementary Equations and Inequalities	25
2.17 Lemma: Existence of Improvers	26
2.18 Definition: Improver	26
2.19 Setting of [Irl80] included	27
2.20 Remark: Infinite Case Includes Finite Case	27

3	Finite Time Horizon	28
3.1	Theorem: Equations and Inequalities	28
3.2	Lemma	30
3.3	Theorem	31
3.4	Conclusion	32
3.5	Example	33
3.6	Lemma	34
3.7	Theorem	35
3.8	Theorem	36
3.9	Remark	37
3.10	Theorem	37
3.11	Lemma	38
3.12	Lemma	38
3.13	Conclusion	39
3.14	Theorem: Optimal Algorithm Termination for $C^{(0)} \equiv \Omega$	39
3.15	Definition and Remark	41
3.16	Example	41
3.17	Generalized Example	41
3.18	Theorem: Algorithm Termination	42
3.19	Example	43
3.20	Lemma	43
3.21	Theorem: Influence of κ	44
4	Different Finite Time Horizons Compared	45
4.1	Idea	45
4.2	Example	45
4.3	Lemma	46
4.4	Lemma	47
4.5	Theorem	48
5	General Case	49
5.1	Theorem from [Irl80] for $\kappa \equiv 1$	49
5.2	General Assumption	50
5.3	Main Theorem	50
5.4	Remark: Limites	51
5.5	Remarks	51
5.6	Lemma	51
5.7	Remarks	52
5.8	Definitions	52
5.9	Lemma: Optimality	52
5.10	Lemma: One Step Improvement	53
5.11	Theorem	57
5.12	Conclusion	60
5.13	Conclusion	61
5.14	Remark	61
5.15	Outlook	61
6	Markovian Case	62
6.1	Setting	62
6.2	Time-Independent Pay-Off	62
	6.2.1 Connection to the General Case	63
6.3	Time-Dependent Pay-Off	64
	6.3.1 Finite Horizon	64
	6.3.2 Linear Costs	64

6.3.3	Constant Discounting	65
6.4	Path-Dependent Costs	65
6.5	Performing the Algorithmic Step	66
6.6	Outlook	66
7	Markovian Case with Random Discounting	67
7.1	Motivation	67
7.2	Setting	68
7.3	Remarks	69
7.4	Lemma: One-Step Improvement	69
7.5	Example: Insufficient Improvement	69
7.6	Lemma: Partial Improving I	71
7.7	Lemma: Partial Improving II	72
7.8	Example: Necessities for Partial Improving II	72
7.9	Proof of Lemma 7.4 (page 69)	74
7.10	Remark	76
7.11	Theorem	76
8	Markovian Case with Random Discounting and Finite State Space	80
8.1	Setting	80
8.2	Theorem	80
8.3	Corollar	81
8.4	FII by Solving Linear Equations	81
9	Numerical Examples	84
9.1	Setting	84
9.2	Results for $\kappa \equiv \{1\}$	85
9.3	Influence of κ	87
10	Monotone Case	90
10.1	Definition	90
10.2	Remark: Appearance in the above Model	91
10.3	Definition and Remark: m -monotone and m -stages look-ahead rules	91
10.4	Lemma: m -monotone	92
10.5	Conclusion	92
10.6	Counter-Example: m -monotone	93
10.7	Lemma	94
10.8	Theorem	94
10.9	Lemma	95
10.10	Markov Case	95
11	No-Information Version of the Best Choice Problem with Random Population Size	96
11.1	General Setting	96
11.2	Special Setting	97
11.3	Definitions and Remarks	97
11.4	Lemma	98
11.5	Lemma	99
11.6	Lemma	99
11.7	Lemma	99
11.8	Lemma	100
11.9	Lemma	100
11.10	Theorem	101
11.11	Theorem	102
11.12	Setting Continued	103

11.13	Theorem	104
11.14	Lemma	104
11.15	Lemma: Optimality	104
11.16	Lemma: Problem-Independent Positive Values of a	105
11.17	Conclusion about a and a_κ	105
11.18	Remark	106
11.19	Lemma	106
11.20	Conclusion: Reasonable Minimal Value for d	106
11.21	Lemma: Bounded Values	106
11.22	Conclusion: Bounded Values	107
11.23	Lemma: Limit Value Zero	107
11.24	Theorem	107
11.25	Theorem	108
11.26	Lemma	108
11.27	Theorem: Deterministic Population Size	109
11.28	Conclusion	110
11.29	Remark	111
11.30	Conclusion	111
	Bibliography	112
	Eidesstattliche Erklärung	113
	Lebenslauf	114
	Curriculum Vitae	115

Thesis Outline

After a detailed introduction we will show necessary basics in Chapter 1 (page 14). The algorithm with all its extensions will be presented in Chapter 2 (page 18). In Chapter 3 (page 28) and Chapter 4 (page 45) its application given finite time horizon is done and in Chapter 5 (page 49) its application given arbitrary time horizon. The algorithm is transferred to the Markovian case in Chapter 6 (page 62), Chapter 7 (page 67) und Chapter 8 (page 80). Numerical examples follow in Chapter 9 (page 84). The application in the “monotone case” is shown in Chapter 10 (page 90) and some advantages of the algorithm are shown for a best choice problem in Chapter 11 (page 96).

Preface

At a stock exchange, not only stocks change hands, but also bets on stocks. These are called financial products or derivatives. Quite common is the following bet: You pay a certain price and receive the right to say stop at one of some given time points and then receive a payoff directly depending on the current price of one or more other financial products, often plain stocks. This bet is called a Bermudan option.

For example the payment is obtained as the amount, by which the actual price of a particular share exceeds a predetermined value (called strike price). This is often called Bermudan call option. Ideally it is to stop on the date on which the payment has its highest value. – But at each time point it is unclear whether tomorrow would be an even better time point. It is therefore necessary to find a rule, when you should stop. The highest expected payoff that a rule can bring is accordingly the fair price of the option, but how is it determined or approximated?

Pricing

Pricing of Bermudan style derivatives on a high dimensional system of underlyings is considered an enduring problem for the last years. Prices for such high dimensional options are difficult, if not impossible, to compute by standard partial differential equation (PDE) methods. For high dimensional European options a good alternative to PDEs is Monte Carlo simulation.¹ Nevertheless, for American options, Monte Carlo simulation is more complicated since the (optimal) exercise boundary is usually unknown.

Various Monte Carlo algorithms for pricing American and Bermudan options have been developed and described in the literature. For a detailed and general overview we refer to [Gla03] and [Put94] and the references therein. Many of these algorithms are related to backward dynamic programming which comes down to a recursive representation of the Snell envelope. They require the evaluation of high order nestings of conditional expectations.

Therefore Monte Carlo estimators for regression functions, which do not run into explosive cost when nested several times, have been proposed by several authors.

As an alternative to backward dynamic programming, one may search for a suitable parametric family of exercise boundaries and then maximize the solutions of the corresponding family of

¹The name goes back to John von Neuman (1903-1957), who used “Monte Carlo” in 1946 as a code name for a secret project, wherein he used these techniques. The name refers to the Monte Carlo casino in Monaco.

boundary value problems over the parameters.

Policy Iteration – Approximations Already After the First Iteration

Another alternative to solving the backward dynamic program recursively are policy iterations. Their main advantage is that they mend one of the main drawbacks of the backward scheme: Assume exercise can take place at one out of k time instances. To obtain the value of the optimal stopping problem via backward dynamic programming, we have to calculate nested conditional expectations of the order k . The k th nested conditional expectation has to be evaluated before any approximation of the time 0 value is possible. This prevents the use of plain Monte-Carlo simulations for approximating the conditional expectations and requires more complicated approximation procedures for these quantities. In contrast, policy iteration algorithms yield already after each iteration step an improved approximation of the Snell envelope which ranges over all exercise dates. This allows to calculate approximations of the Snell envelope of a Markovian process via a plain Monte-Carlo simulation.

Forward Improvement Iteration

A classical method of policy iteration, which is well-known in dynamic programming, was developed by Howard in [How60].² Adapting this method to problems of optimal stopping yields a method first published by Irle in [Irl80] and later called “forward improvement iteration” (FII) (see [Irl06] and [Irl09]). The mathematical model of optimal stopping problems is considerably less involved than the general model of dynamic programming. Hence Howard’s method has a simple form and the optimal stopping problem can be solved by the resulting algorithm without any restrictions on the underlying state space and distribution of the process. We will discuss this algorithm further in this thesis.

In the method of backward induction we have to compare the actual pay-off with the pay-off due to *optimal* continuation first at the last stage, then at the second-last stage and so forth backwards in time. Instead of that the forward improvement algorithm starts with a sequence of attainable sets and compares the pay-off at each stage with the pay-off due to stopping at the next attainable set. Repeating this, one arrives at a sequence of sets which characterizes an optimal stopping rule. By using stopping rules, Bender, Kolodko and Shoenmakers have developed a similar approach in [KS06] and [BS06].³ Some random sets appear in [KS06, part 6] as additional parameters. We will show below that these papers describe in fact the same algorithm and the same idea as already published in [Irl80], just seen from another viewpoint. *We will show the existence of a one to one correspondence between the above mentioned attainable sets, the mentioned sequences of stopping rules and the mentioned random sets.*

The central result in the above mentioned papers ([Irl80], [Irl06], [Irl09], [KS06] and [BS06]) is an iterative construction of the Snell envelope via a sequence of attainable sets and a corresponding sequence of stopping rules which increases to the (first) optimal stopping rule (and equivalently to a sequence of sets and a sequence of random sets), corresponding to the optimal stopping rule. In each iteration step we improve a whole family of stopping rules, attainable sets and random sets. This family is denoted with (τ_i, C_i, Θ_i) where i runs through the set of exercise dates. τ_i is the stopping rule for the Bermudan which is not exercised before date i . The thus obtained sequence of stopping families naturally induces a non-decreasing sequence of lower approximations of the Snell envelope. When looking at a model with finitely many exercise dates, the sequence even coincides with the Snell envelope after finitely many steps: The number of necessary steps is at most the number of exercise dates.

A canonical suboptimal exercise policy for Bermudan swaptions is: exercise as soon as the cash-flow dominates all the Europeans ahead. This is the first iteration step of the FII algo-

²See the explanations in [BKS08, Introduction], [KS06, Part 1] and [Irl80, page 179].

³See the abstracts in [BS06] and [BKS08].

rithm (when starting with the largest possible attainable set). The algorithm is just iterating this policy. Schoenmakers states that even the first iteration is usually “already not far from optimal”. Due to its special nature the FII algorithm gives usually good results with only very few iterations.⁴

Continuous Time

The whole analysis is based on a discrete set of exercise dates, but since a continuous time American option may be approximated by a Bermudan option with a fine grid of exercise dates, one may in principle apply the method as well.

Infinite Time Horizon

In [KS06] it was depicted that the therein shown algorithm for a finite set of exercise dates can be used for an perpetual discrete optimal stopping problem by approximation through instances with a finite set of exercise dates. But this is not necessary! Irle highlighted already in [Irl80] that the procedure has the advantage that problems with infinite time horizon can be treated directly and easily. Clearly, most of the proofs in [KS06] and [BS06] are based on backward induction which cannot be transferred to problems of infinite time-horizon. Other techniques are necessary, the framework for these was developed in [Irl80] and further developed and explained in [Irl06] and [Irl09].

Window Parameter

Irle used a window parameter implicitly set to one for the algorithm when describing it in [Irl80]. Bender and Schoenmakers used a similar parameter, which in their work [BS06] is set to the number of time-steps. In [KS06] Kolodko and Schoenmakers explicitly mentioned this window parameter, called it κ and examined its influence for the finite case on the *first* iteration step. *We will show for all iteration steps that in the finite case a larger κ is at least as good as using a smaller one. We will show empirically that a small κ around 4 increases the speed of the algorithm (when using linear equations). We will examine the influence of κ for the infinite case and show that the choice does not change the convergence results formerly shown by Irle.* The window parameter is roughly spoken the number of time points we look ahead in each iteration step to improve our results:

For each iteration of the algorithm step we fix some window parameter k . Let l be the minimum of k and the number of remaining time points ahead. We look l time points ahead in this manner: We compare the current pay-off with the expected pay-off when we do not use the next 1, 2, 3, ... l exercise possibilities and stop then according to the best stopping rule we have calculated until now. The window parameter k is the maximal number of time points we look ahead. It may be chosen differently in each step and may even depend on the foregoing iteration step.

Scenario Selection Method

In principle, the forward improvement iteration algorithm can be started with a very simple input policy like “exercise immediately”. Given today’s computer power however, more than one degree of nesting conditional expectations is virtually impossible (when using Monte Carlo). This means that in practice only one step of the algorithm can be carried out.⁵ Therefore, the choice of the input stopping family is important.

The basic idea is as follows: Suppose the holder of the option has some pre-information, for example he knows good closed form approximations of the price for the corresponding European options due to their availability in the literature for practically all relevant options. So the investor may rule out some scenarios, at which an optimal strategy cannot exercise, by the

⁴See [KS06, Introduction] and [Irl09, Introduction].

⁵See [BKS08, Introduction].

pre-information. Then the policy improvement is run only at the remaining time points. The set of remaining time points depends on the state of the underlying system, thus we do not simply reduce to an other option with less exercise dates.

This scenario selection method has already been described in [Irl80] and the subsequent papers of Irle and was depicted in [BKS08, part 3] in a slightly different formulation. *We will show their equivalence.*

1 Setting and Definitions

1.1 Conventions

- If we introduce a new term, we write it in bold letters.
- If we introduce a new variable for permanent use in this thesis, we write it in bold letters.
- We write just a for families $(a_i)_{i \in I}$ with index set I , wherever I is obvious by the context.

1.2 Setting

Let us consider the following general situation for a problem of optimal stopping. We start with some probability space (Ω, \mathcal{A}, P) and a filtration $(\mathcal{A}_n)_{0 \leq n < \infty}$ with \mathcal{A}_0 being trivial and an adapted real-valued stochastic process $(X_n)_{0 \leq n < \infty}$, such that X_n is integrable¹ for all $n \in \mathbb{N}_0$ and $X_\infty = \limsup_{n \rightarrow \infty} X_k$. Define $\mathbf{X}^N := (X_n)_{0 \leq n \leq N}$ for all $N \in \mathbb{N}_0 \cup \{\infty\}$.

1.3 Definition and Remark: Stopping Rule and Snell Envelope

A stopping rule is a mapping $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ satisfying $\{\tau = n\} \in \mathcal{A}_n$ for all $n \in \mathbb{N}_0$. Let \mathcal{S} denote the set of all stopping rules. Some subsets of the set of all stopping rules will be of special interest. So define for all $n \in \mathbb{N}_0$

$$\begin{aligned} \mathcal{S}_n^\infty &:= \{\tau \in \mathcal{S} ; n \leq \tau \leq \infty, E(X_\tau^-) < \infty\}, \\ \mathcal{S}_n^b &:= \{\tau \in \mathcal{S} ; \exists N \in \mathbb{N}_0 \ n \leq \tau \leq N\}, \\ \mathcal{S}_n^* &:= \{\tau \in \mathcal{S} ; n \leq \tau < \infty, E(X_\tau^+) < \infty \text{ or } E(X_\tau^-) < \infty\}, \\ \overline{\mathcal{S}}_n &:= \{\tau \in \mathcal{S} ; n \leq \tau < \infty, E(X_\tau^-) < \infty\}, \end{aligned}$$

and for all $n, N \in \mathbb{N}_0$ with $n \leq N$ define

$$\mathcal{S}_n^N := \{\tau \in \mathcal{S} ; n \leq \tau \leq N\}.$$

The definition of \mathcal{S}_n^∞ is already done above and slightly different.

We have

$$(1.3.1) \quad \mathcal{S}_n^b = \bigcup_{n \in \mathbb{N}_0} \mathcal{S}_n^N.$$

Distinguishing different sets of stopping rules in the finite case (meaning N is finite) is not necessary due to the fact that for all $N \in \mathbb{N}_0$, $n \in \mathbb{N}_0$ with $n \leq N$ and for all $\tau \in \mathcal{S}_n^N$ the random variable X_τ is integrable, as we will show in Lemma 1.4 (page 15) below.

Furthermore define for all $N \in \mathbb{N}_0 \cup \{\infty\}$ the process $Y_N^* = (Y_{N,n}^*)_{0 \leq n \leq N}$ by

$$\mathbf{Y}_{N,n}^* := \text{esssup} \{E(X_\tau | \mathcal{A}_n) ; \tau \in \mathcal{S}_n^N\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N,$$

often called **Snell envelope of \mathbf{X}^N** named after James Laurie Snell (born January 15, 1915), a student of Joseph Leo Doob (1910-2004), due to the article [Sne52] he published in 1952.

¹See the Appendix for some common definitions.

1.4 Lemma: Integrability

(1.4.1) X_τ is integrable for all $\tau \in \mathcal{S}_0^b$,

(1.4.2) $\sup_{m \in \mathbb{N}_0} |X_{\sigma_m}|$ is integrable for all $\sigma \in (\mathcal{S}_n^N)^{\mathbb{N}}$ for all $N, n \in \mathbb{N}_0$ with $n \leq N < \infty$

Proof: Let $n \in \mathbb{N}_0$. Due to Formula (1.3.1) (page 14) it suffices to concentrate on \mathcal{S}_n^N for some $N \in \mathbb{N}_0$ instead of examining \mathcal{S}_0^b to prove (1.4.1). Hence consider $N \in \mathbb{N}_0$ with $n \leq N < \infty$. For all $l \in \mathbb{N}_0$ with $l \leq N$ we have $|X_l|$ integrable by Setting 1.2 (page 14). Define $W := \sum_{l=0}^N |X_l|$. As a sum of integrable random variables W is integrable. Since

$$(1.4.3) \quad |X_\tau| \leq \max \{|X_l| ; 0 \leq l \leq N\} \leq \sum_{l=0}^N |X_l| = W \text{ for all } \tau \in \mathcal{S}_n^N,$$

we have the integrability in (1.4.1).

For all $m \in \mathbb{N}_0$ we have $|X_{\sigma_m}| \leq W$ by (1.4.3) and thus

$$0 \leq \sup \{|X_{\sigma_m}| ; m \in \mathbb{N}_0\} \leq W,$$

hence we have the integrability in (1.4.2). □

1.5 Theorem: Snell Envelope

We have for all $n \in \mathbb{N}_0$

$$\begin{aligned} Y_{\infty,n}^* &= \text{esssup} \{E(X_\tau | \mathcal{A}_n) ; \tau \in \mathcal{S}_n^{\infty}\} \\ &= \text{esssup} \{E(X_\tau | \mathcal{A}_n) ; \tau \in \mathcal{S}_n^*\} \\ &= \text{esssup} \{E(X_\tau | \mathcal{A}_n) ; \tau \in \overline{\mathcal{S}}_n\} \end{aligned}$$

and these essential suprema are not enlarged by allowing “randomized” stopping rules.

For all $N \in \mathbb{N}_0$ and under certain assumptions also for $N = \infty$ the process Y_N^* is the smallest X^N -dominating supermartingale.

Proof: See for example [CRS71, pp. 42, 63, 78 for definitions and pp. 63, 80, 81 for results] and [CRS71, p.111,th.5.3]. Beware the different naming of the variables: $\overline{\mathcal{S}}_n \triangleq C_n^{\infty}$, $\mathcal{S}_n^{\infty} \triangleq \overline{C}_n$, $\mathcal{S}_n^* \triangleq C_n^*$. □

1.6 Theorem: Convergence of the Snell Envelope

We have convergence of the Snell Envelope in the sense of

$$Y_{\infty,n}^* = \lim_{N \rightarrow \infty} Y_{N,n}^* = \text{esssup} \left\{ E(X_\tau | \mathcal{A}_n) ; \tau \in \mathcal{S}_n^b \right\} \text{ for all } n \in \mathbb{N}_0,$$

if one of the following conditions is true:

- (a) $X_n \geq 0$ for all $n \in \mathbb{N}_0 \cup \{\infty\}$,
- (b) $(X_n)^- \leq W$ for all $n \in \mathbb{N}_0 \cup \{\infty\}$ for some integrable random variable $W \geq 0$,
- (c) $((X_n)^-)_{n \in \mathbb{N}_0}$ is uniformly integrable
- (d) $\lim_{n \rightarrow \infty} \int_{\{\tau > n\}} (X_n)^- dP = 0$ for all $\tau \in \overline{\mathcal{S}}_0$,
- (e) $\lim_{n \rightarrow \infty} \inf_{k \geq n} \int_{\{\tau > k\}} (\lim_{N \rightarrow \infty} Y_{N,k}^*)^- dP = 0$ for all $\tau \in \overline{\mathcal{S}}_0$,

Remark: We have (a) \implies (b), (b) \implies (c), (c) \implies (d), (d) \implies (e).

Proof: A proof can be found in [CRS71, p. 68-70]. □

1.7 Theorem: Backward Dynamic Programming of the Snell Envelope

For finite N the Snell envelope Y_N^* can be constructed by backward dynamic programming as follows: At the last exercise date N define

$$Y_{N,N}^* := X_N,$$

and for all $n \in \mathbb{N}_0$ with $n < N$ and $Y_{N,n+1}^*$ already defined set

$$Y_{N,n}^* := \max \{X_n, E(Y_{N,n+1}^* | \mathcal{A}_n)\}.$$

Proof: A proof can be found in [Irl02, pp. 92-95]. □

1.8 Convention

For being able to write about the finite and infinite case simultaneously we often use the expression

$$"n < N + 1",$$

by assuming $\infty + 1 := \infty$ it is equivalent to

$$"n \leq N \text{ for finite } N \text{ and } n < \infty \text{ for } N = \infty".$$

1.9 Convention: Fixed N

Fix some $N \in \mathbb{N}_0 \cup \{\infty\}$. Nearly all of the following variables depend on N . For brevity we will often abstain from showing this dependency by writing N as an index. So we write for example Y_n^* for $Y_{N,n}^*$.

1.10 Definition: (Consistent) N -Suitable Family of Stopping Rules

An N -suitable family of stopping rules is a family

$$\tau = (\tau_n)_{0 \leq n < N+1} \text{ with } \tau_n \in \mathcal{S}_n^N \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N + 1.$$

Using a definition introduced in [BKS08, def. 3.1]

it is called **consistent** iff

$$\{\tau_n > n\} \subseteq \{\tau_n = \tau_{n+1}\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N,$$

hence (since $\tau_{n+1} \geq n + 1$) iff

$$\{\tau_n > n\} = \{\tau_n = \tau_{n+1}\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N.$$

Define the set of all consistent N -suitable families of stopping rules by

$$\mathcal{T} := \{(\tau_n)_{0 \leq n < N+1} ; \forall 0 \leq n < N + 1 \tau_n \in \mathcal{S}_n^N, \forall 0 \leq n < N \{\tau_n > n\} \subseteq \{\tau_n = \tau_{n+1}\}\}.$$

1.11 Definition: Optimal

Consider $\hat{\mathcal{S}} \subseteq \mathcal{S}$ and $n \in \mathbb{N}_0$ with $n < N + 1$.

Any $\tau \in \mathcal{S}$ is called **optimal in $\hat{\mathcal{S}} \cap \mathcal{S}_n^N$** iff

$$\tau \in \hat{\mathcal{S}} \cap \mathcal{S}_n^N \text{ and } E(X_\tau | \mathcal{A}_n) = \text{esssup} \left\{ E(X_\sigma | \mathcal{A}_n) ; \sigma \in \hat{\mathcal{S}} \cap \mathcal{S}_n^N \right\}.$$

Hence a $\tau \in \mathcal{S}$ is optimal in \mathcal{S}_n^N iff

$$\tau \in \mathcal{S}_n^N \quad \text{and} \quad \mathbb{E}(X_\tau | \mathcal{A}_n) = Y_n^*.$$

Any $\tau \in \mathcal{T}$ is called **optimal in $\widehat{\mathcal{S}}$** iff

$$\tau_n \text{ is optimal in } \widehat{\mathcal{S}} \cap \mathcal{S}_n^N \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N + 1,$$

and called **optimal** iff

$$\tau_n \text{ is optimal in } \mathcal{S}_n^N \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N + 1.$$

Hence not mentioning $\widehat{\mathcal{S}}$ is a shortcut if $\widehat{\mathcal{S}} = \mathcal{S}$.

1.12 Lemma: Consistent, Optimal

- Each optimal N -suitable family of stopping rules is consistent.
- For all $\tau \in \mathcal{T}$ we have

$$\tau_n \leq \tau_{n+1}$$

for all $n \in \mathbb{N}_0$ with $n < N$.

- For all $\tau \in \mathcal{T}$ and $\omega \in \Omega$ the sequence $(\tau_n(\omega))_{n \in \mathbb{N}_0}$ is non-decreasing and piecewise constant in the following sense. For all $n \in \mathbb{N}_0$ with $n < N$ we have:
If $\tau_n(\omega) < \infty$ then there is a minimal $m = m(\omega) \in \mathbb{N}_0$ such that we have

$$\tau_n(\omega) = \tau_{n+k}(\omega) = \tau_{n+m}(\omega) = n + m \text{ for all } k \in \mathbb{N}_0 \text{ with } k < m.$$

If $\tau_n(\omega) = \infty$ then $\tau_{n+k}(\omega) = \infty$ for all $k \in \mathbb{N}_0$.

Proof: This follows easily by induction from the definition of \mathcal{T} in Definition 1.11 (page 16). \square

1.13 Definition and Theorem:

First and Last Optimal N -Suitable Family of Stopping Rules

Define

$$\begin{aligned} \check{\tau}^* &:= (\check{\tau}_n^*)_{0 \leq n < N+1} := \left(\inf \left\{ n ; n \leq i \leq N, Y_i^* \leq X_i \right\} \right)_{0 \leq n < N+1} \\ &= \left(\inf \left\{ n ; n \leq i \leq N, Y_i^* = X_i \right\} \right)_{0 \leq n < N+1}, \\ \hat{\tau}^* &:= (\hat{\tau}_n^*)_{0 \leq n < N+1} := \left(N \wedge \inf \left\{ n ; n \leq i \leq N, \mathbb{E} \left(Y_{i+1}^* | \mathcal{A}_i \right) < X_i \right\} \right)_{0 \leq n < N+1}. \end{aligned}$$

- Consider finite N . Then $\check{\tau}^*$ is the **first** and $\hat{\tau}^*$ the **last** optimal consistent N -suitable family of stopping rules: For each optimal consistent N -suitable family of stopping rules $\tau = (\tau_n)_{0 \leq n < N+1} \in \mathcal{T}$ and for all $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$\check{\tau}_n^* \leq \tau_n \leq \hat{\tau}_n^*.$$

For all $n \in \mathbb{N}_0$ with $n \leq N$ and ρ being optimal in \mathcal{S}_n^N we have

$$\check{\tau}_n^* \leq \rho \leq \hat{\tau}_n^*$$

- Consider infinite N , $\check{\tau}_n^*$ a.s. finite for all $n \in \mathbb{N}_0$ and $E(\sup_{n \in \mathbb{N}_0} X_n^+) < \infty$, then $\check{\tau}^*$ is the **first** optimal consistent N -suitable family of stopping rules.

Proof: See for example [Nev75, pp. 120-129] and [BKS08, part 2]. \square

2 Forward Improvement Iteration Algorithm

The following algorithm, whose conception goes back to Howard’s policy improvement of dynamic programming [How60], was treated by Irlé in [Irl80], [Irl06] and [Irl09] for any N and therein called *forward improvement iteration* (FII). Schoenmakers et al. treated a similar algorithm for finite N in [BS06], [KS06] and [BKS06] without naming it explicitly. They introduced a window parameter κ equal throughout all iteration steps in their papers. Irlé used in his papers implicitly the window parameter one. We will assume herein a window parameter function, arbitrary chosen in each iteration step. We will call the algorithm, whenever mentioned herein, “forward improvement iteration algorithm” and it will be presented in this section.

In Definition 2.1 (page 18) and Definition 2.2 (page 19) we will define some sets herein called \mathcal{C} and \mathcal{Z} . The first is used with the same name intensively in the papers of Irlé. The latter is implicitly introduced and used in [BKS08, def. 3.1 ff.]. There is an interesting connection between these and \mathcal{T} , not mentioned therein, which we will call *corresponding* and examine it in Definition 2.3 (page 19) and Remark 2.4 (page 20). In Definition 2.6 (page 21) we define the term *essential* and examine it in Lemma 2.7 (page 21).

In Part 2.8 (page 22) we will introduce the above mentioned window parameter in detail. In Part 2.9 (page 22) and Remark 2.10 (page 22) we discuss the other input variables of the algorithm, the output is examined in Part 2.11 (page 23) and Remark 2.12 (page 23). In Algorithm 2.13 (page 24) we present the algorithm, and we state that in each iteration level the algorithm gives corresponding and convergent values in Remark 2.14 (page 25).

In Part 2.15 (page 25) we will compare the here stated algorithm with the similar algorithms stated before in the literature. In Lemma 2.16 (page 25) we state some elementary equations and inequalities. We define *improvers* in Definition 2.18 (page 26) and show their existence in Lemma 2.17 (page 26).

We finish the chapter with the proof in Part 2.19 (page 27) that the setting of the papers of Irlé is included in the setting of this chapter and that whenever we are talking about the infinite case the finite case is included in a natural way in Remark 2.20 (page 27).

Remember that we fixed some N with $0 \leq N \leq \infty$ in Convention 1.9 (page 16).

2.1 Definitions

Define

$$\mathcal{C} := \{(C_n)_{0 \leq n \leq N} ; C_N = \Omega, \forall 0 \leq n < N \ C_n \in \mathcal{A}_n\},$$

and for all $C \in \mathcal{C}$

$$\mathcal{S}(C) := \{\tau \in \mathcal{S} ; \forall 0 \leq n < N \ \{\tau = n\} \subseteq C_n\}.$$

2.2 Definition:

N -suitable Adapted Random Set and Generated Family

A random set $A : \Omega \rightarrow \mathcal{P}ot(\{0, \dots, N\})$ such that

$N \in A$ a.s. and

for all $n \in \mathbb{N}_0$ with $n < N$ the function $1_A(n)$ is \mathcal{A}_n -adapted resp. $\{\omega \in \Omega ; n \in A(\omega)\} \in \mathcal{A}_n$

is called an **N -suitable adapted random set**.

For every N -suitable adapted random set Θ_0 define

$$\Theta := (\Theta_n)_{0 \leq n < N+1} := (\Theta_0 \cap \{j ; n \leq j \leq N\})_{0 \leq n < N+1},$$

called **the family generated by Θ_0** , it is a family of N -suitable adapted random sets.

Define \mathcal{Z} as the set of the families generated by all N -suitable adapted random sets.

2.3 Definition and Theorem: Corresponding

A striking difference between the papers outlining the herein stated algorithm before ([Irl80], [Irl06], [Irl09], [BS06], [KS06] and [BKS06]) is that each of them puts the focus on only one of the sets \mathcal{C} , \mathcal{Z} and \mathcal{T} (see Definition 1.10 (page 16)), and does not explain the connection between them. Thereby it is not easy to see the similarities and differences of the papers.

We will show now that there is a certain one-to-one correspondence between these sets, which conserves all important properties. We will describe below the correspondence in detail by specifying adequate mappings.

We will say that elements of \mathcal{C} , \mathcal{Z} and \mathcal{T} are **corresponding**, if one can be constructed from the other by applying one of these mappings.

First we show that there is a mapping from \mathcal{C} to \mathcal{Z} and another one from \mathcal{Z} to \mathcal{C} so that their composition in both orders is the identity on \mathcal{C} or \mathcal{Z} respectively.

- Consider some $C \in \mathcal{C}$. Define

$$\Theta_0 \text{ by } \Theta_0(\omega) := \{j ; 0 \leq j \leq N, \omega \in C_j\} \text{ for all } \omega \in \Omega$$

or equivalently

$$\Theta_0 := \{j ; 0 \leq j \leq N, \mathbf{1}_{C_j} = 1\}$$

and further for all $n \in \mathbb{N}_0$ with $n < N + 1$

$$\begin{aligned} \Theta_n &:= \Theta_0 \cap \{j ; n \leq j \leq N\} \\ &= \{j ; n \leq j \leq N, \mathbf{1}_{C_j} = 1\}. \end{aligned}$$

Then we have $\Theta \in \mathcal{Z}$.

- Consider some $\Theta \in \mathcal{Z}$. Define

$$C := (C_n)_{0 \leq n \leq N}$$

by

$$\begin{aligned} C_n &:= \{\omega \in \Omega ; n \in \Theta_0(\omega)\} \\ &= \{\omega \in \Omega ; n \in \Theta_n(\omega)\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N. \end{aligned}$$

Then we have $C \in \mathcal{C}$.

- Now we show that their composition in both orders is the identity.

Consider some $C \in \mathcal{C}$. Then we have for all $n \in \mathbb{N}_0$ with $n \leq N$

$$\{\omega \in \Omega ; n \in \{j ; 0 \leq j \leq N, \omega \in C_j\}\} = \{\omega \in \Omega ; \omega \in C_n\} = C_n.$$

Conversely consider some $\Theta \in \mathcal{Z}$. Then we have for all $n \in \mathbb{N}_0$ with $n < N + 1$ and $\omega \in \Omega$

$$\{j ; n \leq j \leq N, \omega \in \{\omega \in \Omega ; j \in \Theta_j(\omega)\}\} = \{j ; n \leq j \leq N, j \in \Theta_0(\omega)\} = \Theta_n(\omega).$$

Next we show that there is a mapping from \mathcal{C} to \mathcal{T} and another one from \mathcal{T} to \mathcal{C} so that their composition in both orders is the identity on \mathcal{C} or \mathcal{T} respectively.

- Consider $C \in \mathcal{C}$. Define for all $n \in \mathbb{N}_0$ with $n < N + 1$

$$\tau_n := \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p} = 1 \right\}.$$

Then we have $\tau \in \mathcal{T}$.

- Consider $\tau \in \mathcal{T}$. Define $C_N := \Omega$ and for all $n \in \mathbb{N}_0$ with $n < N$

$$C_n := \{\tau_n = n\}.$$

Then we have $C \in \mathcal{C}$.

- Now we show that their composition in both orders is the identity.

Consider some $C \in \mathcal{C}$. We have $C_N = \Omega$ and for all $n \in \mathbb{N}_0$ with $n < N$

$$\begin{aligned} & \left\{ n = \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p} = 1 \right\} \right\} \\ &= \left\{ \omega \in \Omega ; n = \inf \left\{ p ; n \leq p \leq N \wedge \omega \in C_p \right\} \right\} \\ &= C_n. \end{aligned}$$

Conversely consider $\tau \in \mathcal{T}$. Let $n \in \mathbb{N}_0$ with $n < N + 1$. Let $\omega \in \Omega$.

If $N = \infty$ we may have $\tau_n(\omega) = \infty$. Then it follows by definition of \mathcal{T} , that $\tau_p(\omega) = \infty$ for all $p \in \mathbb{N}$ with $n \leq p$, hence

$$\inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\{\tau_p = p\}} = 1 \right\} (\omega) = \infty = \tau_n(\omega).$$

Otherwise there is due to Lemma 1.12 (page 17) a minimal $m \in \mathbb{N}_0$ so that we have

$$\tau_n(\omega) = \tau_{n+k}(\omega) = \tau_{n+m}(\omega) = n + m$$

for all $k \in \mathbb{N}_0$ with $k < m$, hence

$$\begin{aligned} \tau_n(\omega) = \tau_{n+m}(\omega) = n + m &= \inf \left\{ p ; n \leq p \leq N \wedge \tau_p(\omega) = p \right\} \\ &= \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\{\tau_p = p\}} = 1 \right\} (\omega). \end{aligned}$$

2.4 Remark: Corresponding

Consider some $\tau \in \mathcal{T}$ and $C \in \mathcal{C}$ corresponding. Then we have $\tau \in \mathcal{S}(C)$.

2.5 Definition: Corresponding To

In the sequel the condition “corresponding” will be too strong several times. So we define a relaxation of that condition: For each $\tau \in \mathcal{T}$, $C \in \mathcal{C}$ and $n \in \mathbb{N}_0$ define

τ_n corresponding to C

iff

$$\tau_n = \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p} = 1 \right\}.$$

Consider some $\tau \in \mathcal{T}$ and $C \in \mathcal{C}$.

They are corresponding iff τ_n is corresponding to C for each $n \in \mathbb{N}_0$ with $n < N + 1$.

2.6 Definition and Remark: Essential

We will consider the general algorithm termination in Theorem 5.11 (page 57). We get better results if the initial parameters have a special type, which we will call essential. This is considered in Conclusion 5.13 (page 61). If the time horizon is finite and the initial parameters are of the most simple nature, there is again a small improvement. This is considered in Theorem 3.14 (page 39).

The special type of being “essential” mentioned above was first defined in [BKS08, part 3.1 ff.], therein only for elements of \mathcal{Z} and called “a-priori-set”. We will enlarge the definition to the corresponding elements of \mathcal{C} and \mathcal{T} .

- $C \in \mathcal{C}$ is called essential if there is an optimal $\sigma \in \mathcal{T}$ such, that for all $n \in \mathbb{N}_0$ with $n < N + 1$ we have $\sigma_n \in \mathcal{S}(C)$.
- $\Theta \in \mathcal{Z}$ is called essential if there is an optimal $\sigma \in \mathcal{T}$ such, that for all $n \in \mathbb{N}_0$ with $n < N + 1$ $\sigma_n \in \Theta$ a.s., meaning $P(\{\omega \in \Omega ; \sigma_n(\omega) \in \Theta(\omega)\}) = 1$.
- $\tau \in \mathcal{T}$ is called essential if there is an optimal $\sigma \in \mathcal{T}$ such, that for all $n \in \mathbb{N}_0$ with $n < N + 1$ $\{\sigma_n = n\} \subseteq \{\tau_n = n\}$.
- Consider corresponding $C \in \mathcal{C}$, $\Theta \in \mathcal{Z}$ and $\tau \in \mathcal{T}$. Then each or none of them is essential.
- $\tau \in \mathcal{T}$ is essential if there is an optimal $\sigma \in \mathcal{T}$ with $\tau \leq \sigma$ point-wise (meaning $\tau_n \leq \sigma_n$ for all $n \in \mathbb{N}_0$ with $n < N + 1$). The following Lemma 2.7 (page 21) will show this.

2.7 Lemma: Elements of \mathcal{T} compared

For all $\sigma, \tau \in \mathcal{T}$ we have the equivalence of the following assertions:

- $\{\sigma_n = n\} \subseteq \{\tau_n = n\}$ for all $n \in \mathbb{N}_0$ with $n < N + 1$,
- $\sigma_n \geq \tau_n$ for all $n \in \mathbb{N}_0$ with $n < N + 1$.

Proof:

Given the second assertion, the first follows instantaneously.

Consider the first assertion is true.

Let $n \in \mathbb{N}_0$ with $n < N + 1$.

We will prove $\sigma_n \geq \tau_n$ separately on $\{\tau_n = \infty\}$ and $\{\tau_n < \infty\}$.

First let $\omega \in \{\tau_n = \infty\}$ and assume $\omega \notin \{\sigma_n = \infty\}$. Due to Lemma 1.12 (page 17) there is $m \in \mathbb{N}_0$ with

$$\sigma_n(\omega) = \sigma_{n+m}(\omega) = n + m.$$

Since $\{\sigma_{n+m} = n + m\} \subseteq \{\tau_{n+m} = n + m\}$ it follows $\tau_{n+m}(\omega) = n + m$. Furthermore we have $\tau_{n+m}(\omega) \geq \tau_n(\omega)$ due to Lemma 1.12 (page 17). Hence $n + m = \sigma_{n+m}(\omega) \geq \tau_n(\omega) = \infty$, a contradiction. So we have $\sigma_n \geq \tau_n$ on $\{\tau_n = \infty\}$.

Due to Lemma 1.12 (page 17) we have

$$\{\tau_n < \infty\} = \bigcup_{m \in \mathbb{N}_0} \left(\{\tau_{n+m} = n + m\} \cap \bigcap_{k=n}^{n+m-1} \{\tau_k > k\} \right).$$

So it suffices to show by induction for all $m \in \mathbb{N}_0$

$$\sigma_n \geq \tau_n \quad \text{on} \quad \{\tau_{n+m} = n + m\} \cap \bigcap_{k=n}^{n+m-1} \{\tau_k > k\}.$$

For $m = 0$ the assertion is true, since $\{\sigma_n = n\} \subseteq \{\tau_n = n\}$ implies $\sigma_n \geq \tau_n$ on $\{\tau_n = n\}$.

Now consider it is true for some $m \in \mathbb{N}_0$. Then we have

$$\sigma_n = \sigma_{n+m+1} = \tau_{n+m+1} = \tau_n \text{ on } \{\tau_{n+m+1} = n + m + 1\} \cap \{\tau_{n+m} > n + m\} \cap \bigcap_{k=n}^{n+m-1} \{\tau_k > k\}.$$

Hence it is true for $m + 1$. So we proved $\sigma_n \geq \tau_n$ on $\{\tau_n < \infty\}$. \square

2.8 The Window Parameter

For each iteration of the algorithm step we fix some window parameter k . Let l be the minimum of k and the number of remaining time points ahead. We look l time points ahead in this manner: We compare the current pay-off with the expected pay-off when we do not use the next 1, 2, 3, ... l exercise possibilities and stop then according to the best stopping rule we have calculated until now. The window parameter k is the maximal number of time points we look ahead. It may be chosen differently in each step, so we therefore define a window parameter function

$$\kappa : \mathbb{N} \longrightarrow \mathbb{N} \cup \{\infty\}$$

as described at the beginning of this chapter, for example $\kappa \equiv 1$ or $\kappa \equiv N$.

Instead of fixing κ in advance, it is possible to choose $\kappa(m + 1)$ dependent of the results of the first m iterations before doing the $(m + 1)$ th iteration. We do not examine this within this thesis because the “goodness” of the choice of κ depends strongly on the individual problem and the individual technique of calculating or approximating the involved conditional expectations.

2.9 Input for FII Algorithm

The FII Algorithm has this input parameters:

- the probability space (Ω, \mathcal{A}, P) with filtration $(\mathcal{A}_n)_{0 \leq n < \infty}$ described in Setting 1.2 (page 14),
- the process $(X_n)_{0 \leq n < \infty}$ described in Setting 1.2 (page 14),
- a window parameter function κ described in Part 2.8 (page 22).
- initial parameters $C^{(0)} \in \mathcal{C}$ and $\Theta^{(0)} \in \mathcal{Z}$ corresponding, and $\tau^{(0)} \in \mathcal{T}$.¹

2.10 Remark: Input for FII Algorithm

The canonical initializing is done by setting

$$\tau_n^{(0)} \equiv n, C_n^{(0)} = \Omega \text{ and } \Theta_n^{(0)} \equiv \{i; n \leq i \leq N\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N + 1$$

and if $N = \infty$ also $C_\infty^{(0)} := \Omega$.

When using any other $C^{(0)}$ the canonical definition for $\Theta^{(0)}$ and $\tau^{(0)}$ is choosing them corresponding by the mappings in Definition 2.3 (page 19).

¹The idea of starting with some consistent family of stopping rules independently of choosing $C^{(0)}$ has been done first in [KS06, p. 34]. There it was done only for $C^{(0)} \equiv \Omega$ and finite N .

2.11 Output of FII Algorithm

In Part 2.13 (page 24) we describe how the algorithm is calculating the output. Here we show which form it has and in Remark 2.12 (page 23) we remark their meaning and sense. The final output of the FII Algorithm is the tuple

$$\left(Y^{(\infty)}, \tau^{(\infty)}, \Theta^{(\infty)}, C^{(\infty)} \right)$$

with

$$\begin{aligned} Y^{(\infty)} &= \left(Y_n^{(\infty)} \right)_{0 \leq n < N+1}, \\ \tau^{(\infty)} &= \left(\tau_n^{(\infty)} \right)_{0 \leq n < N+1}, \\ \Theta^{(\infty)} &= \left(\Theta_n^{(\infty)} \right)_{0 \leq n \leq N}, \\ C^{(\infty)} &= \left(C_n^{(\infty)} \right)_{0 \leq n \leq N}. \end{aligned}$$

This is done by constructing iterative an approximation. In the m th iteration step ($m \in \mathbb{N}_0$) the interim output of the FII Algorithm is the tuple

$$\left(Y^{(m)}, \tilde{Y}^{(m)}, \hat{Y}^{(m)}, \tau^{(m)}, \hat{\tau}^{(m)}, \Theta^{(m)}, \hat{\Theta}^{(m)}, C^{(m)}, \hat{C}^{(m)} \right)$$

with $\tau^{(m)}, \hat{\tau}^{(m)}, Y^{(m)}, \tilde{Y}^{(m)}, \hat{Y}^{(m)}$ being of the form $\tau^{(m)} = \left(\tau_n^{(m)} \right)_{0 \leq n < N+1}$, and $\Theta^{(m)}, \hat{\Theta}^{(m)}, C^{(m)}, \hat{C}^{(m)}$ of the form $\Theta^{(m)} = \left(\Theta_n^{(m)} \right)_{0 \leq n \leq N}$.

2.12 Remarks: Output of FII Algorithm

- Given later discussed conditions for every $m \in \mathbb{N}$ the variable $Y^{(m)}$ is a lower approximation of the process Y^* as shown below.
- $\tau^{(m)}$ is a lower approximation of an (early) optimal stopping family. Under certain conditions it is even a lower approximation of $\check{\tau}^*$, the first optimal stopping family (see Theorem 3.14 (page 39)).

$\tau^{(m)}, C^{(m)}$ and $\Theta^{(m)}$ are corresponding, hence each can be constructed from the other. They are just three different display forms of the same thing. Storing any of these three variables is enough to perform the next iteration step.

- $\hat{C}^{(m)}, \hat{\Theta}^{(m)}, \hat{\tau}^{(m)}$ are also corresponding and hence three different display forms of the same thing. In the finite case (for N finite) we will show in Theorem 3.8 (page 36) that each stopping rule between $\tau^{(m+1)}$ and $\hat{\tau}^{(m+1)}$ is an “improvement” of $\tau^{(m)}$. For getting an approximation of an stopping rule as small as possible we concentrate on the variables without hat.
- $\tilde{Y}^{(m)}$ and $\hat{Y}^{(m)}$ are auxiliary variables to construct the others.

2.13 Performing FII Algorithm

The calculation within the FII Algorithm is done this way:

- **Initializing Step** ($m = 0$): Set $\hat{\tau}^{(0)} := \tau^{(0)}$, $\hat{C}^{(0)} := C^{(0)}$, $\hat{\Theta}^{(0)} := \Theta^{(0)}$, and for all $n \in \mathbb{N}_0$ with $n < N + 1$ set

$$Y_n^{(0)} := \mathbb{E} \left(X_{\tau_n^{(0)}} \middle| \mathcal{A}_n \right) \text{ and } \tilde{Y}_n^{(0)} := \hat{Y}_n^{(0)} := 0.$$

- **Iteration Step** ($m > 0$): Suppose that for some $m \in \mathbb{N}_0$ the above mentioned tuple is constructed, then define for all $n \in \mathbb{N}_0$ with $n < N + 1$

$$\begin{aligned} \hat{Y}_n^{(m+1)} &:= \sup \left\{ \mathbb{E} \left(X_{\tau_p^{(m)}} \middle| \mathcal{A}_n \right) ; n + 1 \leq p < \min \{n + \kappa(m), N\} + 1 \right\} \\ &= \sup \left\{ \mathbb{E} \left(Y_p^{(m)} \middle| \mathcal{A}_n \right) ; n + 1 \leq p < \min \{n + \kappa(m), N\} + 1 \right\}, \end{aligned}$$

$$\begin{aligned} \tilde{Y}_n^{(m+1)} &:= \sup \left\{ \mathbb{E} \left(X_{\tau_p^{(m)}} \middle| \mathcal{A}_n \right) ; n \leq p < \min \{n + \kappa(m), N\} + 1 \right\} \\ &= \sup \left\{ \mathbb{E} \left(Y_p^{(m)} \middle| \mathcal{A}_n \right) ; n \leq p < \min \{n + \kappa(m), N\} + 1 \right\}, \end{aligned}$$

$$\begin{aligned} \hat{C}_n^{(m+1)} &:= \hat{C}_n^{(m)} \cap \left\{ \hat{Y}_n^{(m+1)} < X_n \right\} \\ &= \hat{C}_n^{(m)} \cap \bigcap \left\{ \left\{ \mathbb{E} \left(X_{\tau_p^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} ; n + 1 \leq p < \min \{n + \kappa(m), N\} + 1 \right\}, \end{aligned}$$

$$\begin{aligned} C_n^{(m+1)} &:= C_n^{(m)} \cap \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \\ &= C_n^{(m)} \cap \bigcap \left\{ \left\{ \mathbb{E} \left(X_{\tau_p^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} ; n \leq p < \min \{n + \kappa(m), N\} + 1 \right\}, \end{aligned}$$

$$\Theta_n^{(m+1)} := \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p^{(m+1)}} = 1 \right\} = \Theta_0^{(m+1)} \cap \{p ; n \leq p \leq N\},$$

$$\hat{\Theta}_n^{(m+1)} := \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\hat{C}_p^{(m+1)}} = 1 \right\} = \hat{\Theta}_0^{(m+1)} \cap \{p ; n \leq p \leq N\},$$

$$\begin{aligned} \tau_n^{(m+1)} &:= \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p^{(m+1)}} = 1 \right\} = \inf \Theta_0^{(m+1)} \cap \{p ; n \leq p \leq N\} \\ &= \inf \Theta_n^{(m+1)}, \end{aligned}$$

$$\begin{aligned} \hat{\tau}_n^{(m+1)} &:= \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\hat{C}_p^{(m+1)}} = 1 \right\} = \inf \hat{\Theta}_0^{(m+1)} \cap \{p ; n \leq p \leq N\} \\ &= \inf \hat{\Theta}_n^{(m+1)}, \end{aligned}$$

$$Y_n^{(m+1)} := \mathbb{E} \left(X_{\tau_n^{(m+1)}} \middle| \mathcal{A}_n \right),$$

and if $N = \infty$, then define $\hat{C}_\infty^{(m+1)} := C_\infty^{(m+1)} := \Omega$ and $\Theta_\infty^{(m+1)} := \hat{\Theta}_\infty^{(m+1)} := \{\infty\}$ as well.

- **Final step** ($m = \infty$): Suppose that for all $m \in \mathbb{N}_0$ the above mentioned tuple is constructed. Define for all $n \in \mathbb{N}_0$ with $n < N + 1$

$$C_n^{(\infty)} := \bigcap_{m \in \mathbb{N}_0} C_n^{(m)},$$

$$\Theta_n^{(\infty)} := \bigcap_{m \in \mathbb{N}_0} \Theta_n^{(m)},$$

$$\tau_n^{(\infty)} := \sup_{m \in \mathbb{N}} \tau_n^{(m)},$$

$$Y_n^{(\infty)} := \sup_{m \in \mathbb{N}_0} Y_n^{(m)}$$

and if $N = \infty$ define $C_\infty^{(\infty)} := \Omega$ and $\Theta_\infty^{(\infty)} := \{\infty\}$ as well.

2.14 Remarks: Performing FII Algorithm

- For all $m \in \mathbb{N}$ we have $C^{(m)} \in \mathcal{C}$, $\Theta^{(m)} \in \mathcal{Z}$, and $\tau^{(m)} \in \mathcal{T}$ and they are corresponding.
- For all $n \in \mathbb{N}_0$ with $n \leq N$ we have $C_n^{(\cdot)}$, $\Theta_n^{(\cdot)}$ non-increasing and $\tau_n^{(\cdot)}$ non-decreasing.
- For all $m \in \mathbb{N}$ we have $\widehat{C}^{(m)} \in \mathcal{C}$, $\widehat{\Theta}^{(m)} \in \mathcal{Z}$, and $\widehat{\tau}^{(m)} \in \mathcal{T}$ and they are corresponding.
- For all $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$\tau_n^{(\infty)} = \inf \left\{ p ; n \leq p < N + 1 \wedge \mathbf{1}_{C_p^{(\infty)}} = 1 \right\} \in \mathcal{S}(C^{(\infty)}).$$

2.15 Comparison with the Literature

There are some algorithms being similar to the one above suggested in the literature before. The variables therein examined in each step belong to those described in Part 2.11 (page 23).

The variables \widehat{Y} , \widehat{C} , $\widehat{\Theta}$, and $\widehat{\tau}$ only appear explicitly in [BS06, part 3.2 et seqq.]. The algorithm suggested in [Irl80, p. 180], [Irl06, pp. 102-103] and [Irl09] omits mentioning κ , Θ , Y and \widehat{Y} explicitly. The algorithm suggested in [KS06, pp. 31-35,42] omits mentioning C explicitly, and a Θ very similar to ours appears therein in part six et seqq. The algorithm in [BKS08] omits mentioning κ and instead of examining Θ it just has a look at Θ_0 , named A therein. Irle has a look at the infinite time horizon in his above mentioned papers, while the others focus on finite time horizon.

In the following lemma we collect elementary facts which we will use several times in the subsequent lemmas and theorems. Some of them are used implicitly here and there in [BS06, part 3.2, lemma 3.3].

2.16 Lemma: Elementary Equations and Inequalities

For all $n \in \mathbb{N}_0$ with $n < N + 1$ and for all $m \in \mathbb{N}_0$ we have

$$\begin{aligned} \widetilde{Y}_n^{(m+1)} &= \max \left\{ Y_n^{(m)}, \widehat{Y}_n^{(m+1)} \right\} \\ &= \max \left\{ \mathbb{E} \left(X_{\tau_n^{(m)}} \middle| \mathcal{A}_n \right), \widehat{Y}_n^{(m+1)} \right\} \\ (2.16.1) \quad &= \mathbf{1}_{\{\tau_n^{(m)} > n\}} \widehat{Y}_n^{(m+1)} + \mathbf{1}_{\{\tau_n^{(m)} = n\}} \max \left\{ \widehat{Y}_n^{(m+1)}, X_n \right\} \end{aligned}$$

$$(2.16.2) \quad \geq \widehat{Y}_n^{(m+1)},$$

$$(2.16.3) \quad \left\{ \widetilde{Y}_n^{(m+1)} \leq X_n \right\} = \left\{ \widehat{Y}_n^{(m+1)} \leq X_n \right\},$$

$$\begin{aligned} C_n^{(m+1)} &= C_n^{(m)} \cap \left\{ \widehat{Y}_n^{(m+1)} \leq X_n \right\} \\ (2.16.4) \quad &= C_n^{(m)} \cap \bigcap \left\{ \left\{ \mathbb{E} \left(X_{\tau_p^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} ; n + 1 \leq p < \min \{ n + \kappa(m), N \} + 1 \right\}, \end{aligned}$$

$$(2.16.5) \quad \mathbb{E} \left(\widetilde{Y}_n^{(m)} \middle| \mathcal{A}_{n-1} \right) \geq \widehat{Y}_{n-1}^{(m)}.$$

Proof:

- (2.16.1) follows due to the consistency of $\tau^{(m)}$ defined in Definition 1.10 (page 16),

$$\left\{ \tau_n^{(m)} > n \right\} \subseteq \left\{ \tau_n^{(m)} = \tau_{n+1}^{(m)} \right\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N,$$

by the definitions above and properties of conditional expectations.

- (2.16.3) follows by (2.16.1).
- (2.16.4) is just a combination of the underlying definitions and (2.16.3).
- For (2.16.5) have a look at the definitions. □

In Definition 2.18 (page 26) we will define “improver”. To justify the definition we first state a lemma to show the existence of improvers. The idea of this definition comes from [BS06, definition 3.5].

2.17 Lemma: Existence of Improvers

For all $m \in \mathbb{N}_0$ and for all $n \in \mathbb{N}_0$ with $n < N + 1$ we have

$$\begin{aligned} \widehat{C}_n^{(m)} &\subseteq C_n^{(m)}, \\ \widehat{\Theta}_n^{(m)} &\subseteq \Theta_n^{(m)}, \\ \tau_n^{(m)} &\leq \widehat{\tau}_n^{(m)}. \end{aligned}$$

Proof: Assume $n \in \mathbb{N}_0$ with $n < N + 1$.

Proof of the first statement by induction over m :

It is true for $m = 0$ by definition.

Suppose it is true for some $m \in \mathbb{N}_0$. We have

$$\left\{ \widehat{Y}_n^{(m+1)} < X_n \right\} \subseteq \left\{ \widehat{Y}_n^{(m+1)} \leq X_n \right\} \stackrel{2.16.3}{=} \left\{ \widetilde{Y}_n^{(m+1)} \leq X_n \right\}.$$

Hence it follows

$$C_n^{(m+1)} \stackrel{\text{def.}}{=} \left\{ \widetilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(m)} \supseteq \left\{ \widehat{Y}_n^{(m+1)} < X_n \right\} \cap \widehat{C}_n^{(m)} \stackrel{\text{def.}}{=} \widehat{C}_n^{(m+1)}.$$

The other statements are true since for all $m \in \mathbb{N}$ we have

$$\mathbf{1}_{\widehat{C}_p^{(m+1)}} = 1 \implies \mathbf{1}_{C_p^{(m+1)}} = 1 \text{ for all } n \leq p \leq N,$$

hence

$$\Theta_n^{(m)} \stackrel{\text{def.}}{=} \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p^{(m+1)}} = 1 \right\} \supseteq \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\widehat{C}_p^{(m+1)}} = 1 \right\} \stackrel{\text{def.}}{=} \widehat{\Theta}_n^{(m)}$$

and

$$\tau_n^{(m)} \stackrel{\text{def.}}{=} \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p^{(m)}} = 1 \right\} \leq \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{\widehat{C}_p^{(m)}} = 1 \right\} \stackrel{\text{def.}}{=} \widehat{\tau}_n^{(m)}. \quad \square$$

2.18 Definition: Improver

Consider $m \in \mathbb{N}$ and $\sigma \in \mathcal{T}$.

σ is called **improver of $\tau^{(m)}$** iff

$$\tau_n^{(m+1)} \leq \sigma_n \leq \widehat{\tau}_n^{(m+1)} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N + 1.$$

Accordingly the term improver is defined for elements of \mathcal{C} and \mathcal{Z} .

Remark: By Lemma 2.17 (page 26) $\tau^{(m+1)}$ and $\widehat{\tau}^{(m+1)}$ are improvers of $\tau^{(m)}$.

2.19 Setting of [Irl80] included

If $N = \infty$ and $\kappa \equiv 1$, then for all $m \in \mathbb{N}_0$ and for all $n \in \mathbb{N}_0$ we have

$$C_n^{(m+1)} = \left\{ \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} \cap C_n^{(m)}.$$

Comment: Hence the algorithm of [Irl80] is included in the above described wider setting.

Proof: Consider $N = \infty$, $\kappa \equiv 1$, $m \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$.
Then we have

$$\begin{aligned} \hat{Y}_n^{(m+1)} &= \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right), \\ \tilde{Y}_n^{(m+1)} &= \sup \left\{ \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right), \mathbb{E} \left(X_{\tau_n^{(m)}} \middle| \mathcal{A}_n \right) \right\}. \end{aligned}$$

By Lemma 2.16 (page 25) we have

$$\left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} = \left\{ \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\},$$

hence

$$C_n^{(m+1)} = \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(m)} = \left\{ \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} \cap C_n^{(m)}. \quad \square$$

2.20 Remark: Infinite Case Includes Finite Case

The case $N < \infty$ is of course part of the infinite case by setting $X_{N+n} := X_N$ for all $n \in \mathbb{N}$. Obviously it follows $X_\infty = X_N$.

3 Finite Time Horizon

In this chapter as well as in Section 5 (page 49) we examine further the behaviour of the output variables of our algorithm in dependence of the input variables. In particular we examine under which conditions the terminal output of FII algorithm is optimal in some manner. In this chapter we concentrate on the case of finite time horizon, while we will examine the general case in Section 5 (page 49).

In the first lemmas and theorems we state some fundamental results, we use later. In Remark 3.9 (page 37) and Theorem 3.10 (page 37) we show that $\tau^{(m)}$ is an increasing sequence which is (under weak conditions) lower than each optimal family of stopping times (of a certain class). Theorem 3.14 (page 39) and Theorem 3.18 (page 42) examine the algorithm termination. In Remark 3.15 (page 41) we have a look at an idea of [BKS08, Proposition 3.4] for making generalisations of these results, but we can show by Example 3.16 (page 41) and Example 3.17 (page 41), that this idea is not viable. In Theorem 3.21 (page 44) we show that with increasing κ the number of iterations decreases.

For the whole chapter consider finite time horizon, hence finite N . This finite case has been considered in [KS06] intensively, but always with $C^{(0)} \equiv \Omega$, since this parameter was not used explicitly. In this chapter we will remove this restriction. We will show that relaxing it to $\tau^{(0)}$ and $C^{(0)}$ corresponding is generally possible; for several assertions even this can often be relaxed to the assumption, that $\tau_n^{(0)}$ is corresponding to $C^{(0)}$ for several n .

In the following lemma we collect elementary facts we will use several times in the subsequent lemmas and theorems. Some of them are used implicitly here and there in [BS06, part 3.2] in the more simple form of $C^{(0)} \equiv \Omega$.

3.1 Theorem: Equations and Inequalities

We have for all $n \in \mathbb{N}_0$ with $n \leq N$

$$(3.1.1) \quad X_n = Y_n^{(0)} \text{ on } \left\{ \tau_n^{(0)} = n \right\},$$

$$(3.1.2) \quad \tilde{Y}_n^{(1)} = \max \left\{ \hat{Y}_n^{(1)}, X_n \right\} = \max \left\{ \hat{Y}_n^{(1)}, Y_n^{(0)} \right\} \text{ on } \left\{ \tau_n^{(0)} = n \right\},$$

$$(3.1.3) \quad X_n < \tilde{Y}_n^{(1)} = \hat{Y}_n^{(1)} \text{ on } C_n^{(0)} \setminus C_n^{(1)},$$

$$(3.1.4) \quad \hat{Y}_n^{(1)} = \tilde{Y}_n^{(1)} \text{ on } \left\{ \tau_n^{(0)} > n \right\} \cup C_n^{(0)} \setminus C_n^{(1)},$$

$$(3.1.5) \quad X_n = Y_n^{(0)} = \tilde{Y}_n^{(1)} \geq \hat{Y}_n^{(1)} \text{ on } \left\{ \tau_n^{(0)} = n \right\} \cap C_n^{(1)}.$$

For all $m \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ with $n \leq N$ with $\tau_n^{(m)}$ corresponding to $C^{(m)}$ we have

$$(3.1.6) \quad X_n = Y_n^{(m)} \text{ on } C_n^{(m)},$$

$$(3.1.7) \quad \tilde{Y}_n^{(m+1)} = \max \left\{ \hat{Y}_n^{(m+1)}, X_n \right\} = \max \left\{ \hat{Y}_n^{(m+1)}, Y_n^{(m)} \right\} \text{ on } C_n^{(m)},$$

$$(3.1.8) \quad X_n = Y_n^{(m)} < \tilde{Y}_n^{(m+1)} = \hat{Y}_n^{(m+1)} \text{ on } C_n^{(m)} \setminus C_n^{(m+1)},$$

$$(3.1.9) \quad \hat{Y}_n^{(m+1)} = \tilde{Y}_n^{(m+1)} \text{ on } \left(C_n^{(m+1)} \right)^c,$$

$$(3.1.10) \quad X_n = Y_n^{(m)} = \tilde{Y}_n^{(m+1)} \geq \hat{Y}_n^{(m+1)} \text{ on } C_n^{(m+1)}.$$

Proof: Consider $n \in \mathbb{N}_0$ with $n \leq N$, $m \in \mathbb{N}_0$. We will prove

$$(3.1.11) \quad X_n = Y_n^{(m)} \text{ on } \left\{ \tau_n^{(m)} = n \right\},$$

$$(3.1.12) \quad \tilde{Y}_n^{(m+1)} = \max \left\{ \hat{Y}_n^{(m+1)}, X_n \right\} \text{ on } \left\{ \tau_n^{(m)} = n \right\},$$

$$(3.1.13) \quad X_n < \tilde{Y}_n^{(m+1)} = \hat{Y}_n^{(m+1)} \text{ on } C_n^{(m)} \setminus C_n^{(m+1)},$$

$$(3.1.14) \quad \tilde{Y}_n^{(m+1)} = \hat{Y}_n^{(m+1)} \text{ on } \left\{ \tau_n^{(m)} > n \right\} \cup C_n^{(m)} \setminus C_n^{(m+1)},$$

$$(3.1.15) \quad X_n = Y_n^{(m)} = \tilde{Y}_n^{(m+1)} \geq \hat{Y}_n^{(m+1)} \text{ on } \left\{ \tau_n^{(m)} = n \right\} \cap C_n^{(m+1)}.$$

Before going into the details of the proof we shortly show the consequences:

The equations (3.1.1) to (3.1.5) are evident.

If $\tau_n^{(m)}$ is corresponding to $C^{(m)}$, we have

$$C_n^{(m+1)} \subset C_n^{(m)}, \left\{ \tau_n^{(m)} = n \right\} = C_n^{(m)} \text{ and } \left\{ \tau_n^{(m)} > n \right\} = \left(C_n^{(m)} \right)^c,$$

whereby equations (3.1.6), (3.1.7) and (3.1.8) instantly follow, and we have (3.1.9) due to

$$\left\{ \tau_n^{(m)} > n \right\} \cup C_n^{(m)} \setminus C_n^{(m+1)} = \left(C_n^{(m)} \right)^c \cup C_n^{(m)} \setminus C_n^{(m+1)} = \left(C_n^{(m+1)} \right)^c$$

and (3.1.10) due to

$$\left\{ \tau_n^{(m)} = n \right\} \cap C_n^{(m+1)} = C_n^{(m)} \cap C_n^{(m+1)} = C_n^{(m+1)}.$$

We have by the definitions in Algorithm 2.13 (page 24)

$$(3.1.16) \quad \tau_n^{(m+1)} \stackrel{\text{def}}{=} \inf \left\{ p; n \leq p < N + 1, \mathbf{1}_{C_p^{(m+1)}} = 1 \right\},$$

$$(3.1.17) \quad C_n^{(m+1)} \stackrel{\text{def}}{=} \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(m)},$$

$$(3.1.18) \quad Y_n^{(m)} \stackrel{\text{def}}{=} \mathbb{E} \left(X_{\tau_n^{(m)}} \mid \mathcal{A}_n \right).$$

Proof of (3.1.11):

By properties of conditional expectations we have

$$\left\{ \tau_n^{(m)} = n \right\} \subseteq \left\{ Y_n^{(m)} = X_n \right\}.$$

Proof of (3.1.12):

This follows by Lemma 2.16 (page 25).

Proof of (3.1.13):

We have

$$\begin{aligned} & \left\{ \tau_n^{(m+1)} > n \right\} \cap C_n^{(m)} \\ &= \left(C_n^{(m+1)} \right)^c \cap C_n^{(m)} && \text{by (3.1.16)} \\ &= \left(\left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(m)} \right)^c \cap C_n^{(m)} && \text{by (3.1.17)} \\ &= \left(\left\{ \tilde{Y}_n^{(m+1)} > X_n \right\} \cap C_n^{(m)} \right) \cup \left(\left(C_n^{(m)} \right)^c \cap C_n^{(m)} \right) \\ &= \left\{ \tilde{Y}_n^{(m+1)} > X_n \right\} \cap C_n^{(m)} \cup \emptyset \\ &\subseteq \left\{ \tilde{Y}_n^{(m+1)} > X_n \right\}. \end{aligned}$$

So we have

$$X_n < \tilde{Y}_n^{(m+1)} \stackrel{2.16}{\leq} \max \left\{ X_n, \hat{Y}_n^{(m+1)} \right\} \text{ on } \left(C_n^{(m+1)} \right)^c \cap C_n^{(m)}.$$

and thus

$$X_n < \tilde{Y}_n^{(m+1)} = \hat{Y}_n^{(m+1)} \text{ on } \left(C_n^{(m+1)} \right)^c \cap C_n^{(m)} = C_n^{(m)} \setminus C_n^{(m+1)}.$$

Proof of (3.1.14):

By Lemma 2.16 (page 25) we have

$$(3.1.19) \quad \tilde{Y}_n^{(m+1)} = \hat{Y}_n^{(m+1)} \text{ on } \left\{ \tau_n^{(m)} > n \right\}.$$

(3.1.14) is a combination of this fact with (3.1.13).

Proof of (3.1.15):

We have

$$\left\{ \tau_n^{(m+1)} = n \right\} = C_n^{(m+1)} \quad \text{by (3.1.16)}$$

$$= \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(m)} \quad \text{by (3.1.17)}$$

$$(3.1.20) \quad \stackrel{\text{triv.}}{\subseteq} \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\}.$$

By definition in Algorithm 2.13 (page 24) we have

$$(3.1.21) \quad \tilde{Y}_n^{(m+1)} \geq Y_n^{(m)}.$$

So we have

$$X_n \stackrel{(3.1.20)}{\geq} \tilde{Y}_n^{(m+1)} \stackrel{(3.1.21)}{\geq} Y_n^{(m)} \stackrel{(3.1.11)}{=} X_n \text{ on } C_n^{(m+1)} \cap \left\{ \tau_n^{(m)} = n \right\}.$$

Equality everywhere follows and thus

$$\max \left\{ \hat{Y}_n^{(m+1)}, X_n \right\} \stackrel{(3.1.12)}{=} \tilde{Y}_n^{(m+1)} = Y_n^{(m)} = X_n \text{ on } C_n^{(m+1)} \cap \left\{ \tau_n^{(m)} = n \right\}.$$

So we have

$$\hat{Y}_n^{(m+1)} \leq \tilde{Y}_n^{(m+1)} = Y_n^{(m)} = X_n \text{ on } C_n^{(m+1)} \cap \left\{ \tau_n^{(m)} = n \right\}. \quad \square$$

In the sequel we will show that our algorithm gives improved results in every step. We start with a look at the first iteration step and its specialties in the next lemma, Lemma 3.2 (page 30), and then proceed with the other steps in Theorem 3.3 (page 31) and therein also give an improvement for the first step under some conditions.

3.2 Lemma

For all $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$\tilde{Y}_n^{(1)} \leq Y_n^{(1)} \quad \text{on} \quad \left\{ \tau_n^{(0)} > n \right\} \cup C_n^{(0)}$$

and

$$X_n \leq Y_n^{(1)} \quad \text{on} \quad C_n^{(0)}.$$

Proof: Consider $n \in \mathbb{N}_0$ with $n \leq N$.

We have

$$\begin{aligned}\tilde{Y}_n^{(1)} &\stackrel{\text{def.}}{\leq} X_n = Y_n^{(1)} \text{ on } C_n^{(1)} = \{\tau_n^{(1)} = n\}, \\ X_n &\stackrel{(3.1.3)}{<} \tilde{Y}_n^{(1)} \stackrel{(3.1.3)}{=} \hat{Y}_n^{(1)} \text{ on } C_n^{(0)} \setminus C_n^{(1)}, \\ \hat{Y}_n^{(1)} &\stackrel{(3.1.4)}{=} \tilde{Y}_n^{(1)} \text{ on } \{\tau_n^{(0)} > n\} \cup C_n^{(0)} \setminus C_n^{(1)}.\end{aligned}$$

Consider some $n \in \mathbb{N}_0$ with $n < N - 1$ so that $\tilde{Y}_{n+1}^{(1)} \leq Y_{n+1}^{(1)}$. Then we have

$$\hat{Y}_n^{(1)} \stackrel{(3.3.2)}{\leq} Y_n^{(1)} \text{ on } (C_n^{(1)})^c,$$

$$\left(\{\tau_n^{(0)} > n\} \cup C_n^{(0)} \setminus C_n^{(1)}\right) \cap (C_n^{(1)})^c = \left(\{\tau_n^{(0)} > n\} \cup C_n^{(0)}\right) \cap (C_n^{(1)})^c$$

and thus follows

$$\tilde{Y}_n^{(1)} \stackrel{3.1}{=} \hat{Y}_n^{(1)} \stackrel{(3.3.2)}{\leq} Y_n^{(1)} \text{ on } \left(\{\tau_n^{(0)} > n\} \cup C_n^{(0)}\right) \cap (C_n^{(1)})^c = \{\tau_n^{(0)} > n\} \cup C_n^{(0)}. \quad \square$$

Lemma 3.2 (page 30) above can be generalized: We state and prove here a theorem and a conclusion for arbitrary $C^{(0)}$, which has been proved for $C^{(0)} \equiv \Omega$ in [KS06, pp. 32-35]. We will generalize this theorem again in Theorem 3.8 (page 36) below.

3.3 Theorem

- For all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$\tilde{Y}_n^{(m+1)} \leq Y_n^{(m+1)}.$$

- For all $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

the inequality above is true for $m = 0$.

Proof via Backward Induction over n :

We will proof both points at once. So consider $m \in \mathbb{N}_0$.

We have

$$\tilde{Y}_N^{(m+1)} = \mathbb{E}\left(X_{\tau_N^{(m)}} \middle| \mathcal{A}_N\right) \quad \text{and} \quad Y_N^{(m+1)} = \mathbb{E}\left(X_{\tau_N^{(m+1)}} \middle| \mathcal{A}_N\right) \quad \text{and} \quad \tau_N^{(m)} \equiv N \equiv \tau_N^{(m+1)}.$$

Thus it follows

$$\tilde{Y}_N^{(m+1)} = \mathbb{E}\left(X_{\tau_N^{(m)}} \middle| \mathcal{A}_N\right) = \mathbb{E}(X_N | \mathcal{A}_N) = X_N = \mathbb{E}(X_N | \mathcal{A}_N) = \mathbb{E}\left(X_{\tau_N^{(m+1)}} \middle| \mathcal{A}_N\right) = Y_N^{(m+1)}.$$

Consider some $n \in \mathbb{N}_0$ with $n < N - 1$ so that $\tilde{Y}_{n+1}^{(m+1)} \leq Y_{n+1}^{(m+1)}$.

We have

$$\tilde{Y}_n^{(m+1)} \stackrel{\text{def.}}{\leq} X_n = Y_n^{(m+1)} \text{ on } C_n^{(m+1)} = \{\tau_n^{(m+1)} = n\}.$$

Due to

$$(3.3.1) \quad (C_n^{(m+1)})^c = \{\tau_n^{(m+1)} > n\} \subseteq \{\tau_n^{(m)} = \tau_{n+1}^{(m+1)}\},$$

it follows on $(C_n^{(m+1)})^c$:

$$\begin{aligned}
(3.3.2) \quad Y_n^{(m+1)} &= \mathbb{E} \left(X_{\tau_n^{(m+1)}} \middle| \mathcal{A}_n \right) && \text{by def.} \\
&= \mathbb{E} \left(X_{\tau_{n+1}^{(m+1)}} \middle| \mathcal{A}_n \right) && \text{by (3.3.1)} \\
&= \mathbb{E} \left(\mathbb{E} \left(X_{\tau_{n+1}^{(m+1)}} \middle| \mathcal{A}_{n+1} \right) \middle| \mathcal{A}_n \right) \\
&= \mathbb{E} \left(Y_{n+1}^{(m+1)} \middle| \mathcal{A}_n \right) && \text{by def.} \\
&\geq \mathbb{E} \left(\tilde{Y}_{n+1}^{(m+1)} \middle| \mathcal{A}_n \right) && \text{by ind. ass.} \\
&\geq \hat{Y}_n^{(m+1)} && \text{by 2.16.}
\end{aligned}$$

If $m > 0$ or $m = 0$ and $\tau_n^{(0)}$ is corresponding to $C^{(0)}$ we have

$$\tilde{Y}_n^{(m+1)} \stackrel{(3.1.9)}{=} \hat{Y}_n^{(m+1)} \stackrel{(3.3.2)}{\leq} Y_n^{(m+1)} \text{ on } (C_n^{(m+1)})^c. \quad \square$$

The following conclusion has been proved in [KS06, pp. 32-35] for $C^{(0)} \equiv \Omega$. We will expand it to the case of general $C^{(0)}$.

3.4 Conclusion

For all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$\begin{aligned}
Y_n^{(m)} &\leq \tilde{Y}_n^{(m+1)} \leq Y_n^{(m+1)} \leq Y_n^*, \\
\tilde{Y}_n^{(m+1)} &\leq \tilde{Y}_n^{(m+2)}, \\
\hat{Y}_n^{(m+1)} &\leq \hat{Y}_n^{(m+2)}.
\end{aligned}$$

For all $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

the inequalities above are true for $m = 0$.

Proof: Consider $n \in \mathbb{N}_0$ with $n \leq N$ and $m \in \mathbb{N}_0$. We have by definition

$$Y_n^{(m)} \leq \tilde{Y}_n^{(m+1)} \text{ and } Y_n^{(m+1)} \leq Y_n^*.$$

The proof of

$$\tilde{Y}_n^{(m+1)} \leq Y_n^{(m+1)}$$

is done for $m > 0$ as well as for $m = 0$ under the mentioned condition in Theorem 3.3 (page 31).

The next inequality follows by repeating the arguments of the first set of inequalities.

By the first set of inequalities we have $Y_p^{(m)} \leq Y_p^{(m+1)}$ for all $p \in \mathbb{N}_0$ with $n+1 \leq p \leq N$. Hence the last line follows by the definitions of \hat{Y} in Algorithm 2.13 (page 24). \square

In the following example we will show that Lemma 3.2 (page 30) cannot be expanded to a bigger subset of Ω , and that the condition of correspondency in Theorem 3.3 (page 31) and in Formula (3.1.6) (page 28) cannot be dropped. It will also be helpful for checking that conditions given in the following theorems are sharp.

3.5 Example

The following example will show the possibility of

$$Y_n^{(1)} = \widehat{Y}_n^{(1)} < X_n = \widetilde{Y}_n^{(1)} \text{ on } \left\{ \tau_n^{(0)} = n \right\} \cap \left(C_n^{(0)} \right)^c = \left\{ \tau_n^{(0)} = n \right\} \cap \left(C_n^{(1)} \right)^c.$$

Example: Consider $N \in \mathbb{N}_0 \cup \{\infty\}$ and $n \in \mathbb{N}_0$ with $n < N$ and a setting with

$$X_n, X_{n+1} \text{ deterministic with } X_n > X_{n+1},$$

$$\begin{aligned} \kappa(0) &= 1, \\ \tau_{n+1}^{(0)} &\equiv n + 1, \\ C_n^{(0)} &\not\subseteq \left\{ \tau_n^{(0)} = n \right\}, \end{aligned}$$

then we have

$$\begin{aligned} \widehat{Y}_n^{(1)} &\stackrel{\text{def}}{=} \sup \left\{ \mathbb{E} \left(X_{\tau_p^{(0)}} \middle| \mathcal{A}_n \right) ; n + 1 \leq p < \min \{ n + \kappa(m), N \} + 1 \right\} \\ &= \sup \left\{ \mathbb{E} \left(X_{\tau_p^{(0)}} \middle| \mathcal{A}_n \right) ; n + 1 \leq p < n + 2 \right\} \\ &= \mathbb{E} \left(X_{\tau_{n+1}^{(0)}} \middle| \mathcal{A}_n \right) \\ &\stackrel{\text{def}}{=} \mathbb{E} \left(X_{n+1} \middle| \mathcal{A}_n \right) \\ &= X_{n+1} \\ &< X_n \end{aligned}$$

and it follows

$$\widehat{Y}_n^{(1)} < X_n = \max \left\{ X_n, \widehat{Y}_n^{(1)} \right\} = \widetilde{Y}_n^{(1)} \text{ on } \left(\left\{ \tau_n^{(0)} > n \right\} \cup C_n^{(0)} \right)^c = \left\{ \tau_n^{(0)} = n \right\} \cap \left(C_n^{(0)} \right)^c \neq \emptyset.$$

Assume additionally

$$\begin{aligned} n + 1 &< N, \\ X_{n+2} &= X_{n+1}, \\ \tau_{n+2}^{(0)} &\equiv n + 2, \\ C_n^{(0)} &\not\subseteq C_{n+1}^{(0)} = \Omega. \end{aligned}$$

Then we have

$$\widehat{Y}_{n+1}^{(1)} = X_{n+2} = X_{n+1}$$

and

$$\begin{aligned} \widetilde{Y}_n^{(1)} &\stackrel{\text{def}}{=} \sup \left\{ \mathbb{E} \left(X_{\tau_p^{(0)}} \middle| \mathcal{A}_n \right) ; n \leq p < \min \{ n + \kappa(0), N \} + 1 \right\} \\ &\stackrel{\text{def}}{=} \max \left\{ \mathbb{E} \left(X_{\tau_p^{(0)}} \middle| \mathcal{A}_n \right) ; n \leq p < n + 2 \right\} \\ &= \max \left\{ \mathbb{E} \left(X_{\tau_n^{(0)}} \middle| \mathcal{A}_n \right), \mathbb{E} \left(X_{\tau_{n+1}^{(0)}} \middle| \mathcal{A}_n \right) \right\} \\ &= \max \left\{ X_n \mathbb{1}_{\left\{ \tau_n^{(0)} = n \right\}} + X_{n+1} \mathbb{1}_{\left\{ \tau_n^{(0)} = n+1 \right\}}, X_{n+1} \right\} \\ &= X_n \mathbb{1}_{\left\{ \tau_n^{(0)} = n \right\}} + X_{n+1} \mathbb{1}_{\left\{ \tau_n^{(0)} = n+1 \right\}}, \end{aligned}$$

thus

$$\widetilde{Y}_n^{(1)} = X_n \text{ on } \left\{ \tau_n^{(0)} = n \right\}.$$

We have

$$C_n^{(1)} = \left\{ \tilde{Y}_n^{(1)} \leq X_n \right\} \cap C_n^{(0)} = \Omega \cap C_n^{(0)} = C_n^{(0)}$$

and thus

$$\begin{aligned} C_{n+1}^{(1)} &= \left\{ \tilde{Y}_{n+1}^{(1)} \leq X_n \right\} \cap C_{n+1}^{(0)} \\ &= \left\{ \tilde{Y}_{n+1}^{(1)} \leq X_n \right\} \\ &\stackrel{2.16}{=} \left\{ \hat{Y}_{n+1}^{(1)} \leq X_n \right\} \\ &= \left\{ X_{n+1} \leq X_n \right\} \\ &= \Omega \end{aligned}$$

and

$$\left\{ \tau_n^{(1)} = n + 1 \right\} = C_{n+1}^{(1)} \setminus C_n^{(1)} = \left(C_n^{(1)} \right)^c = \left(C_n^{(0)} \right)^c.$$

Due to

$$X_{n+1} = Y_n^{(1)} \text{ on } \left\{ \tau_n^{(1)} = n + 1 \right\}$$

it follows that

$$\tilde{Y}_n^{(1)} = X_n > X_{n+1} = Y_n^{(1)} \text{ on } \left\{ \tau_n^{(0)} = n \right\} \cap \left(C_n^{(0)} \right)^c. \quad \square$$

3.6 Lemma

Fix some $n \in \mathbb{N}_0$.

For all $p \in \mathbb{N}$ and $m \in \mathbb{N}_0$ with $p \leq m$ we have

$$C_n^{(m+1)} = \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(p)},$$

$$\hat{C}_n^{(m+1)} = \left\{ \hat{Y}_n^{(m+1)} < X_n \right\} \cap \hat{C}_n^{(p)}.$$

If

$$\tau_k^{(0)} \text{ is corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N,$$

we have for all $m \in \mathbb{N}_0$

$$C_n^{(m+1)} = \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(0)},$$

$$\hat{C}_n^{(m+1)} = \left\{ \hat{Y}_n^{(m+1)} < X_n \right\} \cap C_n^{(0)}.$$

Proof by Induction over m :

Fix some $p \in \mathbb{N}_0$.

The hypothesis is trivially true for $m = p$ by definition.

Assume it is true for some $m \in \mathbb{N}_0$ with $p \leq m$. We have by Conclusion 3.4 (page 32)

$$\tilde{Y}_n^{(m+1)} \leq \tilde{Y}_n^{(m+2)},$$

$$\hat{Y}_n^{(m+1)} \leq \hat{Y}_n^{(m+2)}.$$

Thus

$$\left\{ \tilde{Y}_n^{(m+2)} \leq X_n \right\} \subseteq \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\},$$

$$\left\{ \hat{Y}_n^{(m+2)} < X_n \right\} \subseteq \left\{ \hat{Y}_n^{(m+1)} < X_n \right\}.$$

So we have

$$\begin{aligned}
C_n^{(m+2)} &\stackrel{\text{def.}}{=} \left\{ \tilde{Y}_n^{(m+2)} \leq X_n \right\} \cap C_n^{(m+1)} \\
&= \left\{ \tilde{Y}_n^{(m+2)} \leq X_n \right\} \cap \left\{ \tilde{Y}_n^{(m+1)} \leq X_n \right\} \cap C_n^{(p)} \\
&= \left\{ \tilde{Y}_n^{(m+2)} \leq X_n \right\} \cap C_n^{(p)}
\end{aligned}$$

and

$$\begin{aligned}
\hat{C}_n^{(m+2)} &\stackrel{\text{def.}}{=} \left\{ \hat{Y}_n^{(m+2)} < X_n \right\} \cap \hat{C}_n^{(m+1)} \\
&= \left\{ \hat{Y}_n^{(m+2)} < X_n \right\} \cap \left\{ \hat{Y}_n^{(m+1)} < X_n \right\} \cap \hat{C}_n^{(p)} \\
&= \left\{ \hat{Y}_n^{(m+2)} < X_n \right\} \cap \hat{C}_n^{(p)}.
\end{aligned}$$

Due to Algorithm 2.13 (page 24) we have $\hat{C}_n^{(0)} = C_n^{(0)}$. □

The following theorem has been proved for the special case of

$$C^{(0)} \equiv \Omega \quad \text{and} \quad \tau^{(0)} \text{ and } C^{(0)} \text{ corresponding}$$

in [KS06, pp. 32-35]. Here we consider arbitrary $C^{(0)}$ and $\tau^{(0)}$.

3.7 Theorem

For all $n \in \mathbb{N}_0$ with $n \leq N$ and $m \in \mathbb{N}$ we have

$$X_n \leq \tilde{Y}_n^{(m+1)} \text{ on } C_n^{(0)}.$$

For all $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

the set of inequalities above is true for $m = 0$.

Proof: Let $m \in \mathbb{N}$. Then we have

$$\begin{aligned}
X_n &\stackrel{(3.1.6)}{=} Y_n^{(1)} \text{ on } C_n^{(1)}, \\
X_n &\stackrel{(3.1.3)}{<} \tilde{Y}_n^{(1)} \text{ on } C_n^{(0)} \setminus C_n^{(1)}, \\
Y_n^{(1)} &\stackrel{3.4}{\leq} \tilde{Y}_n^{(2)}, \\
\tilde{Y}_n^{(1)} &\stackrel{3.4}{\leq} \tilde{Y}_n^{(2)};
\end{aligned}$$

hence

$$X_n \leq \tilde{Y}_n^{(2)} \stackrel{3.3}{\leq} Y_n^{(2)} \stackrel{3.4}{\leq} \tilde{Y}_n^{(m+1)} \text{ for all } m \in \mathbb{N}.$$

Under the mentioned additional assumption we have

$$X_n \stackrel{(3.1.12)}{\leq} \tilde{Y}_n^{(1)} \text{ on } \left\{ \tau_n^{(0)} = n \right\} = C_n^{(0)}. \quad \square$$

In Theorem 3.3 (page 31) and Theorem 3.7 (page 35) we had a look at $\tau^{(m+1)}$, a special improver of $\tau^{(m)}$ (for each $m \in \mathbb{N}_0$). Now we will have a look at an arbitrary improver and state the same result. This is based on an idea in [BS06, Part 3.2, Theorem 3.4 and Remark 3.4].

3.8 Theorem

Consider $m \in \mathbb{N}_0$ and $\sigma \in \mathcal{T}$ being an improver of $\tau^{(m)}$.

For all $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

we have

$$\tilde{Y}_n^{(m+1)} \leq \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n)$$

and

$$X_n \leq \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \text{ on } C_n^{(0)}.$$

Proof of the first statement via backward induction over n :

We have by definition

$$\begin{aligned} \tilde{Y}_N^{(m+1)} &= \mathbb{E}\left(X_{\tau_N^{(m+1)}} \middle| \mathcal{A}_N\right), \\ \tau_N^{(m+1)} &\equiv N \equiv \sigma_N. \end{aligned}$$

So we have

$$\tilde{Y}_N^{(m+1)} = X_N = \mathbb{E}(X_{\sigma_N} | \mathcal{A}_N).$$

Suppose $n \in \mathbb{N}_0$ with $n < N$ such that

$$\tilde{Y}_{n+1}^{(m+1)} \leq \mathbb{E}(X_{\sigma_{n+1}} | \mathcal{A}_{n+1}).$$

Due to $n \leq \tau_n^{(m+1)} \leq \sigma_n$ we have

$$\{\sigma_n = n\} \subseteq \{\tau_n^{(m+1)} = n\} = C_n^{(m+1)}$$

and by properties of conditional expectations we have

$$\{\sigma_n = n\} \subseteq \{\mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) = X_n\}.$$

Thus it follows

$$\tilde{Y}_n^{(m+1)} \leq \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) = X_n \text{ on } \{\sigma_n = n\}.$$

We have $\{\sigma_n > n\} \subseteq \{\sigma_n = \sigma_{n+1}\}$, hence on $\{\sigma_n > n\}$

$$\begin{aligned} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) &= \mathbb{E}(X_{\sigma_{n+1}} | \mathcal{A}_n) = \mathbb{E}(\mathbb{E}(X_{\sigma_{n+1}} | \mathcal{A}_{n+1}) | \mathcal{A}_n) \\ &\geq \mathbb{E}(\tilde{Y}_{n+1}^{(m+1)} | \mathcal{A}_n) && \text{by ind.hyp.} \\ &\geq \hat{Y}_n^{(m+1)} && \text{by 2.16.} \end{aligned}$$

On $\{\tau_n^{(m)} > n\}$ we have $\hat{Y}_n^{(m+1)} = \tilde{Y}_n^{(m+1)}$.

Hence we have $\mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \geq \hat{Y}_n^{(m+1)} = \tilde{Y}_n^{(m+1)}$ on $\{\sigma_n > n\} \cap \{\tau_n^{(m)} > n\}$.

Furthermore we have

$$\begin{aligned} \{\sigma_n > n\} &\subseteq \{\hat{\tau}_n^{(m+1)} > n\} = \left(\hat{C}_n^{(m+1)}\right)^c \stackrel{3.6}{=} \left(\{\hat{Y}_n^{(m+1)} < X_n\} \cap C_n^{(0)}\right)^c \\ &= \{\hat{Y}_n^{(m+1)} \geq X_n\} \cup \left(C_n^{(0)}\right)^c. \end{aligned}$$

Since $m > 0$ or $\tau_n^{(0)}$ is corresponding to $C^{(0)}$ we have

$$\{\tau_n^{(m)} = n\} = C_n^{(m)} \subseteq C_n^{(0)}$$

and thus

$$\begin{aligned} \{\sigma_n > n\} \cap \{\tau_n^{(m)} = n\} &\subseteq \left(\{\hat{Y}_n^{(m+1)} \geq X_n\} \cup (C_n^{(0)})^c \right) \cap C_n^{(0)} \\ &\subseteq \{\hat{Y}_n^{(m+1)} \geq X_n\}. \end{aligned}$$

So we have

$$E(X_{\sigma_n} | \mathcal{A}_n) \geq \hat{Y}_n^{(m+1)} \geq X_n \text{ on } \{\sigma_n > n\} \cap \{\tau_n^{(m)} = n\}$$

and hence

$$E(X_{\sigma_n} | \mathcal{A}_n) \geq \max\{X_n, \hat{Y}_n^{(m+1)}\} \stackrel{(3.1.12)}{=} \tilde{Y}_n^{(m+1)} \text{ on } \{\sigma_n > n\} \cap \{\tau_n^{(m)} = n\}. \quad \square$$

Proof of the second statement:

We have for all $n \in \mathbb{N}_0$ with $n \leq N$

$$X_n \stackrel{3.7}{\leq} \tilde{Y}_n^{(m+1)} \leq E(X_{\sigma_n} | \mathcal{A}_n) \text{ on } C_n^{(0)}. \quad \square$$

3.9 Remark

Consider $n \in \mathbb{N}_0$ with $n \leq N$ and $m \in \mathbb{N}$.

Then we have

$$\tau_n^{(m)} \leq \tau_n^{(m+1)} \leq \tau_n^{(\infty)}.$$

If

$$\tau_k^{(0)} \text{ is corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

this is true for $m = 0$ as well.

The next theorem shows that under certain conditions the limiting value of this increasing sequence of stopping rules is smaller than certain optimal stopping rules. We shall see later (in Theorem 3.18 (page 42)) that this limit is equal to the smallest optimal stopping rule. This theorem was proved for $C^{(0)} \equiv \Omega$ in [KS06, p. 35; Prop. 4.1], here we state it for a more arbitrary situation.

3.10 Theorem

Consider $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N$$

and some optimal $\sigma \in \mathcal{S}_n^N \cap \mathcal{S}(C^{(0)})$.

Then we have

$$\tau_n^{(\infty)} \leq \sigma.$$

Proof: We have

$$\tau_n^{(m)} = \inf \left\{ p ; n \leq p \leq N \wedge \mathbf{1}_{C_p^{(m)}} = 1 \right\}$$

and due to Lemma 3.6 (page 34)

$$C_n^{(m)} = \left\{ \tilde{Y}_n^{(m)} \leq X_n \right\} \cap C_n^{(0)},$$

thus we have

$$\tau_n^{(m)} = \inf \left\{ p ; n \leq p \leq N \wedge \tilde{Y}_p^{(m)} \leq X_p \wedge \mathbf{1}_{C_p^{(0)}} = 1 \right\}.$$

Combined with the assumption that $\sigma \in \mathcal{S}(C^{(0)})$, we have due to the infimum above

$$\left\{ \tilde{Y}_\sigma^{(m)} \leq X_\sigma \right\} \subseteq \left\{ \tau_n^{(m)} \leq \sigma \right\}.$$

We have

$$X_\sigma = Y_\sigma^*$$

due to the optimality of σ and

$$Y_\sigma^* \geq \tilde{Y}_\sigma^{(m)}$$

due to the definition of Y^* in Definition 1.3 (page 14) and the definition of $\tilde{Y}^{(m)}$ in Algorithm 2.13 (page 24). Hence

$$\tilde{Y}_\sigma^{(m)} \leq X_\sigma,$$

which yields $\tau_n^{(m)} \leq \sigma$. This is true for all $m \in \mathbb{N}$, thus we have $\tau_n^{(\infty)} \leq \sigma$ by definition. \square

The next lemma has been proved for the special case of $C^{(0)} \equiv \Omega$ and $\sigma = \tau^{(m+1)}$ via backward induction in [KS06, p. 36; lemma 4.2] which is generalized here to an arbitrary improver σ .

3.11 Lemma

Consider $m \in \mathbb{N}_0$, $n \in \mathbb{N}_0$ with $n \leq N$ and

$$\tau_k^{(0)} \text{ corresponding to } C^{(0)} \text{ for all } k \in \mathbb{N}_0 \text{ with } n \leq k \leq N.$$

We have for all $\sigma \in \mathcal{T}$ being an improver of $\tau^{(m)}$

$$\mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \geq \tilde{Y}_n^{(m+1)} \geq \mathbb{E}\left(Y_{n+1}^{(m)} | \mathcal{A}_n\right),$$

thus it follows

$$Y_n^{(m+1)} \geq \tilde{Y}_n^{(m+1)} \geq \mathbb{E}\left(Y_{n+1}^{(m)} | \mathcal{A}_n\right).$$

Proof: The inequality on the left side is true due to Theorem 3.8 (page 36) and Theorem 3.3 (page 31) respectively, the other one can easily be verified by a look at the definitions. \square

3.12 Lemma

Let $\sigma \in (\mathcal{S}_0^N)^{\mathbb{N}_0}$ non-decreasing. Then we have

$$\lim_{n \rightarrow \infty} X_{\sigma_n} = X_{\lim_{n \rightarrow \infty} \sigma_n}$$

and

$$\mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} | \mathcal{A}_j\right) = \lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_j) \text{ for all } j \in \mathbb{N}_0.$$

Proof: For all $n \in \mathbb{N}_0$ σ_n is valued in the discrete set $\{0, \dots, N\}$. Hence obviously

$$\left\{ \sup \{ \sigma_n ; n \in \mathbb{N}_0 \} \leq N < \infty \right\} = \Omega,$$

which implies the first line. By Lemma 1.4 (page 15) $\sup_{n \in \mathbb{N}_0} |X_{\sigma_n}|$ is integrable. So dominated convergence for conditional expectations shows the second assertion. (For details of this argument see [Irl05, pp. 126, 127].) \square

Now we want to generalize a statement made in [KS06, p. 36] for $C^{(0)} \equiv \Omega$. This condition can be dropped without any change to the statement itself.

3.13 Conclusion

For all $n \in \mathbb{N}_0$ with $n \leq N$ we have

$$Y_n^{(\infty)} = \mathbb{E} \left(X_{\tau_n^{(\infty)}} \middle| \mathcal{A}_n \right).$$

Proof: This follows directly by Lemma 3.12 (page 38) with $\sigma = (\tau_n^{(m)})_{m \in \mathbb{N}_0}$. \square

Next we show that the algorithm terminates with optimal results when started with $C^{(0)} \equiv \Omega$. A condensed proof can be found in [KS06, pp. 36-37; Proposition 4.3 and Proposition 4.4]; we give herein an extensive proof.

3.14 Theorem: Optimal Algorithm Termination for $C^{(0)} \equiv \Omega$

Consider $C^{(0)} \equiv \Omega$ and $\tau^{(0)}$ and $C^{(0)}$ corresponding.

For all $n \in \mathbb{N}_0$ with $n \leq N$ and $m \in \mathbb{N}_0$ with $N - n \leq m$ we have

$$\begin{aligned} \tau_n^{(m)} &= \tau_n^{(\infty)} = \tau_n^*, \\ Y_n^{(m)} &= Y_n^{(\infty)} = Y_n^*. \end{aligned}$$

Proof: We will prove both sets of equations by backward induction over n .

The latter set of equations:

We have

$$\tau_N^{(0)} = \inf \left\{ p; N \leq p \leq N, \mathbf{1}_{C_p^{(0)}} = 1 \right\} \equiv N$$

and hence for all $m \in \mathbb{N}_0$ with $0 = N - N \leq m$

$$Y_N^{(m)} = Y_N^{(\infty)} = Y_N^* = X_N.$$

Consider some $n \in \mathbb{N}_0$ with $n < N$ and assume that the assertion is proved for $n + 1$.

Then we have $1 \leq N - n$.

Consider $m \in \mathbb{N}_0$ with $N - n \leq m$, hence we have $1 \leq m$.

Due to $N - (n + 1) \leq m - 1$ and $0 \leq m - 1$ we have

$$(3.14.1) \quad Y_{n+1}^{(m-1)} = Y_{n+1}^*$$

by induction hypothesis.

Consider a consistent N -suitable family of stopping rules σ , being an improver of $\tau^{(m-1)}$.

Since $\tau^{(0)}$ and $C^{(0)}$ are corresponding it follows

$$\mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \stackrel{3.11}{\geq} \mathbb{E} \left(Y_{n+1}^{(m-1)} \middle| \mathcal{A}_n \right) \stackrel{(3.14.1)}{=} \mathbb{E} \left(Y_{n+1}^* \middle| \mathcal{A}_n \right)$$

and due to $C_n^{(0)} \equiv \Omega$

$$\mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \stackrel{3.8}{\geq} X_n.$$

Thus it follows

$$Y_n^* \geq \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) \geq \max \left\{ X_n, \mathbb{E} \left(Y_{n+1}^* \middle| \mathcal{A}_n \right) \right\} \stackrel{1.7}{=} Y_n^*,$$

hence

$$Y_n^* = \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n).$$

This is especially true for $\sigma = \tau^{(m)}$, hence we have $Y_n^* = Y_n^{(m)}$.

So $Y_n^* = Y_n^{(m)}$ for all $m \in \mathbb{N}_0$ with $N - n \leq m$ and $1 < m$, which implies

$$Y_n^{(\infty)} = Y_n^*.$$

The first set of equations:

For all $m \in \mathbb{N}_0$ with $0 = N - N \leq m$ we have

$$\tau_N^{(m)} \equiv N \equiv \check{\tau}_N^*.$$

Consider $n \in \mathbb{N}_0$ with $1 \leq n \leq N$ and suppose the assertion is proved for all $j \in \mathbb{N}_0$ with $n \leq j \leq N$. Consider $m \in \mathbb{N}_0$ with $N - (n - 1) \leq m$, hence $m \geq 1$ and $m - 1 \geq N - n \geq 0$. Then we have

$$\begin{aligned} C_{n-1}^{(m)} &= \left\{ X_{n-1} \geq \tilde{Y}_{n-1}^{(m)} \right\} \cap C_{n-1}^{(0)} && \text{by 3.6 and corresponding-assumption} \\ &= \left\{ X_{n-1} \geq \tilde{Y}_{n-1}^{(m)} \right\} && \text{by assumption} \\ &\subseteq \left\{ X_{n-1} \geq \mathbb{E} \left(X_{\tau_n^{(m-1)}} \middle| \mathcal{A}_{n-1} \right) \right\} && \text{by definition} \\ &= \left\{ X_{n-1} \geq \mathbb{E} \left(X_{\check{\tau}_n^*} \middle| \mathcal{A}_{n-1} \right) \right\} && \text{by ind.hyp.} \end{aligned}$$

and by definition

$$\begin{aligned} Y_{n-1}^* &= \max \left\{ X_{n-1}, \mathbb{E} \left(Y_n^* \middle| \mathcal{A}_{n-1} \right) \right\} \\ &= \max \left\{ X_{n-1}, \mathbb{E} \left(\mathbb{E} \left(X_{\check{\tau}_n^*} \middle| \mathcal{A}_n \right) \middle| \mathcal{A}_{n-1} \right) \right\} \\ &= \max \left\{ X_{n-1}, \mathbb{E} \left(X_{\check{\tau}_n^*} \middle| \mathcal{A}_{n-1} \right) \right\}, \end{aligned}$$

hence

$$C_{n-1}^{(m)} \subseteq \left\{ X_{n-1} = Y_{n-1}^* \right\}.$$

This shows

$$(3.14.2) \quad C_{n-1}^{(m)} \subseteq \left\{ \check{\tau}_{n-1}^* = n - 1 \right\}.$$

Due to $C_n^{(0)} = \Omega$ we have by definition

$$C_{n-1}^{(m)} = \left\{ \tau_{n-1}^{(m)} = n - 1 \right\}$$

and hence we have, due to $\tau^{(m)}$ and τ^* being consistent,

$$\begin{aligned} \tau_{n-1}^{(m)} &= \mathbb{1}_{C_{n-1}^{(m)}} \tau_{n-1}^{(m)} + \mathbb{1}_{\Omega \setminus C_{n-1}^{(m)}} \tau_{n-1}^{(m)} \\ &= \mathbb{1}_{C_{n-1}^{(m)}} (n - 1) + \mathbb{1}_{\Omega \setminus C_{n-1}^{(m)}} \tau_n^{(m)} \\ &= \mathbb{1}_{C_{n-1}^{(m)}} (n - 1) + \mathbb{1}_{\Omega \setminus C_{n-1}^{(m)}} \check{\tau}_n^* && \text{by ind.hyp.} \\ &= \mathbb{1}_{C_{n-1}^{(m)}} \tau_{n-1}^* + \mathbb{1}_{\Omega \setminus C_{n-1}^{(m)}} \check{\tau}_n^* && \text{by (3.14.2)} \\ &= \mathbb{1}_{C_{n-1}^{(m)}} \tau_{n-1}^* + \mathbb{1}_{\Omega \setminus C_{n-1}^{(m)}} \check{\tau}_{n-1}^* \\ &= \check{\tau}_{n-1}^*. \end{aligned}$$

So $\check{\tau}_{n-1}^* = \tau_{n-1}^{(m)}$ on $C_n^{(0)} = \Omega$ for all $m \in \mathbb{N}_0$ with $N - n \leq m$, which implies

$$\tau_{n-1}^{(\infty)} = \check{\tau}_{n-1}^*. \quad \square$$

Comment: The proof above does not work for arbitrary $C^{(0)}$, but for corresponding $C^{(0)}$ and $\tau^{(0)}$ we get similar results, mentioned in Theorem 3.18 (page 42) and proved later in Section 5 (page 49).

3.15 Definition and Remark

Now we want to consider in parallel a second instance of the algorithm with other starting parameters. To be able to distinguish it from our normal algorithm we will underline all variables of the second instance twice. We choose

$$\underline{\underline{X}} := \left(X_n \mathbf{1}_{C_n^{(0)}} \right)_{n \in \mathbb{N}_0}, \underline{\underline{C}}^{(0)} := (\Omega)_{n \in \mathbb{N}_0}, \underline{\underline{\tau}}_n^{(0)} := n.$$

Hence $\underline{\underline{\tau}}^{(0)}$ and $\underline{\underline{C}}^{(0)}$ are corresponding and we have by Theorem 3.14 (page 39) for all $n \in \mathbb{N}_0$ with $n \leq N$ and $m \in \mathbb{N}_0$ with $N - n \leq m$

$$\underline{\underline{\tau}}_n^{(m)} = \underline{\underline{\tau}}_n^{(\infty)} = \underline{\underline{\check{\tau}}}_n^*,$$

$$\underline{\underline{Y}}_n^{(m)} = \underline{\underline{Y}}_n^{(\infty)} = \underline{\underline{Y}}_n^*.$$

In [BKS08, Proposition 3.4] it was stated without proof that, given nonnegative X and starting with $\tau^{(0)}$ such that $\tau_n^{(0)} \in \mathcal{S}(C^{(0)})$ for each $n \in \mathbb{N}_0$ with $n \leq N$, we will end up with $\tau^{(\infty)} = \check{\tau}^*$. But the following examples will show that this is wrong, as it may happen that $\check{\tau}^* \notin \mathcal{S}(C^{(0)})$ despite $\tau^{(\infty)} \in \mathcal{S}(C^{(0)})$.

3.16 Example

Consider $N = 1$, $X_0 \equiv 0$, $X_1 \equiv 0$, $C_0^{(0)} = \emptyset$ and $C_1^{(0)} = \Omega$.

Then we have $\underline{\underline{X}} = X$ and $\underline{\underline{\check{\tau}}}_0^* \equiv 0$.

We will start with $\tau^{(0)}$ such that $\tau_n^{(0)} \in \mathcal{S}(C^{(0)})$ for each $n \in \mathbb{N}_0$ with $n \leq N$. And we end up with $\tau_0^{(\infty)} \equiv 1$, obviously $\underline{\underline{\check{\tau}}}_0^* \equiv 0 < 1 \equiv \tau_0^{(\infty)}$, hence $\tau^{(\infty)} \neq \check{\tau}^*$.

This example may be easily extended as follows.

3.17 Generalized Example

Consider nonnegative X and $C^{(0)}$ and $\tau^{(0)}$ such that $\tau_n^{(0)} \in \mathcal{S}(C^{(0)})$ for each $n \in \mathbb{N}_0$ with $n \leq N$. Let σ be corresponding to $C^{(0)}$.

If $\check{\tau}^* \not\leq \sigma$, then $\tau^{(\infty)} \neq \check{\tau}^*$.

This is indeed true if we have

$$P \left(\bigcup_{0 \leq n < N} \bigcap_{N \geq m \geq n+1} \left\{ \underline{\underline{X}}_m = 0 \right\} \setminus C_n^{(0)} \right) > 0.$$

Proof: We have $\underline{\underline{\tau}}^{(\infty)} = \underline{\underline{\check{\tau}}}_0^* \not\leq \sigma \leq \tau^{(0)} \leq \tau^{(\infty)}$. □

3.18 Theorem: Algorithm Termination

- If $\tau^{(0)}$ and $C^{(0)}$ are corresponding, then

$\tau^{(\infty)}$ is (pointwise) smaller than each $\sigma \in \mathcal{T}$, which is optimal in $\mathcal{S}(C^{(0)})$.

- We always have

$\tau^{(\infty)}$ is optimal in $\mathcal{S}(C^{(0)})$.

- If $C^{(0)}$ is essential, we even have

$\tau^{(\infty)}$ is optimal.

Proof: The first statement follows due to Theorem 3.10 (page 37).

We could not prove the latter statements by backward induction. The arguments in [BKS08] use [BKS08, Proposition 3.4] which is not true as the examples above show. So mentioning these results here is in anticipation of Theorem 5.11 (page 57) and Conclusion 5.13 (page 61). \square

Compared with Theorem 3.14 (page 39) the loss for essential $C^{(0)}$ (instead of $C^{(0)} \equiv \Omega$) is having not the smallest optimal stopping rule and the convergence speed has not to be that fast.

Next we examine the parameter κ . After showing by an example the influence of κ on the number of iterations, we will prove that increasing κ does never lead to worse results. Roughly spoken this is: A bigger κ leads to a smaller number of iterations, in doing so the effort of calculating the iteration increases. For the first iteration step this was mentioned in [KS06, Proposition 3.3].

3.19 Example

Consider $N = 12$ and $X \equiv (4, 3, 6, 4, 2, 3, 1, 0, 3, 5, 3, 1)$. The only optimal N-suitable family of stopping rules is $\check{\tau}^* = (3, 3, 3, 10, 10, 10, 10, 10, 10, 10, 11, 12)$. The choice of κ has big influence on the number of iterations. Changes are shown in bold.

Given κ with

$$\begin{array}{l}
\kappa(0) = 1, \kappa(1) = 1, \kappa(2) \leq 2 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (1, \mathbf{3}, 3, 4, \mathbf{6}, 6, 7, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (\mathbf{3}, \mathbf{3}, \mathbf{3}, 4, 6, \mathbf{6}, \mathbf{10}, 10, 10, 10, 11, 12), \\
\tau^{(3)} = (3, 3, 3, 4, \mathbf{10}, \mathbf{10}, 10, 10, 10, 10, 11, 12), \\
\tau^{(4)} = (3, 3, 3, \mathbf{10}, 10, 10, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) = 1, \kappa(1) = 1, \kappa(2) \geq 3 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (1, \mathbf{3}, 3, 4, \mathbf{6}, 6, 7, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (\mathbf{3}, \mathbf{3}, \mathbf{3}, 4, 6, \mathbf{6}, \mathbf{10}, 10, 10, 10, 11, 12), \\
\tau^{(3)} = (3, 3, 3, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) = 1, \kappa(1) = 2 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (1, \mathbf{3}, 3, 4, \mathbf{6}, 6, 7, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (3, 3, 3, 4, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 10, 10, 10, 11, 12), \\
\tau^{(3)} = (3, 3, 3, \mathbf{10}, 10, 10, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) \in \{2, 3\} \text{ and } \kappa(1) \leq 2 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (\mathbf{3}, \mathbf{3}, 3, 4, \mathbf{6}, 6, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (3, 3, 3, 4, \mathbf{10}, \mathbf{10}, 10, 10, 10, 10, 10, 11, 12), \\
\tau^{(3)} = (3, 3, 3, \mathbf{10}, 10, 10, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) \in \{2, 3\} \text{ and } \kappa(1) \geq 3 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (\mathbf{3}, \mathbf{3}, 3, 4, \mathbf{6}, 6, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (3, 3, 3, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) \in \{4, 5\} \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (\mathbf{3}, \mathbf{3}, 3, 4, \mathbf{10}, \mathbf{10}, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 11, 12), \\
\tau^{(2)} = (3, 3, 3, \mathbf{10}, 10, 10, 10, 10, 10, 10, 10, 11, 12) = \check{\tau}^*,
\end{array} \\
\kappa(0) \geq 6 \text{ we have } \begin{array}{l}
\tau^{(0)} = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12), \\
\tau^{(1)} = (\mathbf{3}, \mathbf{3}, 3, \mathbf{10}, \mathbf{10}, \mathbf{10}, \mathbf{10}, \mathbf{10}, \mathbf{10}, 10, 11, 12) = \check{\tau}^*.
\end{array}
\end{array}$$

3.20 Lemma

Consider $\sigma, \tau \in \mathcal{T}$ with $\tau \leq \sigma$ and $i, p \in \mathbb{N}_0$ with $i \leq p$. Then we have

$$X_{\sigma_i} = X_{\sigma_p} \text{ on } \{\tau_i = p\}.$$

Proof: By the definition of \mathcal{T} in Definition 1.10 (page 16) we have

$$\{\sigma_n > n\} \subseteq \{\sigma_n = \sigma_{n+1}\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n < N,$$

hence it follows by induction

$$\{\sigma_i \geq p\} \subseteq \{\sigma_i = \sigma_p\}.$$

Due to $\tau \leq \sigma$ we have $\sigma_i \geq p$ on $\{\tau_i = p\}$ and hence $\sigma_i = \sigma_p$ on $\{\tau_i = p\}$. □

3.21 Theorem: Influence of κ

Assume $C^{(0)} \equiv \Omega$. Consider a second instance of the algorithm, where all variables are marked with a $'$, with identical input variables except for κ' which is only assumed to be nowhere smaller than κ . We have

$$\begin{aligned}
\widehat{Y}_n^{(m)} &\leq \widehat{Y}'_n^{(m)}, \\
\widetilde{Y}_n^{(m)} &\leq \widetilde{Y}'_n^{(m)}, \\
C_n^{(m)} &\supseteq C'_n^{(m)}, \\
\widehat{C}_n^{(m)} &\supseteq \widehat{C}'_n^{(m)}, \\
\Theta_n^{(m)} &\supseteq \Theta'_n^{(m)}, \\
\widehat{\Theta}_n^{(m)} &\supseteq \widehat{\Theta}'_n^{(m)}, \\
\tau_n^{(m)} &\leq \tau'_n^{(m)}, \\
\widehat{\tau}_n^{(m)} &\leq \widehat{\tau}'_n^{(m)}, \\
Y_i^{(m)} &\leq Y'_i^{(m)}.
\end{aligned}$$

Proof: We will show this by induction. The initial assumption of this theorem will not be necessary for $m = 0$ but in the induction step in (3.21.2).

Since we use the same input variables the induction hypothesis is true for $m = 0$. Consider $m \in \mathbb{N}$ and the induction hypothesis being true for $m - 1$. Then it is true for m by definition in Part 2.13 (page 24) except for the last inequality. Since $\tau_n^{(m)} \leq \tau'_n^{(m)}$ for all $n \in \mathbb{N}_0$ with $n \leq N$ we have by Lemma 3.20 (page 43)

$$(3.21.1) \quad X_{\tau'_i^{(m)}} = X_{\tau_p^{(m)}} \text{ on } \left\{ \tau_i^{(m)} = p \right\} \text{ for all } i, p \in \mathbb{N}_0 \text{ with } i \leq p \leq N$$

By Lemma 3.2 (page 30) and Conclusion 3.4 (page 32) we have for all $p \in \mathbb{N}_0$ with $p \leq N$

$$(3.21.2) \quad X_p \leq Y'_p \leq Y_p = \mathbb{E} \left(X_{\tau'_p} \middle| \mathcal{A}_p \right) \quad \text{on} \quad C_p^{(0)} = \Omega.$$

Thus we have for all $i \in \mathbb{N}_0$ with $i \leq N$

$$\begin{aligned}
Y_i^{(m)} &\stackrel{\text{def.}}{=} \mathbb{E} \left(X_{\tau'_i} \middle| \mathcal{A}_i \right) \\
&= \mathbb{E} \left(\sum_{p=i}^N \mathbb{1}_{\{\tau_i^{(m)}=p\}} X_{\tau'_i} \middle| \mathcal{A}_i \right) \\
&\stackrel{(3.21.1)}{=} \mathbb{E} \left(\sum_{p=i}^N \mathbb{1}_{\{\tau_i^{(m)}=p\}} X_{\tau'_p} \middle| \mathcal{A}_i \right) \\
&= \mathbb{E} \left(\sum_{p=i}^N \mathbb{1}_{\{\tau_i^{(m)}=p\}} \mathbb{E} \left(X_{\tau'_p} \middle| \mathcal{A}_p \right) \middle| \mathcal{A}_i \right) \\
&\stackrel{(3.21.2)}{\geq} \mathbb{E} \left(\sum_{p=i}^N \mathbb{1}_{\{\tau_i^{(m)}=p\}} X_p \middle| \mathcal{A}_i \right) \\
&= \mathbb{E} \left(X_{\tau_i} \middle| \mathcal{A}_i \right) \\
&\stackrel{\text{def.}}{=} Y_i^{(m)}. \quad \square
\end{aligned}$$

4 Different Finite Time Horizons Compared

4.1 Idea

As written in Convention 1.9 (page 16) nearly all variables depend on N . For easier reading we mostly did not include this dependence in our notations. Now it will be important to have a look at the number of steps N , so whenever necessary we shall write N as an additional inner subscript. For example we will write $C_{N,k}^{(0)}$ instead of $C_k^{(0)}$.

4.2 Example

A first idea was that we might have $\tau_{N,n}^{(m)} = \tau_{\infty,n}^{(m)} \wedge N$ for m, N, n , at least under reasonable assumptions. But this does not hold in general as the following example shows:

Example: Consider $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$ with $N_1 < N_2$.

Consider $C_{N_i,n}^{(0)} = \Omega$ for all $n \in \mathbb{N}_0$ and $i \in \{1, 2\}$.

Consider a process X given for all $n \in \mathbb{N}$ by

$$X_n = \begin{cases} 1 & \text{if } n < N_1, \\ 0 & \text{if } n = N_1, \\ 2 & \text{if } n > N_1. \end{cases}$$

Then we have for all $i \in \{1, 2\}$ and $n \in \mathbb{N}_0$ with $n \leq N_i$

$$\tau_{N_i,n}^{(0)} = n$$

and

$$Y_{N_i,n}^{(0)} = X_n = \begin{cases} 1 & \text{if } n < N_1, \\ 0 & \text{if } n = N_1, \\ 2 & \text{if } n > N_1. \end{cases}$$

For all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_1$ we have

$$\tilde{Y}_{N_1,n}^{(m)} = \begin{cases} 1 & \text{if } n < N_1, \\ 0 & \text{otherwise.} \end{cases}$$

For all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_2$

$$\tilde{Y}_{N_2,n}^{(m)} = \begin{cases} 1 & \text{if } n \leq N_1 - \sum_{k=0}^{m-1} \kappa(k), \\ 2 & \text{otherwise.} \end{cases}$$

So we have for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_1$

$$C_{N_1,n}^{(m)} = \Omega,$$

and for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_2$

$$C_{N_2,n}^{(m)} = \begin{cases} \emptyset & \text{if } N_1 - \sum_{k=0}^{m-1} \kappa(k) < n \leq N_1, \\ \Omega & \text{otherwise.} \end{cases}$$

Thus we have for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\tau_{N_1, n}^{(m)} = n,$$

and for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_2$

$$\tau_{N_2, n}^{(m)} = \begin{cases} N_1 + 1 & \text{if } N_1 - \sum_{k=0}^{m-1} \kappa(k) < n \leq N_1, \\ n & \text{otherwise.} \end{cases}$$

And we have for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_1$

$$Y_{N_1, n}^{(0)} = \begin{cases} 1 & \text{if } n < N_1, \\ 0 & \text{if } n = N_1, \end{cases}$$

and for all $m \in \mathbb{N}$ and $n \in \mathbb{N}_0$ with $n \leq N_2$

$$Y_{N_2, n}^{(0)} = \begin{cases} 2 & \text{if } n > N_1 - \sum_{k=0}^{m-1} \kappa(k), \\ 1 & \text{otherwise.} \end{cases}$$

So it follows that for all $m \in \mathbb{N}$ and for all $n \in \mathbb{N}_0$ with $N_1 - \sum_{k=0}^{m-1} \kappa(k) \leq n < N_1$

$$\tau_{N_1, n}^{(m)} = n < N_1 = \tau_{N_2, n}^{(m)} \wedge N_1. \quad \square$$

4.3 Lemma

Let $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$ with $N_1 < N_2$, $m \in \mathbb{N}_0$ with

$$(4.3.1) \quad m = 0 \implies X_n \leq Y_{N_2, n}^{(0)} \text{ on } \left\{ \tau_{N_1, n}^{(0)} = n \right\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N_1$$

and

$$(4.3.2) \quad m = 1 \implies \tau_{N_2, n}^{(0)} \equiv n \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N_1.$$

Assume $C_{N_2}^{(0)} \equiv \Omega$.

Suppose that

$$\tau_{N_1, n}^{(m)} \leq \tau_{N_2, n}^{(m)} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N_1.$$

Then we have

$$Y_{N_1, n}^{(m)} \leq Y_{N_2, n}^{(m)} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N_1.$$

Remark: The proof given below does not work for arbitrary $C_{N_2}^{(0)}$.

Proof by Backward Induction:

We have

$$Y_{N_1, N_1}^{(m)} = \mathbb{E} \left(X_{\tau_{N_1, N_1}^{(m)}} \middle| \mathcal{A}_{N_1} \right) = X_{N_1} \stackrel{\text{ass.}}{\underset{3.7}{\leq}} Y_{N_2, N_1}^{(m)} \text{ on } C_{N_2, N_1}^{(0)} = \Omega.$$

Consider some $n \in \mathbb{N}_0$ with $n < N_1$ and $Y_{N_1, n+1}^{(m)} \leq Y_{N_2, n+1}^{(m)}$.

On the set $\left\{ \tau_{N_1, n}^{(m)} = n \right\}$ we have

$$Y_{N_1, n}^{(m)} = X_n \stackrel{(4.3.2), 3.7}{\underset{(4.3.1)}{\leq}} Y_{N_2, n}^{(m)} \text{ on } C_{N_2, n}^{(0)} = \Omega.$$

For all $i \in \{1, 2\}$ we have

$$(4.3.3) \quad \left\{ \tau_{N_i, n}^{(m)} > n \right\} \subseteq \left\{ \tau_{N_i, n}^{(m)} = \tau_{N_i, n+1}^{(m)} \right\}$$

and thus on $\{\tau_{N_i,n}^{(m)} > n\}$

$$\begin{aligned}
Y_{N_i,n}^{(m)} &= \mathbb{E} \left(X_{\tau_{N_i,n}^{(m)}} \middle| \mathcal{A}_n \right) && \text{by def.} \\
&= \mathbb{E} \left(X_{\tau_{N_i,n+1}^{(m)}} \middle| \mathcal{A}_n \right) && \text{by (4.3.3)} \\
&= \mathbb{E} \left(\mathbb{E} \left(X_{\tau_{N_i,n+1}^{(m)}} \middle| \mathcal{A}_{n+1} \right) \middle| \mathcal{A}_n \right) \\
&= \mathbb{E} \left(Y_{N_i,n+1}^{(m)} \middle| \mathcal{A}_n \right) && \text{by def.}
\end{aligned}$$

So we have on $\{\tau_{N_1,n}^{(m)} > n\} \subseteq \{\tau_{N_2,n}^{(m)} > n\}$

$$Y_{N_1,n}^{(m)} = \mathbb{E} \left(Y_{N_1,n+1}^{(m)} \middle| \mathcal{A}_n \right) \stackrel{\text{ind.ass.}}{\leq} \mathbb{E} \left(Y_{N_2,n+1}^{(m)} \middle| \mathcal{A}_n \right) = Y_{N_2,n}^{(m)}. \quad \square$$

4.4 Lemma

Consider $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$ with $N_1 < N_2$.

Assume $C_{N_2}^{(0)} = C_{N_1}^{(0)} \equiv \Omega$.

Consider $m \in \mathbb{N}_0$ satisfying assumptions (4.3.1) (page 46) and (4.3.2) (page 46).

Suppose that for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\tau_{N_1,n}^{(m)} \leq \tau_{N_2,n}^{(m)}.$$

Then we have for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\begin{aligned}
Y_{N_1,n}^{(m)} &\leq Y_{N_2,n}^{(m)}, \\
\tilde{Y}_{N_1,n}^{(m)} &\leq \tilde{Y}_{N_2,n}^{(m)}, \\
C_{N_2,n}^{(m+1)} &\subseteq C_{N_1,n}^{(m+1)}, \\
\tau_{N_1,n}^{(m+1)} &\leq \tau_{N_2,n}^{(m+1)}.
\end{aligned}$$

Proof: The first inequality follows by Lemma 4.3 (page 46).

For the other inequalities define for all $i \in \{1, 2\}$ and for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$A_{N_i,n}^{(m)} := \{p \in \mathbb{N}_0 ; n \leq p < \min\{n + \kappa(m), N_i\} + 1\}.$$

Thus we have for all $n \in \mathbb{N}_0$ with $n \leq N_1$ and for all $p \in A_{N_1,n}^{(m)}$

$$\mathbb{E} \left(Y_{N_1,p}^{(m)} \middle| \mathcal{A}_n \right) \leq \mathbb{E} \left(Y_{N_2,p}^{(m)} \middle| \mathcal{A}_n \right).$$

Furthermore we have for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$A_{N_1,n}^{(m)} \subseteq A_{N_2,n}^{(m)}.$$

So we have for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\begin{aligned}
\tilde{Y}_{N_1,n}^{(m+1)} &= \sup_{p \in A_{N_1,n}^{(m)}} \mathbb{E} \left(Y_{N_1,p}^{(m)} \middle| \mathcal{A}_n \right) \leq \sup_{p \in A_{N_1,n}^{(m)}} \mathbb{E} \left(Y_{N_2,p}^{(m)} \middle| \mathcal{A}_n \right) \\
&\leq \sup_{p \in A_{N_2,n}^{(m)}} \mathbb{E} \left(Y_{N_2,p}^{(m)} \middle| \mathcal{A}_n \right) = \tilde{Y}_{N_2,n}^{(m+1)},
\end{aligned}$$

thus we have

$$\left\{ \tilde{Y}_{N_1,n}^{(m+1)} \leq X_n \right\} \subseteq \left\{ \tilde{Y}_{N_2,n}^{(m+1)} \leq X_n \right\},$$

hence

$$C_{N_2,n}^{(m+1)} \subseteq C_{N_1,n}^{(m+1)}.$$

This implies for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\tau_{N_1,n}^{(m+1)} \leq \tau_{N_2,n}^{(m+1)}.$$

□

4.5 Theorem

Consider $N_1, N_2 \in \mathbb{N} \cup \{\infty\}$ with $N_1 < N_2$.

Assume $C_{N_2}^{(0)} = C_{N_1}^{(0)} \equiv \Omega$, $\tau_{N_1}^{(0)}$ and $\tau_{N_2}^{(0)}$ determined by $C^{(0)}$ and

$$X_n \leq Y_{N_2,n}^{(0)} \text{ on } \left\{ \tau_{N_1,n}^{(0)} = n \right\} \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N_1.$$

Then we have for all $m \in \mathbb{N}_0$ and for all $n \in \mathbb{N}_0$ with $n \leq N_1$

$$\begin{aligned} Y_{N_1,n}^{(m)} &\leq Y_{N_2,n}^{(m)} \text{ on } C_{N_2,n}^{(0)}, \\ \tilde{Y}_{N_1,n}^{(m)} &\leq \tilde{Y}_{N_2,n}^{(m)} \text{ on } C_{N_2,n}^{(0)}, \\ C_{N_2,n}^{(m)} &\subseteq C_{N_1,n}^{(m)}, \\ \tau_{N_1,n}^{(m)} &\leq \tau_{N_2,n}^{(m)}. \end{aligned}$$

Proof: Proved by induction, using Lemma 4.4 (page 47).

□

5 General Case

The FII Algorithm introduced in Section 2 (page 18) gives optimal results for the finite time horizon as seen in Section 3 (page 28). In this section we show this for the infinite time horizon as well. The techniques within the proofs will be different since backward induction is not available here.

First we will repeat the result of [Irl80] in our nomenclature.

5.1 Theorem from [Irl80] for $\kappa \equiv 1$

Assume $\kappa \equiv 1$ and

$$E(X_\tau) \text{ exists for all } \tau \in \mathcal{S},$$

$$E(X_{\sup\{\sigma_n; n \in \mathbb{N}_0\}} | \mathcal{A}_0) = \lim_{n \rightarrow \infty} E(X_{\sigma_n} | \mathcal{A}_0) \text{ for all } \sigma \in \mathcal{S}^{\mathbb{N}_0} \text{ non-decreasing}$$

Then $\mathcal{S}_0^\infty = \mathcal{S}$ and

$$\begin{aligned} \tau_0^{(\infty)} &\text{ is optimal in } \mathcal{S} \\ \tau_0^{(m)} &\leq \tau_0^{(m+1)} \text{ for all } m \in \mathbb{N}_0, \\ Y_0^{(m)} &\leq Y_0^{(m+1)} \text{ for all } m \in \mathbb{N}_0. \end{aligned}$$

Thus it follows

$$\begin{aligned} Y_0^{(\infty)} &\stackrel{\text{def.}}{=} \lim_{\text{incr.}}_{m \rightarrow \infty} Y_0^{(m)} \\ &\stackrel{\text{def.}}{=} \lim_{m \rightarrow \infty} E(X_{\tau_0^{(m)}} | \mathcal{A}_0) \stackrel{\text{ass.}}{=} E\left(X_{\left(\lim_{m \rightarrow \infty} \tau_0^{(m)}\right)} | \mathcal{A}_0\right) \stackrel{\text{def.}}{=} \lim_{\text{incr.}} E(X_{\tau_0^{(\infty)}} | \mathcal{A}_0). \end{aligned}$$

If

$$C^{(l)} = C^{(l+1)} \text{ for some } l \in \mathbb{N}_0,$$

then obviously

$$C^{(n+l)} = C^{(l)} = C^{(\infty)} \text{ for all } n \in \mathbb{N}_0,$$

$$\tau_0^{(n+l)} = \tau_0^{(l)} = \tau_0^{(\infty)} \text{ for all } n \in \mathbb{N}_0.$$

Proof: A direct proof can be found in [Irl80, Part 2, pages 180-182], a proof for a more general situation follows within this section. \square

We will now expand this theorem to arbitrary κ (instead of $\kappa \equiv 1$) and arbitrary timepoint j (instead of looking only at time point 0), see Main Theorem 5.3 (page 50). In Chapter 9 (page 84) we will see that not using $\kappa = 1$ may increase the speed of the algorithm. The general assumption made above in [Irl80] as well as in [Irl06] and [Irl09] has to be slightly enlarged for our purposes to the General Assumption 5.2 (page 50).

5.2 General Assumption

Assume

$$(5.2.1) \quad E(X_\tau) \text{ exists for all } \tau \in \mathcal{S},$$

$$(5.2.2) \quad \lim_{n \rightarrow \infty} E(X_{\sigma_n} | \mathcal{A}_j) = E\left(X_{\lim_{n \rightarrow \infty} \sigma_n} | \mathcal{A}_j\right)$$

for all $j \in \mathbb{N}_0, \sigma \in \mathcal{S}^{\mathbb{N}_0}$ non-decreasing.

5.3 Main Theorem

$$\tau^{(\infty)} \text{ is optimal in } \mathcal{S}(C^{(0)})$$

and for all $m, j \in \mathbb{N}_0$ we have

$$\begin{aligned} \tau_j^{(m)} &\leq \tau_j^{(m+1)}, \\ Y_j^{(m)} &\leq Y_j^{(m+1)}. \end{aligned}$$

We have for all $j \in \mathbb{N}_0$

$$Y_j^{(\infty)} = E\left(X_{\tau_j^{(\infty)}} | \mathcal{A}_j\right).$$

If

$$C^{(l)} = C^{(l+1)} \quad \text{for some } l \in \mathbb{N}_0,$$

then we have

$$\begin{aligned} C^{(l)} = C^{(l+n)} = C^{(\infty)} &\text{ for all } n \in \mathbb{N}_0, \\ \tau^{(l)} = \tau^{(l+n)} = \tau^{(\infty)} &\text{ for all } n \in \mathbb{N}_0, \\ \Theta^{(l)} = \Theta^{(l+n)} = \Theta^{(\infty)} &\text{ for all } n \in \mathbb{N}_0, \\ Y^{(l)} = Y^{(l+n)} = Y^{(\infty)} &\text{ for all } n \in \mathbb{N}_0. \end{aligned}$$

If $C^{(0)} \equiv \Omega$ or $C^{(0)}$ is essential, then

$$\tau^{(\infty)} \text{ is optimal.}$$

We now give a reason why it has not been necessary to mention the General Assumption 5.2 (page 50) before with finite N in Remark 5.4 (page 51) and Remark 5.5 (page 51). We will further discuss the General Assumption 5.2 (page 50) in Lemma 5.6 (page 51) and Remark 5.7 (page 52).

The nomenclature used in this section is mainly defined in Definition 2.1 (page 18) and is then enlarged in Definition 5.8 (page 52). We then give the proof of Main Theorem 5.3 (page 50) in several steps. These are the necessary Lemma 5.10 (page 53), followed by the very detailed Theorem 5.11 (page 57), whereafter the latter parts are proved in Conclusion 5.12 (page 60), Conclusion 5.13 (page 61) and Remark 5.14 (page 61).

5.4 Remark: Limites

Consider a non-decreasing $\sigma \in \mathcal{S}^{\mathbb{N}_0}$. We have

$$\lim_{n \rightarrow \infty} X_{\sigma_n} = X_{\lim_{n \rightarrow \infty} \sigma_n} \text{ on } \{\sup \{\sigma_n ; n \in \mathbb{N}_0\} < \infty\}.$$

If $X_\infty = \lim_{n \rightarrow \infty} X_n$ then we have

$$\lim_{n \rightarrow \infty} X_{\sigma_n} = X_{\lim_{n \rightarrow \infty} \sigma_n}.$$

5.5 Remarks

- If N is finite, we set $X_{N+n} = X_N$ for all $n \in \mathbb{N}$ and $X_\infty = X_N$ (see Remark 2.20 (page 27)). Hence we have

$$X_\infty = \lim_{n \rightarrow \infty} X_n$$

and for all non-decreasing $\sigma \in \mathcal{S}^{\mathbb{N}_0}$ there is non-decreasing $\tau \in (\mathcal{S}_0^N)^{\mathbb{N}_0}$ with $X_{\sigma_n} = X_{\tau_n}$ for all $n \in \mathbb{N}$. Thus we have by Lemma 3.12 (page 38)

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_j) = \mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} | \mathcal{A}_j\right)$$

for all $j \in \mathbb{N}_0, \sigma \in \mathcal{S}^{\mathbb{N}_0}$ non-decreasing.

- This is not generally true for infinite N . Consider the following example:
For all $n \in \mathbb{N}_0$ let $X_n = n \mathbf{1}_{]0, \frac{1}{n}]}$. Then we have

$$X_\infty = \lim_{n \rightarrow \infty} X_n = 0.$$

Now consider $\sigma \in \mathcal{S}^{\mathbb{N}_0}$ with $\sigma_n = n$ for all $n \in \mathbb{N}_0$. Then we have for all $j \in \mathbb{N}_0$

$$\mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} | \mathcal{A}_j\right) = \mathbb{E}(X_\infty | \mathcal{A}_j) = 0 \neq 1 = \lim_{n \rightarrow \infty} \mathbb{E}(X_n | \mathcal{A}_j) = \lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_j).$$

5.6 Lemma

A sufficient condition for (5.2.1) and (5.2.2) is

$$(5.6.1) \quad \mathbb{E}\left(\sup_{n \in \mathbb{N}} |X_n|\right) < \infty \text{ and } X_\infty = \lim_{n \rightarrow \infty} X_n.$$

Proof: Assume (5.6.1). We have for each $\tau \in \mathcal{S}$

$$|X_\tau| \leq \sup_{n \in \mathbb{N}} |X_n|,$$

where the rhs is an integrable random variable, hence we have (5.2.1).

Now consider $j \in \mathbb{N}_0, \sigma \in \mathcal{S}^{\mathbb{N}_0}$ non-decreasing.

By Conditional Lebesgue Dominated Convergence we have

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_j) = \mathbb{E}\left(\lim_{n \rightarrow \infty} X_{\sigma_n} | \mathcal{A}_j\right)$$

and by Remark 5.4 (page 51) we have

$$\mathbb{E}\left(\lim_{n \rightarrow \infty} X_{\sigma_n} | \mathcal{A}_j\right) = \mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} | \mathcal{A}_j\right)$$

and thus (5.2.2) holds. □

5.7 Remarks

- If N is finite and embedded in the infinite case as written in Remark 2.20 (page 27), then we have (5.2.2) by Remark 5.5 (page 51) and (5.2.1) by Lemma 1.4 (page 15).
- The assumption (5.2.2) is equivalent to the following assumption:

$$\lim_{n \rightarrow \infty} \int_G X_{\sigma_n} dP = \int_G X_{\lim_{n \rightarrow \infty} \sigma_n} dP$$

for all $j \in \mathbb{N}_0$, $G \in \mathcal{A}_j$, $\sigma \in \mathcal{S}^{\mathbb{N}_0}$ non-decreasing.

- Due to the assumption (5.2.1) we have $\mathcal{S} = \mathcal{S}_0^*$ (see Definition 1.3 (page 14)).

5.8 Definitions

For all $C \in \mathcal{C}$ and $\sigma \in \mathcal{T}$ define

$$\tau_\sigma(C) := \inf \{k \geq \sigma ; \mathbf{1}_{C_k} = 1\}.$$

For all $n \in \mathbb{N}_0$ write $\tau_n(C)$ for $\tau_\sigma(C)$ with $\sigma \equiv n$.

For all $C \in \mathcal{C}$ and $\kappa \in \mathbb{N} \cup \{\infty\}$ define $C^{*\kappa}$ by

$$(5.8.1) \quad C_n^{*\kappa} := C_n \cap \bigcap_{j=1}^{\kappa} \left\{ X_n \geq \mathbb{E} \left(X_{\tau_{n+j}(C)} \middle| \mathcal{A}_n \right) \right\} \text{ for all } n \in \mathbb{N}_0, C_\infty^{*\kappa} := \Omega.$$

For all $C \in \mathcal{C}$, $\kappa \in \mathbb{N} \cup \{\infty\}$, $\rho \in \mathcal{S}(C)$ with $\sigma \leq \rho \leq \tau_\sigma(C^{*\kappa})$ define

$$\widehat{\rho}^\kappa := \rho \mathbf{1}_{\{\rho = \tau_\sigma(C^{*\kappa})\}} + \sum_{n=0}^{\infty} \sum_{j=1}^{\kappa} \mathbf{1}_{\{\rho = n\} \cap \{j = \min\{l \in \mathbb{N}; X_n < \mathbb{E}(X_{\tau_{n+l}(C)} | \mathcal{A}_n)\}\}} (\tau_{n+j}(C) \wedge \tau_\sigma(C^{*\kappa})).$$

If ambiguity is impossible, we do not write the index κ .

The improvement step in (5.8.1) coincides with that in Algorithm 2.13 (page 24) due to (2.16.4).

5.9 Lemma: Optimality

Consider $C \in \mathcal{C}$ with $\tau(C)$ being optimal in $\mathcal{S}(C)$. Then for each $\kappa \in \mathbb{N} \cup \{\infty\}$ we have $C = C^{*\kappa}$.

Proof: We have for all $j \in \mathbb{N}$ $\tau_{n+j}(C) \in \mathcal{S}(C) \cap \mathcal{S}_{n+j}^N \subseteq \mathcal{S}(C) \cap \mathcal{S}_n^N$ and hence by Definition 1.11 (page 16)

$$X_n = \mathbb{E}(X_{\tau_n(C)} | \mathcal{A}_n) = \mathbb{E}(X_{\tau_{n+j}(C)} | \mathcal{A}_n) \text{ on } C_n.$$

The assertion follows by the definition of $C^{*\kappa}$. □

We make adaptations of the ideas in [Irl80, Part 2, Theorem 2.1, Proof part (i), page 181] to conditional expectations and arbitrary κ . The idea for arbitrary κ , but with finite N , can be found in [KS06] and [BS06] and has been extended in Section 3 (page 28) where backward induction was the main tool of the proofs.

5.10 Lemma: One Step Improvement

Suppose $C \in \mathcal{C}$, $i \in \mathbb{N}_0$, $\kappa \in \mathbb{N} \cup \{\infty\}$, and $\sigma \in \mathcal{S}(C)$ with

$$(5.10.1) \quad \sigma \geq i.$$

For all

$$(5.10.2) \quad \rho \in \mathcal{S}(C) \text{ with } \sigma \leq \rho \leq \tau_\sigma(C^*)$$

we have

$$(5.10.3) \quad \hat{\rho} \in \mathcal{S}(C) \text{ with } \sigma \leq \hat{\rho} \leq \tau_\sigma(C^*),$$

$$(5.10.4) \quad \rho \leq \hat{\rho},$$

$$(5.10.5) \quad \rho + 1 \leq \hat{\rho} \text{ on } \{\rho < \tau_\sigma(C^*)\},$$

$$(5.10.6) \quad \mathbb{E}(X_\rho | \mathcal{A}_i) \leq \mathbb{E}(X_{\hat{\rho}} | \mathcal{A}_i),$$

and using assumption (5.2.2) (page 50) also

$$(5.10.7) \quad \mathbb{E}(X_\rho | \mathcal{A}_i) \leq \mathbb{E}(X_{\tau_\sigma(C^*)} | \mathcal{A}_i).$$

Remark: If $\kappa = 1$, then

$$(5.10.8) \quad C_n^* = C_n \cap \{X_n \geq \mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n)\} \text{ for all } n \in \mathbb{N}_0,$$

$$\hat{\rho} = \rho \mathbf{1}_{\{\rho = \tau_\sigma(C^*)\}} + \sum_{n=i}^{\infty} \mathbf{1}_{\{\rho = n\} \cap \{\mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n) > X_n\}} \tau_{n+1}(C).$$

We will have much shorter proofs for $\kappa = 1$ than for arbitrary κ . We included both proofs for comparance in the sequel.

Proof of (5.10.7) using (5.10.3) to (5.10.6):

Define $\sigma_0 := \rho$ and inductively

$$\sigma_{k+1} := \hat{\sigma}_k \text{ for all } k \in \mathbb{N}_0.$$

Now we will prove by induction that

$$(5.10.9) \quad \rho \leq \sigma_k \leq \sigma_{k+1} \leq \tau_\sigma(C^*) \text{ for all } k \in \mathbb{N}_0$$

We have $\sigma_0 = \rho \leq \tau_\sigma(C^*)$, hence we have by (5.10.3) and (5.10.4) $\sigma_0 \leq \sigma_1 \leq \tau_\sigma(C^*)$.

Consider $k \in \mathbb{N}_0$ with $\sigma_k \leq \tau_\sigma(C^*)$. It follows $\sigma_k \leq \sigma_{k+1} \leq \tau_\sigma(C^*)$ by the same equations.

We have (by induction) for all $k \in \mathbb{N}$

$$(5.10.10) \quad \mathbb{E}(X_{\sigma_0} | \mathcal{A}_i) \stackrel{(5.10.6)}{\leq} \mathbb{E}(X_{\sigma_k} | \mathcal{A}_i).$$

By (5.10.9) and (5.10.5) we have

$$\lim_{k \rightarrow \infty} \sigma_k \leq \tau_\sigma(C^*)$$

and

$$\sigma_k + 1 \leq \sigma_{k+1} \text{ on } \{\sigma_k < \tau_\sigma(C^*)\} \text{ for all } k \in \mathbb{N}.$$

So we have

$$(5.10.11) \quad \sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \dots \leq \lim_{k \rightarrow \infty} \sigma_k = \tau_\sigma(C^*).$$

It follows

$$\begin{aligned} \mathbb{E}(X_\rho | \mathcal{A}_i) &= \mathbb{E}(X_{\sigma_0} | \mathcal{A}_i) && \text{by def.} \\ &\leq \lim_{k \rightarrow \infty} \mathbb{E}(X_{\sigma_k} | \mathcal{A}_i) && \text{by (5.10.10)} \\ &= \mathbb{E}\left(X_{\lim_{k \rightarrow \infty} \sigma_k} \middle| \mathcal{A}_i\right) && \text{by (5.2.2) (page 50)} \\ &= \mathbb{E}(X_{\tau_\sigma(C^*)} | \mathcal{A}_i) && \text{by (5.10.11).} \quad \square \end{aligned}$$

Proof of (5.10.3) to (5.10.6) for $\kappa = 1$:

(This is a special version for $\kappa = 1$ of the arbitrary case below and due to this much shorter.)
Obviously we have

$$(5.10.12) \quad \rho = \hat{\rho} = \tau_\sigma(C^*) \text{ on } \{\rho = \tau_\sigma(C^*)\} \stackrel{(5.10.2)}{=} \{\rho = \infty\} \cup \bigcup_{n=i}^{\infty} \{\rho = n\} \cap C_n^*.$$

For all $n \in \mathbb{N}_0$ with $i \leq n$ we have

$$\begin{aligned} \rho < \rho + 1 = n + 1 &\leq \overbrace{\tau_{n+1}(C)}^{=\hat{\rho}} \leq \tau_\sigma(C^*) \\ \text{on } \{\rho = n\} &\cap \{\mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n) > X_n\}. \end{aligned}$$

Furthermore we have (where \uplus is here and in the following used for disjoint unions)

$$(5.10.13) \quad \{\rho < \tau_\sigma(C^*)\} \stackrel{(5.10.12)}{=} \biguplus_{n=i}^{\infty} \{\rho = n\} \cap \{\mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n) > X_n\}.$$

For (5.10.6) it suffices to prove for all $G \in \mathcal{A}_i$

$$\int_G X_\rho dP \leq \int_G X_{\hat{\rho}} dP.$$

So consider $G \in \mathcal{A}_i$.

Now we will split G in subsets and prove the inequality above for each of these subsets.

Thus let us define for all $n \in \mathbb{N}_0$ with $i \leq n$

$$(5.10.14) \quad A^{n,1} := G \cap \{\rho = n\} \cap \{\mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n) > X_n\}.$$

Let $n \in \mathbb{N}_0$ with $i \leq n$.

We have $A^{n,1} \in \mathcal{A}_i \subseteq \mathcal{A}_n$ and

$$(5.10.15) \quad \int_{A^{n,1}} X_\rho dP = \int_{A^{n,1}} X_n dP \leq \int_{A^{n,1}} \mathbb{E}(X_{\tau_{n+1}(C)} | \mathcal{A}_n) dP = \int_{A^{n,1}} X_{\tau_{n+1}(C)} dP = \int_{A^{n,1}} X_{\hat{\rho}} dP.$$

(5.10.13) shows that

$$G \cap \{\rho < \tau_\sigma(C^*)\} = \biguplus_{n=i}^{\infty} A^{n,1}.$$

Hence we have by (5.10.15) and by dominated convergence

$$\int_{G \cap \{\rho < \tau_\sigma(C^*)\}} X_\rho dP \leq \int_{G \cap \{\rho < \tau_\sigma(C^*)\}} X_{\hat{\rho}} dP.$$

We have by (5.10.12)

$$\rho = \hat{\rho} \text{ on } G \cap \{\rho = \tau_\sigma(C^*)\},$$

hence

$$\int_{G \cap \{\rho = \tau_\sigma(C^*)\}} X_\rho dP = \int_{G \cap \{\rho = \tau_\sigma(C^*)\}} X_{\hat{\rho}} dP.$$

So we have

$$\int_G X_\rho dP \leq \int_G X_{\hat{\rho}} dP. \quad \square$$

Proof of (5.10.3) to (5.10.5) for arbitrary κ :

Obviously we have

$$(5.10.16) \quad \rho = \hat{\rho} = \tau_\sigma(C^*) \text{ on } \{\rho = \tau_\sigma(C^*)\} \stackrel{(5.10.2)}{=} \{\rho = \infty\} \cup \bigcup_{n=i}^{\infty} \{\rho = n\} \cap C_n^*.$$

For all $n \in \mathbb{N}_0$ and for all $j \in \mathbb{N}$ with $j \leq \kappa$ we have

$$(5.10.17) \quad \text{on } \{\rho = n\} \cap \{j = \min \{l \in \mathbb{N}; X_n < \mathbb{E}(X_{\tau_{n+l}(C)} | \mathcal{A}_n)\}\} \\ \rho < \rho + 1 = n + 1 \leq \underbrace{(\tau_{n+j}(C) \wedge \tau_\sigma(C^*))}_{=\hat{\rho}} \leq \tau_\sigma(C^*).$$

Furthermore we have

$$(5.10.18) \quad C_n \setminus C_n^* = \bigoplus_{j=1}^{\kappa} \{j = \min \{l \in \mathbb{N}; X_n < \mathbb{E}(X_{\tau_{n+l}(C)} | \mathcal{A}_n)\}\} \text{ for all } n \in \mathbb{N}_0,$$

hence

$$(5.10.19) \quad \{\rho < \tau_\sigma(C^*)\} \\ \stackrel{(5.10.16)}{=} \bigoplus_{n=i}^{\infty} \{\rho = n\} \setminus C_n^*$$

$$(5.10.20) \stackrel{(5.10.18)}{=} \bigoplus_{n=i}^{\infty} \bigoplus_{j=1}^{\kappa} \{\rho = n\} \cap \{j = \min \{l \in \mathbb{N}; X_n < \mathbb{E}(X_{\tau_{n+l}(C)} | \mathcal{A}_n)\}\}. \quad \square$$

Proof of (5.10.6) for arbitrary κ :

It suffices to prove for all $G \in \mathcal{A}_i$

$$\int_G X_\rho dP \leq \int_G X_{\hat{\rho}} dP.$$

So consider $G \in \mathcal{A}_i$.

Now we will split G in subsets and prove the inequality above for each of these subsets.

Thus let us define for all $n \in \mathbb{N}_0$ with $i \leq n$, $j \in \mathbb{N}$ with $j \leq \kappa$

$$(5.10.21) \quad A^{n,j} := G \cap \{\rho = n\} \cap \{j = \min \{l \in \mathbb{N}; X_n < \mathbb{E}(X_{\tau_{n+l}(C)} | \mathcal{A}_n)\}\}.$$

Let $n \in \mathbb{N}_0$ with $i \leq n$, $j \in \mathbb{N}$ with $j \leq \kappa$.

We have $G \in \mathcal{A}_i \subseteq \mathcal{A}_n$ and hence $A^{n,j} \in \mathcal{A}_n$.

Since

$$(5.10.22) \quad A^{n,j} \stackrel{(5.10.21)}{\subseteq} \{X_n < \mathbb{E}(X_{\tau_{n+j}(C)} | \mathcal{A}_n)\}$$

we have

$$(5.10.23) \quad \int_{A^{n,j}} X_\rho dP \stackrel{(5.10.21)}{=} \int_{A^{n,j}} X_n dP \\ \stackrel{(5.10.22)}{\leq} \int_{A^{n,j}} \mathbb{E}(X_{\tau_{n+j}(C)} | \mathcal{A}_n) dP \\ \stackrel{A^{n,j} \in \mathcal{A}_n}{=} \int_{A^{n,j}} X_{\tau_{n+j}(C)} dP.$$

Now we will define a disjoint decomposition of $A^{n,j}$. Thus we define

$$D := \{\tau_{n+j}(C) \leq \tau_\sigma(C^*)\}$$

and for all $m \in \mathbb{N}$ with $m < j$ define

$$A_m := A^{n,j} \cap \left(C_{n+m}^* \setminus \bigcup_{l=1}^{m-1} C_{n+l}^* \right).$$

Since $A^{n,j} \in \mathcal{A}_n$ we have

$$(5.10.24) \quad A_m \in \mathcal{A}_{m+n} \text{ for all } m \in \mathbb{N} \text{ with } m < j.$$

We will prove now

$$A^{n,j} \cap D^c = \bigsqcup_{m=1}^{j-1} A_m.$$

Disjoint: For all $m, l \in \mathbb{N}$ with $m, l < j$, $m \neq l$ and (w.l.o.g.) $m \leq l$ we have $m \leq l - 1$ and thus

$$A_m \cap A_l \subseteq C_{n+m}^* \setminus C_{n+m}^* = \emptyset,$$

hence for all $m, l \in \mathbb{N}$ with $m, l < j$

$$m \neq l \implies A_m \cap A_l = \emptyset.$$

“ \supseteq ”: For all $m \in \mathbb{N}$ with $m < j$ we have

$$A_m \subseteq D^c.$$

“ \subseteq ”: Consider $\omega \in A^{n,j} \cap D^c$. Define

$$r := \tau_\sigma(C^*)(\omega) = \inf \{l \geq \sigma(\omega) ; \omega \in C_l^*\}.$$

We have $r = \tau_\sigma(C^*)(\omega) \stackrel{(5.10.2)}{\geq} \rho(\omega) = n$.

Due to the definition of $A^{n,j}$ in (5.10.21) we have by (5.10.18) $\omega \in C_n \setminus C_n^*$, hence $r \neq n$.

Define

$$q := \tau_{n+j}(C)(\omega) = \inf \{l \geq n+j ; \omega \in C_l\}.$$

For all $p \in \mathbb{N}_0$ with $n+j \leq p < q$ we have $\omega \notin C_l \supseteq C_l^*$.

Since $\omega \in D^c = \{\tau_{n+j}(C) > \tau_\sigma(C^*)\}$ we have $q > r$.

This yields $r < n+j$.

So we have $n < r < n+j$, thus $1 \leq r - n \leq j - 1$ and it follows $\omega \in A_{r-n}$ by definition of r .

By the above part of the proof we can deduce for all $m \in \mathbb{N}$ with $m < j$

$$(5.10.25) \quad \tau_\sigma(C^*) = m + n \text{ on } A_m = A_{(m+n)-n}.$$

It follows

$$(5.10.26) \quad A^{n,j} = (A^{n,j} \cap D) \uplus \bigsqcup_{m=1}^{j-1} A_m.$$

We have

$$(5.10.27) \quad \tau_{n+j}(C) = \tau_{n+j}(C) \wedge \tau_\sigma(C^*) = \hat{\rho} \text{ on } A^{n,j} \cap D.$$

For all $m \in \mathbb{N}$ with $m < j$ we have

$$\begin{aligned}
\int_{A_m} X_{\tau_{n+j}(C)} dP &\stackrel{(5.10.24)}{=} \int_{A_m} \mathbb{E} \left(X_{\tau_{n+j}(C)} \middle| \mathcal{A}_{n+m} \right) dP \\
&\stackrel{(5.8.1), j-m \leq \kappa}{\leq} \int_{A_m} X_{n+m} dP \\
&\stackrel{(5.10.25)}{=} \int_{A_m} X_{\tau_\sigma(C^*)} dP \\
(5.10.28) \quad &\stackrel{A_m \subseteq D^c}{=} \int_{A_m} X_{\hat{\rho}} dP.
\end{aligned}$$

By (5.10.26), (5.10.27) and (5.10.28) and summing over m we have

$$\int_{A^{n,j}} X_{\tau_{n+j}(C)} dP = \int_{A^{n,j}} X_{\hat{\rho}} dP.$$

Now we can combine this with (5.10.23) and so we have

$$\begin{aligned}
\int_{A^{n,j}} X_{\rho} dP &\stackrel{(5.10.23)}{\leq} \int_{A^{n,j}} X_{\tau_{n+j}(C)} dP \\
(5.10.29) \quad &= \int_{A^{n,j}} X_{\hat{\rho}} dP.
\end{aligned}$$

(5.10.20) (page 55) shows that

$$G \cap \{\rho < \tau_\sigma(C^*)\} = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{\kappa} A^{n,j}.$$

Hence we have by (5.10.29) and by dominated convergence

$$\int_{G \cap \{\rho < \tau_\sigma(C^*)\}} X_{\rho} dP \leq \int_{G \cap \{\rho < \tau_\sigma(C^*)\}} X_{\hat{\rho}} dP.$$

We have by (5.10.16) (page 55)

$$\rho = \hat{\rho} \text{ on } G \cap \{\rho = \tau_\sigma(C^*)\}.$$

hence

$$\int_{G \cap \{\rho = \tau_\sigma(C^*)\}} X_{\rho} dP = \int_{G \cap \{\rho = \tau_\sigma(C^*)\}} X_{\hat{\rho}} dP.$$

So we have

$$\int_G X_{\rho} dP \leq \int_G X_{\hat{\rho}} dP. \quad \square$$

5.11 Theorem

$\tau^{(\infty)}$ is optimal in $\mathcal{S}(C^{(0)})$

and for all $m, j \in \mathbb{N}_0$ we have

$$\begin{aligned}
\tau_j^{(m)} &\leq \tau_j^{(m+1)}, \\
Y_j^{(m)} &\leq Y_j^{(m+1)}.
\end{aligned}$$

Proof: Consider $j \in \mathbb{N}_0$.

Step 1:

Consider $(C^{(m)})_{m \in \mathbb{N}_0}$ as constructed above.

Let $\rho \in \mathcal{S}(C^{(0)})$ with $\rho \geq j$.

Define $\rho' := \inf \left\{ n ; \rho \leq n, \mathbb{1}_{C_n^{(\infty)}} = 1 \right\}$, thus $\rho' \in \mathcal{S}(C^{(\infty)})$.

For all $m \in \mathbb{N}_0$ define $\rho_m := \inf \left\{ n ; \rho \leq n, \mathbb{1}_{C_n^{(m)}} = 1 \right\}$.

We have $\rho_0 = \inf \left\{ n ; \rho \leq n, \mathbb{1}_{C_n^{(0)}} = 1 \right\} = \inf \{ n ; \rho \leq n \} = \rho$,

because for all $n \in \mathbb{N}_0$ we have $\{ \rho = n \} \subseteq C_n^{(0)}$.

Consider $m \in \mathbb{N}_0$.

For all $n \in \mathbb{N}_0 \cup \{ \infty \}$ we have $C_n^{(m+1)} \subseteq C_n^{(m)}$ and thus

$$\left\{ \mathbb{1}_{C_n^{(m+1)}} = 1 \right\} \subseteq \left\{ \mathbb{1}_{C_n^{(m)}} = 1 \right\}.$$

So we have

$$\rho_{m+1} = \inf \left\{ n ; \rho \leq n, \mathbb{1}_{C_n^{(m+1)}} = 1 \right\} \geq \inf \left\{ n ; \rho \leq n, \mathbb{1}_{C_n^{(m)}} = 1 \right\} = \rho_m.$$

It follows by Lemma 5.10 (page 53), equation (5.10.7)

(adaption: $i = j$, $\sigma = \rho_m$, $\sigma^* = \rho_{m+1}$, $C = C^{(m)}$, $C^* = C^{(m+1)}$)

$$\mathbb{E}(X_{\rho_m} | \mathcal{A}_j) \leq \mathbb{E}(X_{\rho_{m+1}} | \mathcal{A}_j).$$

The facts $\rho_m \leq \rho_{m+1}$ for all $m \in \mathbb{N}_0$ and $\lim \rho_m = \rho'$ imply

$$\mathbb{E}(X_\rho | \mathcal{A}_j) \leq \lim_{m \rightarrow \infty} \mathbb{E}(X_{\rho_m} | \mathcal{A}_j) \stackrel{(5.2.2)}{=} \mathbb{E}\left(X_{\lim_{m \rightarrow \infty} \rho_m} | \mathcal{A}_j\right) = \mathbb{E}(X_{\rho'} | \mathcal{A}_j).$$

So for every $\rho \in \mathcal{S}(C^{(0)})$ with $\rho \geq j$ there exists $\rho' \in \mathcal{S}(C^{(\infty)})$ with $\mathbb{E}(X_\rho | \mathcal{A}_j) \leq \mathbb{E}(X_{\rho'} | \mathcal{A}_j)$.

For all $n \in \mathbb{N}_0 \cup \{ \infty \}$ we have $C_n^{(\infty)} \subseteq C_n^{(0)}$ and thus $\mathcal{S}(C^{(\infty)}) \subseteq \mathcal{S}(C^{(0)})$.

The last two expressions yield:

$$\text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_j) ; \rho \in \mathcal{S}(C^{(\infty)}), \rho \geq j \right\} = \text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_j) ; \rho \in \mathcal{S}(C^{(0)}), \rho \geq j \right\}.$$

Step 2A:

Consider $\rho, \tau \in \mathcal{S}(C^{(\infty)})$ with $j \leq \rho \leq \tau$.

For all $m \in \mathbb{N}_0$ define

$$\rho_m := \tau \mathbb{1}_{\{\rho = \tau\}} + \sum_{n=j}^{\infty} \tau_{n+1}^{(m)} \mathbb{1}_{\{\rho = n\}} \mathbb{1}_{\{\rho < \tau\}}.$$

Define further

$$\tilde{\rho} := \tau \mathbb{1}_{\{\rho = \tau\}} + \sum_{n=j}^{\infty} \tau_{n+1}^{(\infty)} \mathbb{1}_{\{\rho = n\}} \mathbb{1}_{\{\rho < \tau\}}.$$

We have $\rho \leq \rho_m \leq \rho_{m+1}$ for all $m \in \mathbb{N}_0$ and $\lim_{m \rightarrow \infty} \rho_m = \tilde{\rho} \leq \tau$.

Consider $G \in \mathcal{A}_j$, $m \in \mathbb{N}_0$, $n \in \mathbb{N}_0$ with $j \leq n$.

We have

$$\begin{aligned}
G \cap \{\rho = n\} \cap \{n < \tau\} &= G \cap \{\rho = n\} \cap \{\rho < \tau\} \\
&\subseteq \{\rho = n\} \\
&\stackrel{\rho \in \mathcal{S}(C^{(\infty)})}{\subseteq} C_n^{(\infty)} \\
&\stackrel{2.13}{\subseteq} C_n^{(m+1)} \\
&\stackrel{2.13}{\subseteq} \left\{ \mathbb{E} \left(X_{\tau_{n+1}^{(m)}} \middle| \mathcal{A}_n \right) \leq X_n \right\}
\end{aligned}$$

and $G \cap \{\rho = n\} \cap \{n < \tau\} \in \mathcal{A}_n$, because $G \in \mathcal{A}_j \subseteq \mathcal{A}_n$ and $\{\rho = n\}, \{n < \tau\} \in \mathcal{A}_n$.

Thus

$$\int_{G \cap \{\rho = n\} \cap \{n < \tau\}} X_{\rho_m} dP = \int_{G \cap \{\rho = n\} \cap \{n < \tau\}} X_{\tau_{n+1}^{(m)}} dP \leq \int_{G \cap \{\rho = n\} \cap \{n < \tau\}} X_{\rho} dP.$$

So we have

$$\begin{aligned}
\int_G X_{\rho_m} dP &= \int_{G \cap \{\rho = \tau\}} X_{\rho_m} dP + \sum_{n \in \mathbb{N}_0} \int_{G \cap \{\rho = n\} \cap \{n < \tau\}} X_{\rho_m} dP \\
&\leq \int_{G \cap \{\rho = \tau\}} X_{\rho} dP + \sum_{n \in \mathbb{N}_0} \int_{G \cap \{\rho = n\} \cap \{n < \tau\}} X_{\rho} dP \\
&= \int_G X_{\rho} dP.
\end{aligned}$$

Combining these facts yields

$$\mathbb{E}(X_{\tilde{\rho}} | \mathcal{A}_j) = \mathbb{E} \left(X_{\lim_{m \rightarrow \infty} \rho_m} \middle| \mathcal{A}_j \right) \stackrel{(5.2.2)}{=} \lim_{m \rightarrow \infty} \mathbb{E}(X_{\rho_m} | \mathcal{A}_j) \leq \mathbb{E}(X_{\rho} | \mathcal{A}_j)$$

Step 2B:

Consider $\rho, \tau \in \mathcal{S}(C^{(\infty)})$ with $j \leq \rho \leq \tau$.

Define $\rho_0 := \rho$ and for all $m \in \mathbb{N}_0$ define $\rho_{m+1} := \tilde{\rho}_m$.

For all $m \in \mathbb{N}_0$ we have

$$\begin{aligned}
\rho_m &\leq \rho_{m+1} \leq \tau, \\
\rho_m + 1 &\leq \tilde{\rho}_m = \rho_{m+1} \text{ on } \{\rho_m < \tau\}.
\end{aligned}$$

It follows

$$\lim_{m \rightarrow \infty} \rho_m = \tau.$$

So for all $G \in \mathcal{A}_j$ we have

$$\mathbb{E}(X_{\tau} | \mathcal{A}_j) = \mathbb{E} \left(X_{\lim_{m \rightarrow \infty} \rho_m} \middle| \mathcal{A}_j \right) \stackrel{(5.2.2)}{=} \lim_{m \rightarrow \infty} \mathbb{E}(X_{\rho_m} | \mathcal{A}_j) \leq \mathbb{E}(X_{\rho} | \mathcal{A}_j).$$

Step 2C:

Let $\sigma \in \mathcal{S}(C^{(\infty)})$ with $\sigma \geq j$. Let $\omega \in \Omega$, $l := \sigma(\omega)$. Then $\omega \in C_l^{(\infty)}$ and $1 = \mathbf{1}_{C_l^{(\infty)}}(\omega)$, thus

$$\tau_j^{(\infty)}(\omega) = \inf \left\{ p \geq j ; \mathbf{1}_{C_p^{(\infty)}}(\omega) = 1 \right\} \leq l = \sigma(\omega).$$

This yields $\tau_j^{(\infty)} \leq \sigma$.

So for all $\tau \in \mathcal{S}(C^{(\infty)})$ with $\tau \geq j$ we have $\tau \geq \tau_j^{(\infty)}$, thus

$$\mathbb{E}(X_{\tau} | \mathcal{A}_j) \leq \mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right).$$

This yields

$$\mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right) = \text{esssup} \left\{ \mathbb{E} (X_\tau | \mathcal{A}_j) ; \tau \in \mathcal{S}(C^{(\infty)}), \tau \geq j \right\}.$$

The above steps yield

$$\mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right) = \text{esssup} \left\{ \mathbb{E} (X_\rho | \mathcal{A}_j) ; \rho \in \mathcal{S}(C^{(0)}), \rho \geq j \right\}.$$

So for any family $C^{(0)} \in \mathcal{C}$ we have

$$\tau_j^{(\infty)} \in \mathcal{S}(C^{(0)}) \quad \text{for all } j \in \mathbb{N}_0. \quad \square$$

5.12 Conclusion

We have for all $j \in \mathbb{N}_0$

$$Y_j^{(\infty)} = \mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right).$$

If there is some $k \in \mathbb{N}_0$ such that

$$\tau_k \text{ is optimal in } \mathcal{S}(C^{(k)})$$

or

$$C^{(k)} = C^{(k+1)}$$

then we have

$$\begin{aligned} C^{(k)} &= C^{(k+n)} = C^{(\infty)} \quad \text{for all } n \in \mathbb{N}_0, \\ \tau^{(k)} &= \tau^{(k+n)} = \tau^{(\infty)} \quad \text{for all } n \in \mathbb{N}_0, \\ \Theta^{(k)} &= \Theta^{(k+n)} = \Theta^{(\infty)} \quad \text{for all } n \in \mathbb{N}_0, \\ Y^{(k)} &= Y^{(k+n)} = Y^{(\infty)} \quad \text{for all } n \in \mathbb{N}_0. \end{aligned}$$

Proof: For all $j \in \mathbb{N}_0$ we have

$$\begin{aligned} Y_j^{(\infty)} &\stackrel{\text{def.}}{=} \lim_{\text{incr. } m \rightarrow \infty} Y_j^{(m)} \stackrel{\text{def.}}{=} \lim_{m \rightarrow \infty} \mathbb{E} \left(X_{\tau_j^{(m)}} \middle| \mathcal{A}_j \right) \\ &\stackrel{\text{ass.}}{=} \mathbb{E} \left(X_{\left(\lim_{m \rightarrow \infty} \tau_j^{(m)} \right)} \middle| \mathcal{A}_j \right) \stackrel{\text{def.}}{=} \mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right). \end{aligned}$$

The other statement is true by Lemma 5.9 (page 52) and the definitions in Part 2.13 (page 24). \square

5.13 Conclusion

If $C^{(0)}$ is essential, then

$\tau^{(\infty)}$ is optimal.

Proof: Assume $C^{(0)}$ is essential. Then there is some optimal $\sigma \in \mathcal{T}$ with $\sigma_n \in \mathcal{S}(C^{(0)})$ for all $n \in \mathbb{N}_0$. So we have for all $n \in \mathbb{N}_0$

$$\begin{aligned} \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n) &\leq \text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_n) ; \rho \in \mathcal{S}_n^\infty \cap \mathcal{S}(C^{(0)}) \right\} \\ &\leq \text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_n) ; \rho \in \mathcal{S}_n^\infty \right\} \\ &= \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n). \end{aligned}$$

By Theorem 5.11 (page 57)

$\tau^{(\infty)}$ is optimal in $\mathcal{S}(C^{(0)})$,

hence for all $n \in \mathbb{N}_0$ we have

$$\mathbb{E} \left(X_{\tau_n^{(\infty)}} \middle| \mathcal{A}_n \right) = \text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_n) ; \rho \in \mathcal{S}_n^\infty \cap \mathcal{S}(C^{(0)}) \right\}.$$

Hence we have for all $n \in \mathbb{N}_0$

$$\mathbb{E} \left(X_{\tau_n^{(\infty)}} \middle| \mathcal{A}_n \right) = \mathbb{E}(X_{\sigma_n} | \mathcal{A}_n),$$

and

$\tau^{(\infty)}$ is optimal. □

5.14 Remark

If $C^{(0)} \equiv \Omega$, then

$\tau^{(\infty)}$ is optimal.

Proof: Consider $C^{(0)} \equiv \Omega$. There are two ways to prove the optimality:

(1) We have $\mathcal{S} = \mathcal{S}(C^{(0)})$ and hence for all $j \in \mathbb{N}_0$ we have

$$\mathbb{E} \left(X_{\tau_j^{(\infty)}} \middle| \mathcal{A}_j \right) = \text{esssup} \left\{ \mathbb{E}(X_\rho | \mathcal{A}_j) ; \rho \in \mathcal{S}, \rho \geq j \right\}.$$

(2) $C^{(0)}$ is essential, hence $\tau^{(\infty)}$ is optimal by Conclusion 5.13 (page 61). □

5.15 Outlook

The above theorems and proofs are adapted to the Markovian case with random discounting in Section 7 (page 67). The Markovian case itself is introduced in the following Section 6 (page 62).

6 Markovian Case

In this chapter we adapt the FII Algorithm stated in Section 2 (page 18) to Markovian stopping problems.

After repeating the setting in Part 6.1 (page 62), we consider the case of time-independent pay-off for the case of arbitrary κ in Subsection 6.2 (page 62). For $\kappa \equiv 1$ this was done before in [Irl06, Part 2] and [Irl09, Part 1]. Some cases of time-dependent pay-off for $\kappa \equiv 1$ have been mentioned in [Irl09, p. 2] and in detail considered in [Irl06, part 3.1]. Here we will show their extensions to arbitrary κ in Subsection 6.3 (page 64). Part 6.4 (page 65) shows that the case of certain path-dependent pay-offs, considered in [Pre10], fits in our setting as well.

6.1 Setting

Let $(Z_n)_{n \in \mathbb{N}_0}$ be a time-homogeneous Markov process with respect to the underlying filtration with state space (S, \mathcal{S}) . In the Markovian case the FII algorithm only has to work in the systems of subsets of the state space S , as we will show exemplary for the case of time-independent pay-off in Part 6.2.1 (page 63).

6.2 Time-Independent Pay-Off

For all $B \in \mathcal{S}$ define

$$\begin{aligned} \mathcal{T}(B) &:= \{\tau \in \mathcal{T} ; Z_\tau \in B \text{ on } \{\tau < \infty\}\}, \\ \tau_n(B) &:= \inf \{k \geq n : Z_k \in B\} \text{ for all } n \in \mathbb{N}_0. \end{aligned}$$

Given a measurable mapping $g : S \rightarrow \mathbb{R}$ consider the problem of optimal stopping for the pay-off

$$X_n := g(Z_n) \text{ for all } n \in \mathbb{N}_0,$$

defining for simpler notation

$$X_\infty := g(Z_\infty) := \limsup_{n \rightarrow \infty} \sup_{k \geq n} g(Z_k),$$

so that we have

$$X_\tau = g(Z_\tau) \text{ for all } \tau \in \mathcal{T}.$$

Assume as usual

$$\mathbb{E}(X_\tau | Z_0 = z) \text{ exists for all } \tau \in \mathcal{T} \text{ and } z \in S.$$

Then the iterative step is performed for given κ by

$$B^{*\kappa} := \left\{ z \in B ; g(z) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right) \right\} \text{ for all } B \in \mathcal{S}.$$

6.2.1 Connection to the General Case

Consider N infinite and some $n \in \mathbb{N}_0$ with $C^{(n)}$ given such that there is some B with

$$C_k^{(n)} = \{Z_k \in B\} \text{ for all } k \in \mathbb{N}_0.$$

Then we have

$$C_k^{(n+1)} = \left\{ Z_k \in B^{*\kappa(n)} \right\} \text{ for all } k \in \mathbb{N}_0.$$

Therefore the FII algorithm only has to work in the systems of subsets of the state space S .

Proof: We have $\tau_k^{(n)} = \inf \{p \geq k ; Z_p \in B\} = \tau_k(B)$ for all $k \in \mathbb{N}_0$.
Assume $(Z'_n)_{n \in \mathbb{N}_0}$ being an independent copy of $(Z_n)_{n \in \mathbb{N}_0}$ and

$$\tau'_p{}^{(n)} = \tau'_p(B) = \inf \{q \geq p ; Z'_q \in B\} \text{ for all } p \in \mathbb{N}_0.$$

We have for all $p, k \in \mathbb{N}_0$ with $p \leq k$

$$\begin{aligned} & \left\{ \mathbb{E} \left(g(Z_{\tau_p^{(n)}}) \middle| \mathcal{A}_k \right) \leq g(Z_k) \right\} \\ &= \left\{ Z_k \in \left\{ \mathbb{E} \left(g(Z'_{\tau'_p{}^{(n)}}) \middle| Z'_k = z \right) \leq g(z) \right\} \right\} \\ &= \left\{ Z_k \in \left\{ \mathbb{E} \left(g(Z'_{\tau'_{p-k}{}^{(n)}}) \middle| Z'_0 = z \right) \leq g(z) \right\} \right\} \\ &= \left\{ Z_k \in \left\{ \mathbb{E} \left(g(Z'_{\tau'_{p-k}(B)}) \middle| Z'_0 = z \right) \leq g(z) \right\} \right\}. \end{aligned}$$

Hence we have

$$\begin{aligned} C_k^{(n+1)} &= C_k^{(n)} \cap \bigcap \left\{ \left\{ \mathbb{E} \left(X_{\tau_p^{(n)}} \middle| \mathcal{A}_k \right) \leq X_k \right\} ; k \leq p \leq k + \kappa(n) \right\} \\ &= \{Z_k \in B\} \cap \bigcap \left\{ Z_k \in \left\{ g(z) \geq \mathbb{E} \left(g(Z'_{\tau'_p{}^{(n)}}) \middle| Z'_0 = z \right) \right\} ; 0 \leq p \leq \kappa(n) \right\} \\ &= \{Z_k \in B\} \cap \left\{ Z_k \in \left\{ g(z) \geq \sup \left\{ \mathbb{E} \left(g(Z'_{\tau'_p{}^{(n)}}) \middle| Z'_0 = z \right) ; 0 \leq p \leq \kappa(n) \right\} \right\} \right\} \\ &= \left\{ Z_k \in \left\{ z \in B ; g(z) \geq \sup \left\{ \mathbb{E} \left(g(Z'_{\tau'_p{}^{(n)}}) \middle| Z'_0 = z \right) ; 0 \leq p \leq \kappa(n) \right\} \right\} \right\} \\ &= \left\{ Z_k \in B^{*\kappa(n)} \right\}. \end{aligned} \quad \square$$

6.3 Time-Dependent Pay-Off

We can apply the algorithm and its results to the “time-independent” space-time Markov process $(Z_n, n)_{n \in \mathbb{N}_0}$ on $S \times \mathbb{N}_0$ instead of S . Hence the time-dependent case can be treated as follows:

Consider $f : S \times \mathbb{N}_0 \rightarrow \mathbb{R}$ and the problem of optimal stopping for the pay-off

$$X_n := f(Z_n, n) \text{ for all } n \in \mathbb{N}_0,$$

defining (as above) for simpler notation

$$X_\infty := f(Z_\infty, \infty) := \lim_{n \rightarrow \infty} \sup_{k \geq n} f(Z_k, k),$$

so that we have

$$X_\tau = f(Z_\tau, \tau) \text{ for all stopping rules } \tau.$$

Assume as usual

$$\mathbb{E}(X_\tau | Z_0 = z) \text{ exists for all } \tau \in \mathcal{T} \text{ and } z \in S.$$

Then the iterative step is done for any κ and any measurable $B \subseteq S \times \mathbb{N}_0$ by

$$B^{*\kappa} := \left\{ (z, n) \in B ; f(z, n) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(f \left(Z_{\tau_j(B)}, n + \tau_j(B) \right) \middle| Z_0 = z \right) \right\}$$

with

$$\tau_n(B) := \inf \{ k \geq n ; (Z_k, k) \in B \} \text{ for all } n \in \mathbb{N}_0.$$

6.3.1 Finite Horizon

In the case of a stopping problem with finite time horizon $N \in \mathbb{N}$ and pay-off

$$f(Z_n, n) \text{ for all } n \in \mathbb{N}_0 \text{ with } n \leq N$$

just define

$$f(z, k) = -\infty \text{ for all } z \in S \text{ and for all } k \in \mathbb{N}_0 \text{ with } N + 1 \leq k.$$

Then for any κ and any measurable $B \subseteq S \times \mathbb{N}_0$ we have $B^{*\kappa} \subseteq S \times \{0, \dots, N\}$.

6.3.2 Linear Costs

Consider some measurable mapping $g : S \rightarrow \mathbb{R}$, some $c \in \mathbb{R}_{>0}$ (often representing the costs of observation for one time step) and

$$f(z, n) = g(z) - cn \text{ for all } z \in S \text{ and for all } n \in \mathbb{N}_0.$$

Then for all $z \in S$, $n \in \mathbb{N}_0$ and $j \in \mathbb{N}$ we have

$$\begin{aligned} & \mathbb{E} \left(f \left(Z_{\tau_j(B)}, n + \tau_j(B) \right) \middle| Z_0 = z \right) \\ &= \mathbb{E} \left(g \left(Z_{\tau_j(B)} \right) - cn - c\tau_j(B) \middle| Z_0 = z \right) \\ &= \mathbb{E} \left(g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right) - cn - c\mathbb{E}(\tau_j(B) | Z_0 = z), \end{aligned}$$

hence the iterative step is done for any κ and any measurable $B \subseteq S \times \mathbb{N}_0$ by

$$\begin{aligned} B^{*\kappa} &:= \left\{ (z, n) \in B ; f(z, n) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(f \left(Z_{\tau_j(B)}, n + \tau_j(B) \right) \middle| Z_0 = z \right) \right\} \\ &= \left\{ (z, n) \in B ; g(z) \geq \sup_{1 \leq j < \kappa+1} \left(\mathbb{E} \left(g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right) - c\mathbb{E}(\tau_j(B) | Z_0 = z) \right) \right\}. \end{aligned}$$

Obviously there is no time-dependence any more.

This example is extended in Part 6.4 (page 65).

6.3.3 Constant Discounting

Consider some measurable mapping $g : S \rightarrow \mathbb{R}$, some (constant discount factor) $\alpha \in]0, 1]$ and

$$f(z, n) = \alpha^n g(z) \text{ for all } z \in S \text{ and for all } n \in \mathbb{N}_0$$

Then for all $z \in S$, $n \in \mathbb{N}_0$ and $j \in \mathbb{N}$ we have

$$\begin{aligned} & \mathbb{E} \left(f \left(Z_{\tau_j(B)}, n + \tau_j(B) \right) \middle| Z_0 = z \right) \\ &= \mathbb{E} \left(\alpha^{n+\tau_j(B)} g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right) \\ &= \alpha^n \mathbb{E} \left(\alpha^{\tau_j(B)} g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right), \end{aligned}$$

hence the iterative step is done for any κ and any measurable $B \subseteq S \times \mathbb{N}_0$ by

$$\begin{aligned} B^{*\kappa} &:= \left\{ (z, n) \in B ; f(z, n) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(f \left(Z_{\tau_j(B)}, n + \tau_j(B) \right) \middle| Z_0 = z \right) \right\} \\ &= \left\{ (z, n) \in B ; g(z) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(\alpha^{\tau_j(B)} g \left(Z_{\tau_j(B)} \right) \middle| Z_0 = z \right) \right\}. \end{aligned}$$

Again there is no time-dependence any more.

We consider the case of random discounting in detail in Section 7 (page 67).

6.4 Path-Dependent Costs

Consider some measurable mappings $g, c : S \rightarrow \mathbb{R}$. $g(z)$ is the payoff for stopping the observation and $c(z)$ is the cost of continuation for each point z . The simplification of constant c boils down to the case of linear costs considered in Part 6.3.2 (page 64). Consider the pay-off

$$X_n := g(Z_n) - \sum_{i=0}^{n-1} c(Z_i) \text{ for all } n \in \mathbb{N}_0.$$

The payoff is path-dependent. Now we apply the algorithm and its results to a time-independent “space-cost” Markov process $(Z_n, \sum_{i=0}^{n-1} c(Z_i))$ on $S \times \mathbb{R}$ instead of S . So we can treat the problem as follows:

Consider

$$f : S \times \mathbb{R} \longrightarrow \mathbb{R}, (z, d) \longmapsto g(z) - d$$

and the problem of optimal stopping for the pay-off

$$X_n := f \left(Z_n, \sum_{i=0}^{n-1} c(Z_i) \right) = g(Z_n) - \sum_{i=0}^{n-1} c(Z_i) \text{ for all } n \in \mathbb{N}_0,$$

defining (as above) for simpler notation

$$X_\infty := f \left(Z_\infty, \sum_{i=0}^{\infty-1} c(Z_i) \right) := \lim_{n \rightarrow \infty} \sup_{k \geq n} f \left(Z_k, \sum_{i=0}^{k-1} c(Z_i) \right),$$

so that we have

$$X_\tau = \left(Z_\tau, \sum_{i=0}^{\tau-1} c(Z_i) \right) \text{ for all stopping rules } \tau.$$

Assume as usual

$$\mathbb{E} (X_\tau | Z_0 = z) \text{ exists for all } \tau \in \mathcal{T} \text{ and } z \in S.$$

Then the iterative step is done for any κ and any measurable $B \subseteq S \times \mathbb{R}$ by

$$\begin{aligned}
B^{*\kappa} &:= \left\{ (z, d) \in B ; f(z, d) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(f \left(Z_{\tau_j(B)}, \sum_{i=0}^{\tau_j(B)-1} c(Z_i) + d \right) \middle| Z_0 = z \right) \right\} \\
&= \left\{ (z, d) \in B ; g(z) \geq \sup_{1 \leq j < \kappa+1} \mathbb{E} \left(g(Z_{\tau_j(B)}) - \sum_{i=0}^{\tau_j(B)-1} c(Z_i) \middle| Z_0 = z \right) \right\}.
\end{aligned}$$

Obviously there is no path-dependence any more.

This problem has been considered by Presman in [Pre10]. The iteration is therein only explicitly defined under certain additional assumptions for B of the form $B = S \setminus C_k$ in Lemma 1 as follows

$$\begin{aligned}
D_{k+1} = S \setminus C_{k+1} &= S \setminus (C_k \cup \{z \in S ; g_k(z) < Tg_k(z)\}) \\
&= S \setminus (C_k \cup \{z \in S ; g_{C_k}(z) < Tg_{C_k}(z)\}) \\
&= (S \setminus C_k) \cap (S \setminus \{z \in S ; g_{C_k}(z) < Tg_{C_k}(z)\}) \\
&= \{z \in S \setminus C_k ; g_{C_k}(z) < Tg_{C_k}(z)\}
\end{aligned}$$

This can be rewritten in the form of our algorithm above by checking that

$$B^{*1} = \{z \in B ; g_B(z) < Tg_B(z)\}$$

for each B and the appropriately defined functions g_B and T in [Pre10].

6.5 Performing the Algorithmic Step

In [Irl06, part 4] the algorithmic step of going from B to B^{*1} was performed by providing numerical values for the conditional expectations by *path-wise simulations of the Markov chain*. In the examples of [Irl06, part 4] the state space was finite and the algorithm found the solution within a surprisingly low number of iterations.

Rewriting a problem with finite state space as a *linear equation* as done in [Irl09, part 2] is another approach. It is therein done for $\kappa \equiv 1$. We will show this here for arbitrary κ in Section 8 (page 80).

6.6 Outlook

The theorems and proofs given in Section 5 (page 49) for the general case will be adapted to the Markovian case with random discounting in Section 7 (page 67). The same situation under finite state space is considered in Section 8 (page 80): Finding the solution can be rewritten as solving linear equations. Numerical examples calculated this way will be given in Section 9 (page 84).

7 Markovian Case with Random Discounting

The proofs given in Section 5 (page 49) for the general case will be adapted to the Markovian case with random discounting in this section. Herein we will also give some examples to explain the origin and to confirm the necessity of some definitions within the algorithm. Simplifications of the calculations for the case of finite state space will be examined in Section 8 (page 80). Numerical examples can be found in Section 9 (page 84).

We motivate in Part 7.1 (page 67) the setting we will give in Setting 7.2 (page 68). The main theorem of this chapter is Theorem 7.11 (page 76). This theorem is based on Lemma 7.4 (page 69), which is proved in Proof 7.9 (page 74) and uses therein Lemma 7.7 (page 72) and Lemma 7.6 (page 71). Example 7.5 (page 69) and Example 7.8 (page 72) substantiate necessities in these lemmata.

Lemma 7.4 (page 69) is adapted from [Irl09, Part 3, Theorem 1, Proof, Part (a)], where the case of $D = \{1\}$ is considered, now proved for almost arbitrary D . Theorem 7.11 (page 76) is based on ideas of [Irl09, Part 3, Theorem 1], where $\kappa = 1$ is used. The basic ideas are from [Irl80].

7.1 Motivation

Consider a (time-homogeneous) Markov chain in continuous time $(\widehat{Z}_t)_{t \in [0, \infty[}$ operating on the discrete state space $(S, \mathcal{P}ot(S))$ and a payoff

$$\widehat{X}_t := \widehat{\alpha}^t g(\widehat{Z}_t) \text{ for each } t \in [0, \infty[\text{ and } \widehat{X}_\infty := \lim_{t \rightarrow \infty} \sup_{s \geq t} \widehat{X}_s.$$

Assuming no instantaneous jumps (i. e. between two jump times there is always a positive amount of time) and no absorption, the above Markov chain has a pure jump structure, it is a special Markov jump process and we can describe it up to explosion time completely by a sequence of jump times and a sequence of states visited with adequate parameters, see [Asm03, p. 39-40]. When focussing on the problem of optimal stopping of the above Markov chain, we can reduce this problem to one of the following setting as described in [Irl09, part 3].

7.2 Setting

Let $(Z_n)_{n \in \mathbb{N}_0}$ be a time-homogeneous Markov process with respect to the underlying filtration and the discrete state space $(S, \mathcal{P}ot(S))$, $g : S \rightarrow \mathbb{R}$, $\alpha : S \rightarrow [0, 1]$.

We look at the problem of optimal stopping for the pay-off

$$X_n := \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) g(Z_n) \text{ for all } n \in \mathbb{N}_0, \quad X_\infty := \limsup_{n \rightarrow \infty} \sup_{k \geq n} X_k.$$

We write for simpler notation formally

$$X_\infty = \left(\prod_{i=0}^{\infty} \alpha(Z_i) \right) g(Z_\infty)$$

so that we have

$$X_\tau = \left(\prod_{i=0}^{\tau-1} \alpha(Z_i) \right) g(Z_\tau) \text{ for all stopping rules } \tau.$$

Assume as usual

$$(7.2.1) \quad \mathbb{E}(X_\tau | Z_0 = z) \text{ exists for all } \tau \in \mathcal{T} \text{ and } z \in S.$$

For all $B \in \mathcal{S}$ define

$$\mathcal{T}(B) := \{\tau \in \mathcal{T} ; Z_\tau \in B \text{ on } \{\tau < \infty\}\}.$$

For all $B \in \mathcal{S}$ and $\sigma \in \mathcal{T}$ define

$$(7.2.2) \quad \tau_\sigma(B) := \inf \{k \geq \sigma ; Z_k \in B\},$$

the time of first entrance in B at or after time σ ,
and if $\sigma \equiv n$ for some $n \in \mathbb{N}_0$ write $\tau_n(B)$ for $\tau_\sigma(B)$.

Define for all $B \in \mathcal{S}$ and $k \in \mathbb{N}$

$$B^{*k} := \{z \in B ; g(z) \geq \sup \{\mathbb{E}(X_{\tau_l(B)} | Z_0 = z) ; l \in \mathbb{N}_{\leq k}\}\}.$$

While running numerical examples (see Chapter 9 (page 84)) it appeared that using any $D \subseteq \mathbb{N}$ with $1 \in D$ for the improvement step instead of an initial sequence of \mathbb{N} gives also optimal results and the convergence speed depends only on the cardinality of D not on the largest element of D . Hence let us define for all $B \in \mathcal{S}$ and $D \subseteq \mathbb{N}$

$$B^{*D} := \{z \in B ; g(z) \geq \sup \{\mathbb{E}(X_{\tau_k(B)} | Z_0 = z) ; k \in D\}\}.$$

This extends the definition above since $B^{*k} = B^{*\mathbb{N}_{\leq k}}$.

We tried to prove all results for arbitrary D — and for the first part of Lemma 7.4 (page 69) this worked. But already for the second part we had to draw back to an initial sequence of \mathbb{N} . The reason is shown in Example 7.8 (page 72).

For all $B \in \mathcal{S}$, $D \subseteq \mathbb{N}$, $\sigma \in \mathcal{T}(B)$ and $\rho \in \mathcal{T}(B)$ with $\sigma \leq \rho \leq \tau_\sigma(B^{*D})$ define

$$(7.2.3) \quad \hat{\rho} := \rho \mathbf{1}_{\{\rho = \tau_\sigma(B^{*D})\}} + \sum_{n=0}^{\infty} \sum_{j \in D} \mathbf{1}_{\{\rho = n\} \cap \{j = \inf\{i \in D ; Z_n \notin B^{*D}_{\leq i}\}\}} (\tau_{n+j}(B) \wedge \tau_\sigma(B^{*D})).$$

7.3 Remarks

- The time-independent case is included by setting $\alpha \equiv 1$.
- The situation of constant discounting as described in Part 6.3.3 (page 65) is included by assuming the function α being constant.
- If $D = \{1\}$, then we have by (7.2.3)

$$\hat{\rho} = \rho \mathbb{1}_{\{\rho = \tau_\sigma(B^{*\{1\}})\}} + \sum_{n=0}^{\infty} \mathbb{1}_{\{\rho = n\}} \tau_{n+1}(B).$$

7.4 Lemma: One-Step Improvement

Let $B \in \mathcal{S}$, $D \subseteq \mathbb{N}$, $\sigma \in \mathcal{T}(B)$ and

$$(7.4.1) \quad \rho \in \mathcal{T}(B) \text{ with } \sigma \leq \rho \leq \tau_\sigma(B^{*D}).$$

Then we have

$$(7.4.2) \quad \hat{\rho} \in \mathcal{T}(B) \text{ with } \sigma \leq \hat{\rho} \leq \tau_\sigma(B^{*D}),$$

$$(7.4.3) \quad \rho \leq \hat{\rho},$$

$$(7.4.4) \quad \rho + 1 \leq \hat{\rho} \text{ on } \{\rho < \tau_\sigma(B^{*D})\}.$$

Assume additionally the existence of $k \in \mathbb{N} \cup \{\infty\}$ with $D = \mathbb{N}_{\leq k}$.

Then we have

$$(7.4.5) \quad \mathbb{E}(X_\rho | Z_0 = z) \leq \mathbb{E}(X_{\hat{\rho}} | Z_0 = z) \text{ for all } z \in S,$$

and if additionally

$$(7.4.6) \quad \mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} \middle| Z_0 = z\right) = \lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | Z_0 = z)$$

for all $z \in S$ and all non-decreasing sequences $(\sigma_n)_{n \in \mathbb{N}_0}$ of stopping rules,

then we have

$$(7.4.7) \quad \mathbb{E}(X_\sigma | Z_0 = z) \leq \mathbb{E}\left(X_{\tau_\sigma(B^{*D})} \middle| Z_0 = z\right) \text{ for all } z \in S.$$

Remark:

- This lemma is an adaption of Lemma 5.10 (page 53) above to Markovian stopping problems.
- The proof is postponed, see Proof 7.9 (page 74).

7.5 Example: Insufficient Improvement

Using the definition

$$\hat{\rho} := \rho \mathbb{1}_{\{\rho = \tau_\sigma(B^{*D})\}} + \sum_{n=0}^{\infty} \sum_{j \in D} \mathbb{1}_{\{\rho = n\}} \cap \{j = \inf\{i \in D; Z_n \notin B^{*D \leq i}\}\} \tau_{n+j}(B)$$

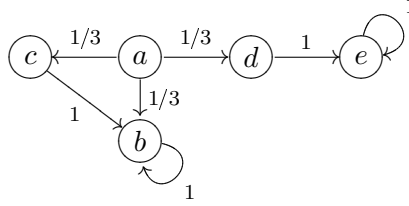
instead of (7.2.3) may contradict (7.4.2) and (7.4.7).

We can see this by the following example:

Imagine a Markov chain with five states a, b, c, d, e and the following transition matrix

	a	b	c	d	e
a	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0
b	0	1	0	0	0
c	0	1	0	0	0
d	0	0	0	0	1
e	0	0	0	0	1

as drawn in the following picture



and the payments, i.e. the function g

state	payment
a	3
b	4
c	1.5
d	2.5
e	2

and we will use $D := \{1, 2\}$, $\alpha := 1$. We initialize the algorithm with

$$B = \{a, b, c, d, e\}.$$

Hence we have

$$B^{\{1\}} = \{a, b, d, e\}$$

and

$$B^{\{1,2\}} = \{b, d, e\}.$$

Now we determine $\hat{\rho}$ for $\rho := \tau_\sigma(B^{\{1\}})$. On $\{Z_0 \neq a\}$ we have $\hat{\rho} = \tau_\sigma(B^{\{1,2\}})$.

If $Z_0 = a$, then $\rho = n = 0$, hence

$$j := \inf \{i \in D ; Z_n \notin B^{*D \leq i}\} = \inf \{i \in D ; a \notin B^{*D \leq i}\} = 2$$

and thus on $\{Z_0 = a\}$

$$\begin{aligned} \hat{\rho} &= \tau_{n+j}(B) = \tau_{0+2}(B) = \tau_2(B) = \inf \{n \geq 2 ; Z_n \in \{b, e\}\} \\ &> \inf \{n \geq 0 ; Z_n \in \{b, d, e\}\} = \tau_\sigma(B^{\{1,2\}}). \end{aligned}$$

This is contradicting (7.4.2), and this $\hat{\rho}$ gives a smaller expected revenue, contradicting (7.4.7).

Hence it is necessary to take the minimum with σ^{*D} for a sufficient improvement, in the sense of having $\hat{\rho}$ always smaller than σ^{*D} and giving not a smaller expected revenue.

7.6 Lemma: Partial Improving I

Assume $B \in \mathcal{S}$, $D \subseteq \mathbb{N}$, $n \in \mathbb{N}_0$, $j \in D$, $A \in \mathcal{A}_n$ with $A \subseteq \{Z_n \in B^{*D < j} \setminus B^{*D \leq j}\}$, $z \in S$. Then we have

$$\int_A X_n dP_z \leq \int_A X_{\tau_{n+j}(B)} dP_z.$$

Proof: For all $E \subseteq D$, $j \in E$ we have

$$(7.6.1) \quad \begin{aligned} B^{*E} &\subseteq B^{*E \setminus \{j\}}, \\ B^{*E} &= \left\{ z \in B^{*E \setminus \{j\}} ; g(z) \geq \mathbb{E} \left(X_{\tau_j(B)} \middle| Z_0 = z \right) \right\}, \\ B^{*E \setminus \{j\}} \setminus B^{*E} &= \left\{ z \in B^{*E \setminus \{j\}} ; g(z) < \mathbb{E} \left(X_{\tau_j(B)} \middle| Z_0 = z \right) \right\}. \end{aligned}$$

Assume $(Z'_m)_{m \in \mathbb{N}_0}$ being an independent copy of $(Z_m)_{m \in \mathbb{N}_0}$ and

$$\tau'_p(B) = \inf \{ q \geq p ; Z'_q \in B \} \text{ for all } p \in \mathbb{N}_0.$$

Then we have

$$\begin{aligned} &\int_A X_n dP_z \\ \stackrel{\text{def. of } X_n}{=} &\int_A \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) g(Z_n) dP_z \\ \stackrel{Z_n \in B^{*D < j} \setminus B^{*D \leq j}}{\leq} &\int_A \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) \mathbb{E} \left(g \left(Z'_{\tau'_j(B)} \right) \prod_{i=0}^{\tau'_j(B)-1} \alpha(Z'_i) \middle| Z'_0 = Z_n \right) dP_z \\ \stackrel{(7.6.1)}{\leq} &\int_A \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) \mathbb{E} \left(g \left(Z_{\tau_{n+j}(B)} \right) \prod_{i=n}^{\tau_{n+j}(B)-1} \alpha(Z_i) \middle| \mathcal{A}_n \right) dP_z \\ \stackrel{\text{M.C.}}{=} &\int_A \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) \mathbb{E} \left(g \left(Z_{\tau_{n+j}(B)} \right) \prod_{i=n}^{\tau_{n+j}(B)-1} \alpha(Z_i) \middle| \mathcal{A}_n \right) dP_z \\ \stackrel{A \in \mathcal{A}_n}{=} &\int_A \left(\prod_{i=0}^{n-1} \alpha(Z_i) \right) g \left(Z_{\tau_{n+j}(B)} \right) \prod_{i=n}^{\tau_{n+j}(B)-1} \alpha(Z_i) dP_z \\ = &\int_A \left(\prod_{i=0}^{\tau_{n+j}(B)-1} \alpha(Z_i) \right) g \left(Z_{\tau_{n+j}(B)} \right) dP_z \\ \stackrel{\text{def.}}{=} &\int_A X_{\tau_{n+j}(B)} dP_z. \end{aligned}$$

□

7.7 Lemma: Partial Improving II

Assume $B \in \mathcal{S}$, $D \subseteq \mathbb{N}$, $s, t \in \mathbb{N}$ with $s \leq t$ and $t - s \in D$, $A \in \mathcal{A}_s$ with $A \subseteq \{Z_s \in B^{*D}\}$. Then we have

$$(7.7.1) \quad \int_A X_{\tau_t(B)} dP_z \leq \int_A X_s dP_z.$$

Proof: Assume $(Z'_n)_{n \in \mathbb{N}_0}$ being an independent copy of $(Z_n)_{n \in \mathbb{N}_0}$ and

$$\tau'_p(B) = \inf \{q \geq p ; Z'_q \in B\} \text{ for all } p \in \mathbb{N}_0.$$

Then we have

$$\begin{aligned} & \int_A X_{\tau_t(B)} dP_z \\ \stackrel{\text{def.}}{=} & \int_A \left(\prod_{i=0}^{\tau_t(B)-1} \alpha(Z_i) \right) g(Z_{\tau_t(B)}) dP_z \\ = & \int_A \left(\prod_{i=0}^{s-1} \alpha(Z_i) \right) g(Z_{\tau_t(B)}) \prod_{i=s}^{\tau_t(B)-1} \alpha(Z_i) dP_z \\ \stackrel{A \in \mathcal{A}_s}{=} & \int_A \left(\prod_{i=0}^{s-1} \alpha(Z_i) \right) \mathbb{E} \left(g(Z_{\tau_t(B)}) \prod_{i=s}^{\tau_t(B)-1} \alpha(Z_i) \middle| \mathcal{A}_s \right) dP_z \\ \stackrel{\text{M.C.}}{=} & \int_A \left(\prod_{i=0}^{s-1} \alpha(Z_i) \right) \mathbb{E} \left(g(Z'_{\tau'_{t-s}(B)}) \prod_{i=0}^{\tau'_{t-s}(B)-1} \alpha(Z'_i) \middle| Z'_0 = Z_s \right) dP_z \\ \stackrel{\substack{Z_s \in B^{*D} \\ t-s \in D}}{=} & \int_A \left(\prod_{i=0}^{s-1} \alpha(Z_i) \right) g(Z_s) dP_z \\ \stackrel{\text{def. of } X_s}{=} & \int_A X_s dP_z. \end{aligned}$$

□

Remark: The assumption “ $t - s \in D$ ” is necessary as shown in Example 7.8 (page 72).

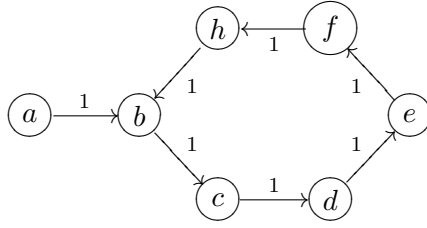
7.8 Example: Necessities for Partial Improving II

This example will show the necessity of the condition “ $t - s \in D$ ” for (7.7.1).

Imagine a Markov chain with seven states a, b, c, d, e, f, h and the following transition matrix

	a	b	c	d	e	f	h
a	0	1	0	0	0	0	0
b	0	0	1	0	0	0	0
c	0	0	0	1	0	0	0
d	0	0	0	0	1	0	0
e	0	0	0	0	0	1	0
f	0	0	0	0	0	0	1
h	0	1	0	0	0	0	0

as drawn in the following picture



and the payments, i.e. the function g

state	payment
a	4
b	2
c	7
d	5
e	5
f	3
h	8

and we will use $D := \{1, 3\}$ and $\alpha := 1$. We initialize the algorithm with

$$B := \{a, b, c, d, e, f, h\}.$$

Hence we have

$$B^{*D} = \{c, e, h\}$$

and

$$F = \{h\}.$$

Now consider $\rho := 0$, $A := \{Z_0 \in \{d\}\}$, $t := 3$, $s := 1$.

We have $A \in \mathcal{A}_1$ and $A \subseteq \{Z_1 = e\} \subseteq \{Z_1 \in B^{*D}\} = \{Z_s \in B^{*D}\}$
 (With the definitions of the proof of (7.4.5) and (7.9.4) below we have $A^{0,3} = \{Z_0 \in \{a, d\}\}$
 splitting into the sets $A = A_1 = \{Z_0 \in \{d\}\}$ and $A_2 = \{Z_0 \in \{a\}\}$.)

We have $t - s = 3 - 1 = 2 \notin D$ and

$$\begin{aligned}
 & \mathbf{E}(\mathbf{1}_A X_{\tau_t(B)} | Z_0 = d) \\
 &= \mathbf{E}(\mathbf{1}_A X_{\tau_3(B)} | Z_0 = d) \\
 &= \mathbf{E}(X_{\tau_3(B)} | Z_0 = d) \\
 &= \mathbf{E}(X_3 | Z_0 = d) \\
 &= \mathbf{E}(g(Z_3) | Z_0 = d) \\
 &= g(h) \\
 &= 8 \\
 &> 5 \\
 &= g(e) \\
 &= \mathbf{E}(X_1 | Z_0 = d) \\
 &= \mathbf{E}(\mathbf{1}_A X_1 | Z_0 = d) \\
 &= \mathbf{E}(\mathbf{1}_A X_s | Z_0 = d),
 \end{aligned}$$

contradicting (7.7.1) and (7.9.8).

7.9 Proof of Lemma 7.4 (page 69)

Proof of (7.4.2), (7.4.3), (7.4.4): We have

$$\rho = \hat{\rho} = \tau_\sigma(B^{*D}) \text{ on } \{\rho = \tau_\sigma(B^{*D})\} = \{\rho = \infty\} \cup \{Z_\rho \in B^{*D}\}.$$

Furthermore we have $B^{*\emptyset} = B$,

$$(7.9.1) \quad B^{*D} = \bigcap_{j \in D} B^{*D \leq j}.$$

We have

$$\{Z_n \notin B^{*D}\} \stackrel{(7.9.1)}{=} \bigcup_{j \in D} \{j = \inf \{i \in D; Z_n \notin B^{*D \leq i}\}\} \text{ for all } n \in \mathbb{N}_0.$$

So for all $n \in \mathbb{N}_0$ and $j \in D$ we have

$$\rho < \rho + 1 = n + 1 \leq \overbrace{(\tau_{n+j}(B) \wedge \tau_\sigma(B^{*D}))}^{=\hat{\rho}} \leq \tau_\sigma(B^{*D})$$

$$\text{on } \{\rho = n\} \cap \{j = \inf \{i \in D; Z_n \notin B^{*D \leq i}\}\}.$$

Hence

$$(7.9.2) \quad \{\rho < \tau_\sigma(B^{*D})\} = \{\rho < \infty\} \cap \{Z_\rho \notin B^{*D}\} = \bigcup_{n=0}^{\infty} \{\rho = n\} \cap \{Z_n \notin B^{*D}\}. \quad \square$$

Proof of (7.4.5): Let $z \in S$ and define

$$A := \{\rho < \tau_\sigma(B^{*D})\} \stackrel{(7.9.2)}{=} \{\rho < \infty\} \cap \{Z_\rho \notin B^{*D}\}.$$

We have

$$A^c = \{\rho = \infty\} \cup \{Z_\rho \in B^{*D}\} = \{\rho = \tau_\sigma(B^{*D})\} \subseteq \{\rho = \hat{\rho}\}$$

and thus obviously

$$\mathbb{E}(\mathbf{1}_{A^c} X_\rho | Z_0 = z) \leq \mathbb{E}(\mathbf{1}_{A^c} X_{\hat{\rho}} | Z_0 = z).$$

Define for all $n \in \mathbb{N}_0$ and $j \in D$

$$A^{n,j} := \{\rho = n\} \cap \{j = \inf \{i \in D; Z_n \notin B^{*D \leq i}\}\}.$$

We have

$$A = \bigcup_{n=0}^{\infty} \bigcup_{j \in D} A^{n,j}.$$

For all $n \in \mathbb{N}_0$ and $j \in D$ we have by Lemma 7.6 (page 71)

$$(7.9.3) \quad \mathbb{E}(\mathbf{1}_{A^{n,j}} X_\rho | Z_0 = z) = \mathbb{E}(\mathbf{1}_{A^{n,j}} X_n | Z_0 = z) \leq \mathbb{E}(\mathbf{1}_{A^{n,j}} X_{\tau_{n+j}(B)} | Z_0 = z).$$

Due to dominated convergence it suffices to show for all $n \in \mathbb{N}_0$ and $j \in D$ the inequality

$$(7.9.4) \quad \mathbb{E}(\mathbf{1}_{A^{n,j}} X_{\tau_{n+j}(B)} | Z_0 = z) \leq \mathbb{E}(\mathbf{1}_{A^{n,j}} X_{\hat{\rho}} | Z_0 = z). \quad \square$$

Remark:

- (7.9.3) stays correct even for arbitrary $D \subseteq \mathbb{N}$.
- As shown in Example 7.8 (page 72) the inequality (7.9.4) may be false if there is no $k \in \mathbb{N} \cup \{\infty\}$ with $D = \mathbb{N}_{\leq k}$.

- Nevertheless it seems that the inequality

$$(7.9.5) \quad \mathbb{E}(\mathbf{1}_{A^{n,j}} X_\rho | Z_0 = z) \leq \mathbb{E}(\mathbf{1}_{A^{n,j}} X_{\hat{\rho}} | Z_0 = z)$$

is true for all D with $1 \in D$ as studies with our computer program show, but this could not be proved.

Proof of (7.9.4): Define

$$E := \{\tau_{n+j}(B) \leq \tau_\sigma(B^{*D})\}.$$

We have

$$\tau_{n+j}(B) = \tau_{n+j}(B) \wedge \tau_\sigma(B^{*D}) = \hat{\rho} \text{ on } A^{n,j} \cap E,$$

hence

$$(7.9.6) \quad \mathbb{E}(\mathbf{1}_{A^{n,j} \cap E} X_{\tau_{n+j}(B)} | Z_0 = z) \leq \mathbb{E}(\mathbf{1}_{A^{n,j} \cap E} X_{\hat{\rho}} | Z_0 = z).$$

For all $m \in \mathbb{N}$ with $m < j$ define

$$(7.9.7) \quad A_m := A^{n,j} \cap \{Z_{n+m} \in B^{*D}\} \cap \bigcap_{l=1}^{m-1} \{Z_{n+l} \notin B^{*D}\}.$$

For all $m \in \mathbb{N}$ with $m < j$ and $j - m \in D$ we have by Lemma 7.7 (page 72) (with setting $A := A_m$, $s := n + m$, $t := n + j$)

$$(7.9.8) \quad \int_{A_m} X_{\tau_{n+j}(B)} dP_z \stackrel{7.7}{\leq} \int_{A_m} X_{n+m} dP_z \stackrel{A_m \subseteq \{n+m = \tau_\sigma(B^{*D})\}}{=} \int_{A_m} X_{\tau_\sigma(B^{*D})} dP_z \stackrel{A_m \subseteq E^c}{=} \int_{A_m} X_{\hat{\rho}} dP_z.$$

We will prove next

$$(7.9.9) \quad (A^{n,j} \cap E^c) = \bigsqcup_{m=1}^{j-1} A_m.$$

Due to linearity of expectation this finishes the proof. \square

Proof of (7.9.9):

Disjoint: For all $m, l \in \mathbb{N}$ with $m, l < j$, $m \neq l$ we have

$$\emptyset \subseteq A_m \cap A_l \subseteq \{Z_{n+m} \in B^{*D}\} \cap \{Z_{n+m} \notin B^{*D}\} = \emptyset, \text{ hence } A_m \cap A_l = \emptyset.$$

“ \supseteq ”: For all $m \in \mathbb{N}$ with $m < j$ we have due to $n + m < n + j \leq \tau_{n+j}(B)$

$$A_m \stackrel{(7.9.7)}{\subseteq} \{\tau_\sigma(B^{*D}) = n + m\} \subseteq \{\tau_\sigma(B^{*D}) < \tau_{n+j}(B)\} \stackrel{\text{def.}}{=} E^c.$$

“ \subseteq ”: Consider $\omega \in A^{n,j} \cap E^c$.

For all $l \in \mathbb{N}_0$ with $n + j \leq l < \tau_{n+j}(B)(\omega) = \inf\{l \geq n + j; Z_l(\omega) \in B\}$ we obviously have $Z_l(\omega) \notin B$ and due to $B^{*D} \subseteq B$ we have $Z_l(\omega) \notin B^{*D}$.

Since we consider

$$\omega \in A^{n,j} \cap E^c \subseteq \{\tau_\sigma(B^{*D}) < \tau_{n+j}(B)\} \cap \{\rho = n\},$$

hence

$$\tau_{n+j}(B)(\omega) > \tau_\sigma(B^{*D})(\omega) = \inf\{l \geq \sigma(\omega); Z_l(\omega) \in B^{*D}\},$$

we obtain $n + j > \tau_\sigma(B^{*D})(\omega)$ and by (7.4.1) $\tau_\sigma(B^{*D})(\omega) \geq \rho(\omega) = n$. By the definition of $A^{n,j}$ we have $Z_n(\omega) \notin B^{*D \leq j} \supseteq B^{*D}$, hence $\tau_\sigma(B^{*D})(\omega) \neq n$. So we have $n < \tau_\sigma(B^{*D})(\omega) < n + j$, thus it follows $\omega \in A_{[\tau_\sigma(B^{*D})(\omega) - n]}$ by (7.9.7). \square

Proof of (7.4.7): Define $\sigma_0 := \sigma$ and inductively

$$\sigma_{k+1} := \hat{\sigma}_k \text{ for all } k \in \mathbb{N}_0.$$

First we will prove by induction that

$$(7.9.10) \quad \sigma_k \leq \sigma_{k+1} \leq \tau_\sigma(B^{*D}) \text{ for all } k \in \mathbb{N}_0.$$

We have $\sigma = \sigma_0 \leq \tau_\sigma(B^{*D})$.

Then we have $\sigma = \sigma_0 \leq \sigma_1 \leq \tau_\sigma(B^{*D})$ by (7.4.2) and (7.4.3).

Consider $k \in \mathbb{N}_0$ with $\sigma \leq \sigma_k \leq \tau_\sigma(B^{*D})$.

It follows $\sigma \leq \sigma_k \leq \sigma_{k+1} \leq \tau_\sigma(B^{*D})$ by (7.4.2) and (7.4.3).

By (7.9.10) we have

$$\lim_{k \rightarrow \infty} \sigma_k \leq \tau_\sigma(B^{*D})$$

and by (7.4.4)

$$\sigma_k + 1 \leq \sigma_{k+1} \text{ on } \{\sigma_k < \tau_\sigma(B^{*D})\} \text{ for all } k \in \mathbb{N},$$

hence we have

$$\sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \dots \leq \lim_{k \rightarrow \infty} \sigma_k = \tau_\sigma(B^{*D}).$$

Let $z \in S$. We have for all $k \in \mathbb{N}_0$

$$(7.9.11) \quad \mathbb{E}(X_{\sigma_k} | Z_0 = z) \leq \mathbb{E}(X_{\sigma_{k+1}} | Z_0 = z) \quad \text{by (7.4.5).}$$

We infer by induction and assuming (7.4.6)

$$\begin{aligned} \mathbb{E}(X_\sigma | Z_0 = z) &= \mathbb{E}(X_{\sigma_0} | Z_0 = z) \\ &\leq \lim_{k \rightarrow \infty} \mathbb{E}(X_{\sigma_k} | Z_0 = z) && \text{by (7.9.11)} \\ &= \mathbb{E}\left(X_{\lim_{k \rightarrow \infty} \sigma_k} \middle| Z_0 = z\right) && \text{by (7.4.6)} \\ &= \mathbb{E}\left(X_{\tau_\sigma(B^{*D})} \middle| Z_0 = z\right). \end{aligned} \quad \square$$

7.10 Remark

We used the superscript $*D$ above to depict that some parts of the proofs and lemmatas are true regardless of the structure of D , but we could show the following results only for the special form of D being an initial sequence of \mathbb{N} . Therefore we defined above in Setting 7.2 (page 68) for all $k \in \mathbb{N}$ the shortcut B^{*k} for $B^{*\mathbb{N}_{\leq k}}$.

7.11 Theorem

Let $\kappa : \mathbb{N} \rightarrow \mathbb{N} \cup \{\infty\}$ and $B^0 \in \mathcal{S}$.

Define inductively

$$(7.11.1) \quad B^k := (B^{k-1})^{*\kappa(k)} \text{ for all } k \in \mathbb{N}, \quad F := \bigcap_{k \in \mathbb{N}_0} B^k.$$

Assume (as above)

$$(7.4.6) \quad \mathbb{E}\left(X_{\lim_{n \rightarrow \infty} \sigma_n} \middle| Z_0 = z\right) = \lim_{n \rightarrow \infty} \mathbb{E}(X_{\sigma_n} | Z_0 = z)$$

for all $z \in S$ and all non-decreasing sequences $(\sigma_n)_{n \in \mathbb{N}_0}$ of stopping rules.

Then we have for all $j \in \mathbb{N}_0$

$$(7.11.2) \quad \tau_j(B^k) \leq \tau_j(B^{k+1}) \text{ for all } k \in \mathbb{N}_0,$$

$$(7.11.3) \quad \lim_{k \rightarrow \infty} \tau_j(B^k) = \tau_j(F),$$

$$(7.11.4) \quad \mathbb{E} \left(X_{\tau_j(B^k)} \mid Z_0 = z \right) \leq \mathbb{E} \left(X_{\tau_j(B^{k+1})} \mid Z_0 = z \right) \text{ for all } k \in \mathbb{N}_0 \text{ for all } z \in S,$$

$$(7.11.5) \quad \lim_{k \rightarrow \infty} \mathbb{E} \left(X_{\tau_j(B^k)} \mid Z_0 = z \right) = \mathbb{E} \left(X_{\tau_j(F)} \mid Z_0 = z \right) = \sup_{\substack{\tau \in \mathcal{T}(B^0) \\ \tau \geq j}} \mathbb{E}(X_\tau \mid Z_0 = z) \text{ for all } z \in S,$$

hence $\tau_j(F)$ is optimal in $\{\tau \in \mathcal{T}(B^0) ; \tau \geq j\}$, and thus $\tau_0(F)$ is optimal in $\mathcal{T}(B^0)$.

Remark: The herein made assumption is the same as in Lemma 7.4 (page 69).

Remark: For getting the optimal stopping rule, set $B^0 = S$.

Proof of (7.11.2) and (7.11.4): Let $j, k \in \mathbb{N}_0$. By Lemma 7.4 (page 69) we have

$$B^{k+1} = (B^k)^{* \kappa(k)} \supseteq B^k$$

and

$$\tau_j(B^k) \leq \tau_{\tau_j(B^k)}((B^k)^{* \kappa(k)}) = \tau_{\tau_j(B^k)}(B^{k+1}) = \tau_j(B^{k+1}).$$

(Set B and σ of the cited lemma to B^k and $\tau_j(B^k)$.)

Hence we can infer (7.11.2) and (7.11.4). □

Proof of (7.11.3): (7.11.3) follows by the definitions in (7.11.1) and (7.2.2). □

Proof of (7.11.5):

(1): At first we will show that for any $\rho \in \mathcal{T}(B^0)$ there exists

$$\rho^\infty \in \mathcal{T}(F) \text{ with } \rho^\infty \geq \rho \text{ and } \mathbb{E}(X_\rho \mid Z_0 = z) \leq \mathbb{E}(X_{\rho^\infty} \mid Z_0 = z) \text{ for all } z \in S.$$

Let $\rho \in \mathcal{T}(B^0)$. Define

$$\begin{aligned} \rho^k &:= \inf \left\{ n \geq \rho ; Z_n \in B^k \right\} \text{ for all } k \in \mathbb{N}_0, \\ \rho^\infty &:= \inf \left\{ n \geq \rho ; Z_n \in F \right\}. \end{aligned}$$

Since $(B^k)_{k \in \mathbb{N}_0}$ is a non-increasing sequence of sets with limit F ,
 $(\rho^k)_{k \in \mathbb{N}_0}$ is a non-decreasing sequence of stopping rules with limit ρ^∞ .
Let $k \in \mathbb{N}_0$.

By Lemma 7.4 (page 69) we have

$$\mathbb{E}(X_{\rho^k} \mid Z_0 = z) \leq \mathbb{E}(X_{\rho^{k+1}} \mid Z_0 = z) \text{ for all } z \in S.$$

(Set B and σ of the cited lemma to B^k and ρ^k , then we have $\tau_{\rho^k}((B^k)^{* \kappa(k)}) = \rho^{k+1}$.)

Since $\rho \in \mathcal{T}(B^0)$, we have $\rho = \rho^0$.

So it follows

$$\begin{aligned} \mathbb{E}(X_\rho \mid Z_0 = z) &\leq \lim_{k \rightarrow \infty} \mathbb{E}(X_{\rho^k} \mid Z_0 = z) \\ &\stackrel{(7.4.6)}{\leq} \mathbb{E} \left(X_{\lim_{k \rightarrow \infty} \rho^k} \mid Z_0 = z \right) = \mathbb{E}(X_{\rho^\infty} \mid Z_0 = z) \text{ for all } z \in S. \end{aligned}$$

(2): Let $\rho, \tau \in \mathcal{T}(F)$ such that $\rho \leq \tau$. We will show

$$\mathbb{E}(X_\tau | Z_0 = z) \leq \mathbb{E}(X_\rho | Z_0 = z) \text{ for all } z \in S.$$

Define

$$\rho_k := \rho \mathbf{1}_{\{\rho=\tau\}} + \mathbf{1}_{\{\rho<\tau\}} \sum_{n=0}^{\infty} \mathbf{1}_{\{\rho=n\}} \tau_{n+1}(B^k) \text{ for all } k \in \mathbb{N}_0,$$

$$\rho_\infty := \rho \mathbf{1}_{\{\rho=\tau\}} + \mathbf{1}_{\{\rho<\tau\}} \sum_{n=0}^{\infty} \mathbf{1}_{\{\rho=n\}} \tau_{n+1}(F).$$

Then we have $\rho \leq \rho_k \leq \rho_{k+1} \leq \rho_\infty \leq \tau$ for all $k \in \mathbb{N}_0$ and $\lim_{k \rightarrow \infty} \rho_k = \rho_\infty$.

Furthermore we have

$$\rho + 1 = n + 1 \leq \tau_{n+1}(F) = \rho_\infty \text{ on } \{\rho < \tau\} \cap \{\rho = n\} \text{ for all } n \in \mathbb{N}_0,$$

hence

$$(7.11.6) \quad \rho + 1 \leq \rho_\infty \text{ on } \bigcup_{n=0}^{\infty} \{\rho = n\} \cup \{\rho < \tau\} = \{\rho \neq \tau\}.$$

Let $(Z'_n)_{n \in \mathbb{N}_0}$ be an independent copy of $(Z_n)_{n \in \mathbb{N}_0}$ and for all $p \in \mathbb{N}_0$ and $B \in \mathcal{S}$ let $\tau'_p(B) = \inf \{q \geq p ; Z'_q \in B\}$.

Let $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0$.

Since $\rho \in \mathcal{T}(F)$ we have $Z_\rho \in F$ on $\{\rho < \infty\}$ and thus

$$\begin{aligned} \{\rho = n\} \cap \{\rho < \tau\} &\subseteq \{Z_n \in F\} \\ &\subseteq \{Z_n \in (B^k)^{* \kappa(k)}\} \\ &\subseteq \left\{ g(Z_n) \geq \mathbb{E} \left(\left(\prod_{i=0}^{\tau'_1(B^k)-1} \alpha(Z'_i) \right) g(Z'_{\tau'_{n+1}(B^k)}) \middle| Z'_0 = Z_n \right) \right\}, \end{aligned}$$

hence we have for all $z \in S$

$$\begin{aligned} \int_{\{\rho=n<\tau\}} X_{\rho_k} dP_z &= \int_{\{\rho=n<\tau\}} \left(\prod_{i=0}^{\tau_{n+1}(B^k)-1} \alpha(Z_i) \right) g(Z_{\tau_{n+1}(B^k)}) dP_z \\ &= \int_{\{\rho=n<\tau\}} \mathbb{E} \left(\left(\prod_{i=0}^{\tau'_1(B^k)-1} \alpha(Z'_i) \right) g(Z'_{\tau'_{n+1}(B^k)}) \middle| Z'_0 = Z_n \right) \prod_{i=0}^{n-1} \alpha(Z_i) dP_z \\ &\leq \int_{\{\rho=n<\tau\}} g(Z_n) \prod_{i=0}^{n-1} \alpha(Z_i) dP_z \\ &= \int_{\{\rho=n<\tau\}} X_\rho dP_z. \end{aligned}$$

By dominated convergence follows

$$\mathbb{E}(X_{\rho_k} | Z_0 = z) \leq \mathbb{E}(X_\rho | Z_0 = z) \text{ for all } z \in S.$$

Letting $k \rightarrow \infty$ we obtain

$$(7.11.7) \quad \mathbb{E}(X_{\rho_\infty} | Z_0 = z) \leq \mathbb{E}(X_\rho | Z_0 = z) \text{ for all } z \in S.$$

Define $\rho^0 = \rho$ and $\rho^k = (\rho^{k-1})_\infty$ for all $k \in \mathbb{N}$. Then obviously $\rho \leq \rho^k \leq \rho^{k+1} \leq \tau$, and $\rho^k \in \mathcal{T}(F)$ for all $k \in \mathbb{N}$. Hence we have by (7.11.6) $\lim_{k \rightarrow \infty} \rho^k = \tau$.

It follows for all $z \in S$

$$\mathbb{E}(X_\tau | Z_0 = z) = \mathbb{E}\left(X_{\lim_{k \rightarrow \infty} \rho^k} \middle| Z_0 = z\right) \stackrel{(7.4.6)}{=} \lim_{k \rightarrow \infty} \mathbb{E}(X_{\rho^k} | Z_0 = z) \stackrel{(7.11.7)}{\leq} \mathbb{E}(X_\rho | Z_0 = z).$$

(3): Let $j \in \mathbb{N}_0$ and $\sigma \in \mathcal{T}(B^0)$ with $\sigma \geq j$. By (1), there exists $\tau \in \mathcal{T}(F)$ with $\tau \geq \sigma$ such that

$$\mathbb{E}(X_\sigma | Z_0 = z) \leq \mathbb{E}(X_\tau | Z_0 = z) \text{ for all } z \in S.$$

Obviously we have $\tau \geq j$ and thus $\tau_j(F) \leq \tau$ due to $\tau \in \mathcal{T}(F)$. Hence it follows by (2)

$$\mathbb{E}(X_\sigma | Z_0 = z) \leq \mathbb{E}(X_\tau | Z_0 = z) \leq \mathbb{E}\left(X_{\tau_j(F)} \middle| Z_0 = z\right) \text{ for all } z \in S. \quad \square$$

8 Markovian Case with Random Discounting and Finite State Space

In this chapter we consider the Markovian case with random discounting as discussed in Section 7 (page 67) and assume additionally finite state space. The algorithm requires the calculation of many conditional expectations. We show in this section that these expectations can be calculated by solving linear equations under the given assumptions. Numerical examples can be found in Section 9 (page 84).

8.1 Setting

We continue with the setting of Setting 7.2 (page 68) and additionally assume finite S . Let Π be the transition matrix and define Ψ by $\Psi_{z,y} := \alpha(z) \cdot \Pi_{z,y}$ for all $z, y \in S$. Define for all $p \in \mathbb{N}$ and all $B \in \mathcal{S}$

$$h'_{B,p} : S \longrightarrow \mathbb{R}, z \longmapsto \mathbb{E} \left(X_{\tau_p(B)} \mid Z_0 = z \right) = \mathbb{E} \left(g \left(Z_{\tau_p(B)} \right) \prod_{j=0}^{\tau_p(B)-1} \alpha(Z_j) \mid Z_0 = z \right).$$

8.2 Theorem

For all $B \in \mathcal{S}$, $p \in \mathbb{N}$ and $z \in S$ we have

$$h'_{B,p}(z) = \sum_{y \in S} (\Psi^p)_{z,y} h'_{B,0}(y).$$

Proof: Let $B \in \mathcal{S}$.

We have for all $p \in \mathbb{N}$ and $z \in S$

$$\begin{aligned} h'_{B,p}(z) &= \mathbb{E} \left(g \left(Z_{\tau_p(B)} \right) \prod_{j=0}^{\tau_p(B)-1} \alpha(Z_j) \mid Z_0 = z \right) \\ &= \alpha(z) \mathbb{E} \left(g \left(Z_{\tau_p(B)} \right) \prod_{j=1}^{\tau_p(B)-1} \alpha(Z_j) \mid Z_0 = z \right) \\ &= \alpha(z) \sum_{y \in S} P(Z_1 = y \mid Z_0 = z) \mathbb{E} \left(g \left(Z_{\tau_p(B)} \right) \prod_{j=1}^{\tau_p(B)-1} \alpha(Z_j) \mid Z_1 = y, Z_0 = z \right) \\ &\stackrel{\text{M.C.}}{=} \sum_{y \in S} \alpha(z) P(Z_1 = y \mid Z_0 = z) \mathbb{E} \left(g \left(Z_{\tau_{p-1}(B)} \right) \prod_{j=0}^{\tau_{p-1}(B)-1} \alpha(Z_j) \mid Z_0 = y \right) \\ &= \sum_{y \in S} \alpha(z) \Pi_{z,y} \mathbb{E} \left(g \left(Z_{\tau_{p-1}(B)} \right) \prod_{j=0}^{\tau_{p-1}(B)-1} \alpha(Z_j) \mid Z_0 = y \right) \\ &= \sum_{y \in S} \Psi_{z,y} \mathbb{E} \left(g \left(Z_{\tau_{p-1}(B)} \right) \prod_{j=0}^{\tau_{p-1}(B)-1} \alpha(Z_j) \mid Z_0 = y \right) \\ &= \sum_{y \in S} \Psi_{z,y} h'_{B,p-1}(y). \end{aligned}$$

Hence we have

$$h'_{B,p} = \Psi \cdot h'_{B,p-1}.$$

By induction we have for all $p \in \mathbb{N}$

$$h'_{B,p} = \Psi^p \cdot h'_{B,0},$$

hence for all $p \in \mathbb{N}$ and $z \in S$ we have

$$h'_{B,p}(z) = \sum_{y \in S} (\Psi^p)_{z,y} h'_{B,0}(y). \quad \square$$

8.3 Corollar

Using the definitions of Setting 7.2 (page 68) we have for all $B \in \mathcal{S}$ and $D \subseteq \mathbb{N}$

$$B^{*D} = \left\{ z \in B ; g(z) \geq \sup_{p \in D} \sum_{y \in S} (\Psi^p)_{z,y} h'_{B,0}(y) \right\}.$$

Proof: Due to the definitions in Setting 8.1 (page 80) the iterative step is performed for all $B \in \mathcal{S}$ and $D \subseteq \mathbb{N}$ by

$$B^{*D} = \left\{ z \in B ; g(z) \geq \sup_{p \in D} h'_{B,p}(z) \right\},$$

so the statement follows by Theorem 8.2 (page 80). □

8.4 FII by Solving Linear Equations¹

Let $B \in \mathcal{S}$ and assume

$$(8.4.1) \quad \text{for all } z \in S \text{ } ([\alpha(z) < 1] \text{ or } [P(\tau_0(B) < \infty \mid Z_0 = z) = 1]).$$

Then $h'_{B,0}$ is the unique solution of the system of linear equations which is constructed in the following:

Define the vector b by

$$b_i := \begin{cases} 1 & i \in B, \\ 0 & i \notin B, \end{cases} \quad \text{for all } i \in S,$$

the vector d by

$$d_i := b_i \cdot g(i) \text{ for all } i \in S$$

and the matrix A by

$$A_{i,j} := \delta_{i,j} - (1 - b_i) \Psi_{i,j} \text{ for all } i, j \in S.$$

Proof: For all $h \in \mathbb{R}^S$ we have by definition

$$d = Ah \iff \forall z \in B \ h(z) = g(z) \wedge \forall z \notin B \ h(z) = \alpha(z) \sum_{y \in S} P(Z_1 = y \mid Z_0 = z) h(y).$$

Now we prove in two steps that $\{h \in \mathbb{R}^S ; d = Ah\} = \{h'_{B,0}\}$.

¹[Irl09, part 2 and part 3]

Proof of $\{h'_{B,0}\} \subseteq \{h \in \mathbb{R}^S ; d = Ah\}$:

For all $z \in B$ we have $Z_0 = z \in B$ and thus $\tau_0(B) = 0$ and $\prod_{j=0}^{\tau_0(B)-1} \alpha(Z_j) = 1$, hence

$$h'_{B,0}(z) = \mathbb{E} \left(g(Z_{\tau_0(B)}) \prod_{j=0}^{\tau_0(B)-1} \alpha(Z_j) \middle| Z_0 = z \right) = \mathbb{E} (g(Z_0) \cdot 1 | Z_0 = z) = g(z)$$

and for all $z \notin B$ we have

$$\begin{aligned} h'_{B,0}(z) &= \mathbb{E} \left(g(Z_{\tau_0(B)}) \prod_{j=0}^{\tau_0(B)-1} \alpha(Z_j) \middle| Z_0 = z \right) \\ &= \alpha(z) \mathbb{E} \left(g(Z_{\tau_0(B)}) \prod_{j=1}^{\tau_0(B)-1} \alpha(Z_j) \middle| Z_0 = z \right) \\ &= \alpha(z) \sum_{y \in S} P(Z_1 = y | Z_0 = z) \mathbb{E} \left(g(Z_{\tau_1(B)}) \prod_{j=1}^{\tau_1(B)-1} \alpha(Z_j) \middle| Z_1 = y \right) \\ &\stackrel{\text{M.C.}}{=} \alpha(z) \sum_{y \in S} P(Z_1 = y | Z_0 = z) \mathbb{E} \left(g(Z_{\tau_0(B)}) \prod_{j=0}^{\tau_0(B)-1} \alpha(Z_j) \middle| Z_0 = y \right) \\ &= \alpha(z) \sum_{y \in S} P(Z_1 = y | Z_0 = z) h'(y). \end{aligned}$$

So we have $d = Ah'_{B,0}$.

Proof of $|\{h \in \mathbb{R}^S ; d = Ah\}| = 1$:

By (8.4.1) we have

$$\left(\left(\alpha(z) < 1 \right) \text{ or } \left(z \notin B \Rightarrow \sum_{y \notin B} \Pi_{z,y} < 1 \right) \right) \text{ for all } z \in S,$$

hence

$$(8.4.2) \quad \sum_{y \notin B} \Psi_{z,y} = \alpha(z) \sum_{y \notin B} \Pi_{z,y} < 1 \text{ for all } z \notin B.$$

We look at the corresponding homogeneous system of linear equations. Hence consider some $h \in \mathbb{R}^S$ with $0 = Ah$. It is obvious that for all $y \in B$ we have $h(y) = 0$. For all $z \notin B$ we have

$$\begin{aligned} h(z) &= \alpha(z) \sum_{y \in S} P(Z_1 = y | Z_0 = z) h(y) \\ &= \sum_{y \notin B} \alpha(z) P(Z_1 = y | Z_0 = z) h(y) \\ &= \sum_{y \notin B} \Psi_{z,y} h(y). \end{aligned}$$

We have for all $z \notin B$

$$\begin{aligned} |h(z)| &= \left| \sum_{y \notin B} \Psi_{z,y} h(y) \right| \leq \sum_{y \notin B} \Psi_{z,y} |h(y)| \\ &\leq \sum_{y \notin B} \Psi_{z,y} \max \{ |h(x)| ; x \notin B \} \\ &\leq \max \{ |h(x)| ; x \notin B \} \sum_{y \notin B} \Psi_{z,y}. \end{aligned}$$

Hence

$$\max \{ |h(x)| ; x \notin B \} \leq \max \{ |h(x)| ; x \notin B \} \sum_{y \notin B} \Psi_{z,y}.$$

Due to (8.4.2) we have

$$0 = \max \{ |h(x)| ; x \notin B \} \sum_{y \notin B} \Psi_{z,y}.$$

Hence $h \equiv 0$ and the solution for the homogenous system is unique and thus the solution for the inhomogeneous system is unique as well. \square

9 Numerical Examples

In the section we give some numerical examples for the Markovian case with constant discounting (see Part 6.3.3 (page 65)) and finite state space using the technique shown in Section 8 (page 80). At first this is done for $\kappa \equiv \{1\}$ in Part 9.2 (page 85) and then (using the definitions of Setting 7.2 (page 68)) for more arbitrary $\kappa : \mathbb{N} \rightarrow \mathcal{Pot}(\mathbb{N})$ in Part 9.3 (page 87).

9.1 Setting

The examples we focus on in this section are the same as in [Irl06, part 4], but here the technique used is different.

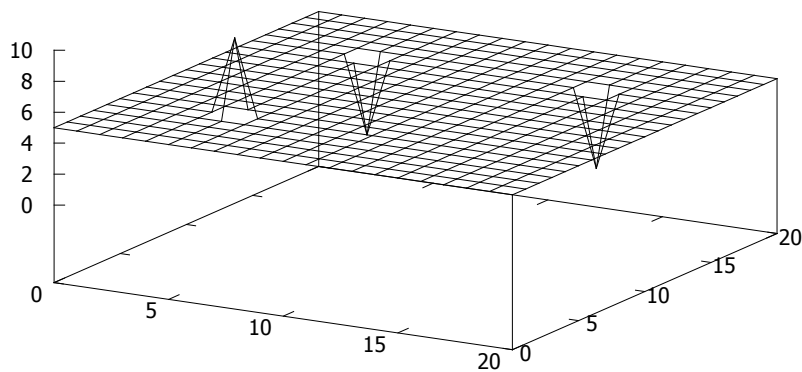
Consider $\{0, \dots, 20\} \times \{0, \dots, 20\}$ as the state space for the chain and a payoff function g with $g(z, n) = \alpha^n f(z)$ with $f((5, 5)) = 10$, $f((5, 15)) = f((15, 15)) = 0$, $f((x, y)) = 5$ otherwise. The transition probabilities are given in the form

$$p(x + 1, y \mid x, y) = 0.5 \cdot p_x, \quad p(x - 1, y \mid x, y) = 0.5 \cdot (1 - p_x),$$

$$p(x, y + 1 \mid x, y) = 0.5 \cdot p_y, \quad p(x, y - 1 \mid x, y) = 0.5 \cdot (1 - p_y),$$

with reflecting boundaries.

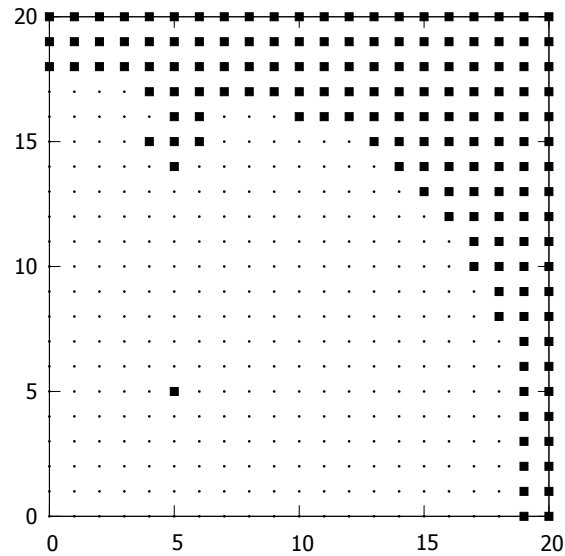
For better understanding the following graphic shows the undiscounted pay-offs:



Undiscounted Pay-Offs

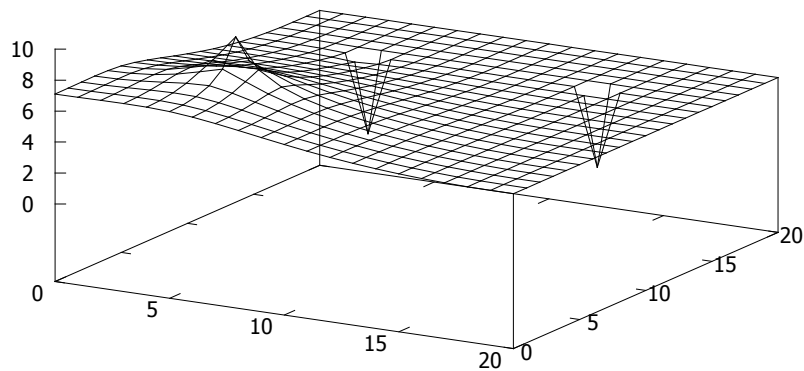
9.2 Results for $\kappa \equiv \{1\}$

With a discount-factor $\alpha = 0.98^{1/20}$ our program recommends a stopping rule visible in the next picture. The stopping points are big squares.



Stopping-Points of the best calculated stopping rule

Using this stopping rule our expected values are the optimal values shown in the next picture:



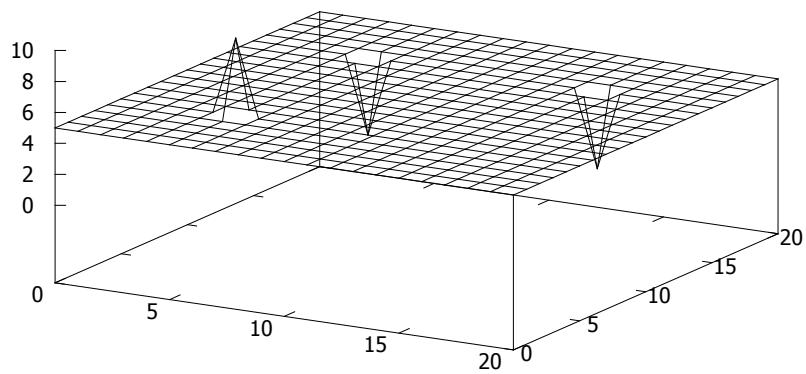
Expected Values

The number of iterations and the time for the calculation depends on κ , but this did not influence the result.

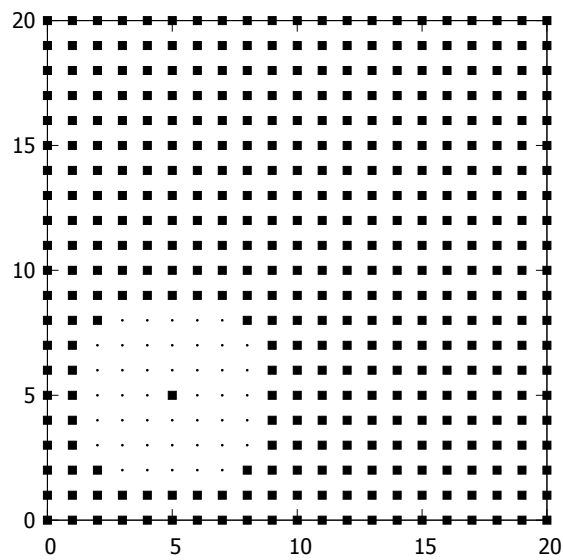
Nevertheless in other examples we found out that due to machine precision the powers of the transition matrix could not be calculated exactly enough, so even the results changed to too small or too big stopping areas.

For constant κ equivalent to $\{1\}$ the program is doing the problem above within 19 iterations and 0.16 seconds.

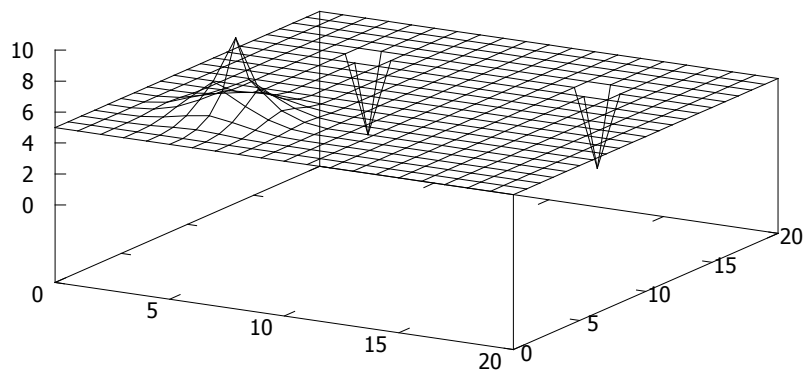
With $\alpha = 0.99$ and $\kappa \equiv \{1\}$ the following result appears within 6 iterations and 0.04 seconds.



Pay-Offs



Stopping Points of the Recommended Stopping Rule



Expected Values, when using the stopping rule above

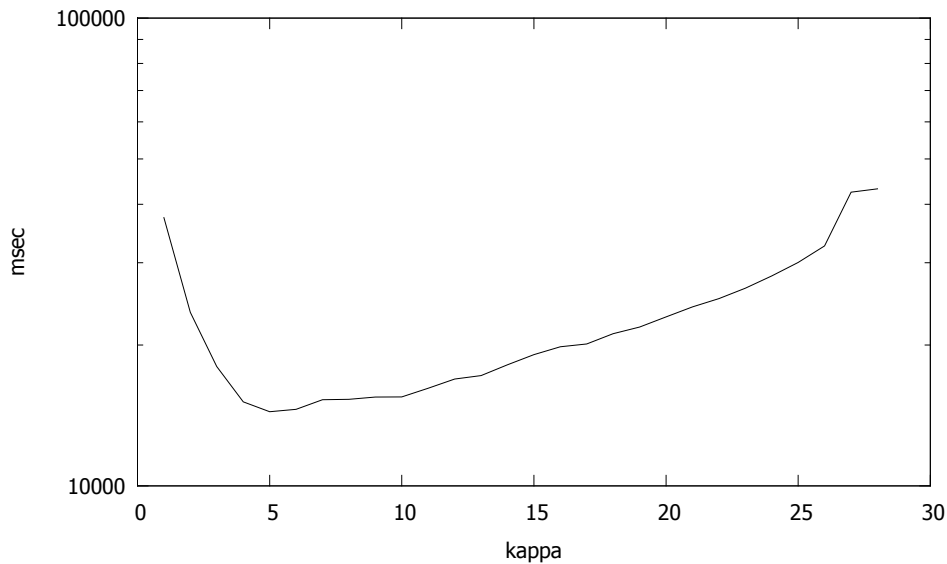
The results above are similar to those of Irlé shown in [Irl06, part 4], where they were calculated by simulation studies.

The speed of the example above is so high that the effect of changing κ cannot be examined therein. Thus we will increase the resolution by factor 10 in each dimension to study the effect of κ , hence we study a state space of $\{0, \dots, 200\} \times \{0, \dots, 200\}$ with pay-off 10 at $(50, 50)$, pay-off 0 at $(50, 150)$ and $(150, 150)$. We set $\alpha = 0.9999$ so that we have a trivial optimal stopping rule of stopping at $(50, 150)$, $(150, 150)$ and their four surrounding points as well as at $(50, 50)$ and continuing at all other points.

9.3 Influence of κ

Using the definitions of Setting 7.2 (page 68) we will examine more arbitrary $\kappa : \mathbb{N} \rightarrow \mathcal{Pot}(\mathbb{N})$.

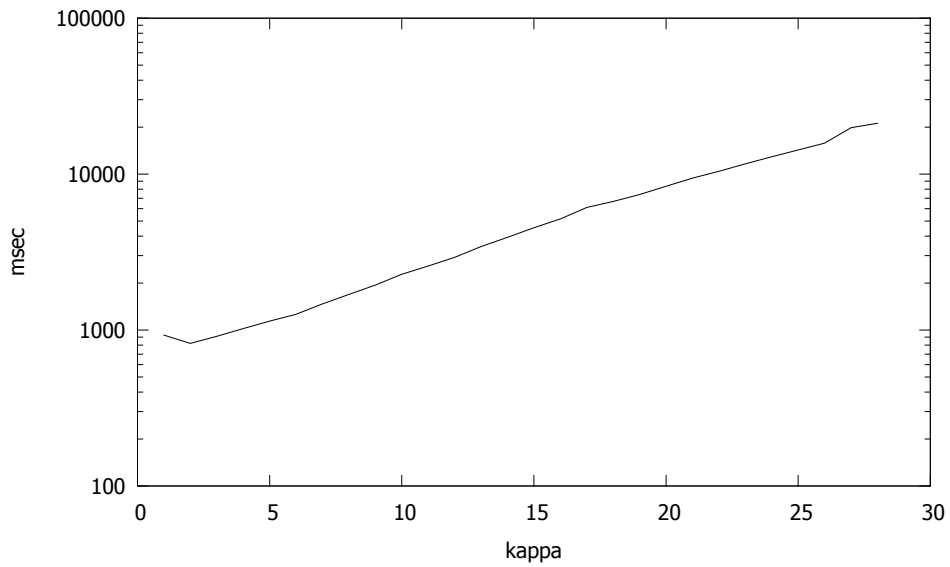
At first we will examine $\kappa \equiv \mathbb{N}_{\leq k}$ with some $k \in \mathbb{N}$ in the text and graphics called “kappa”.



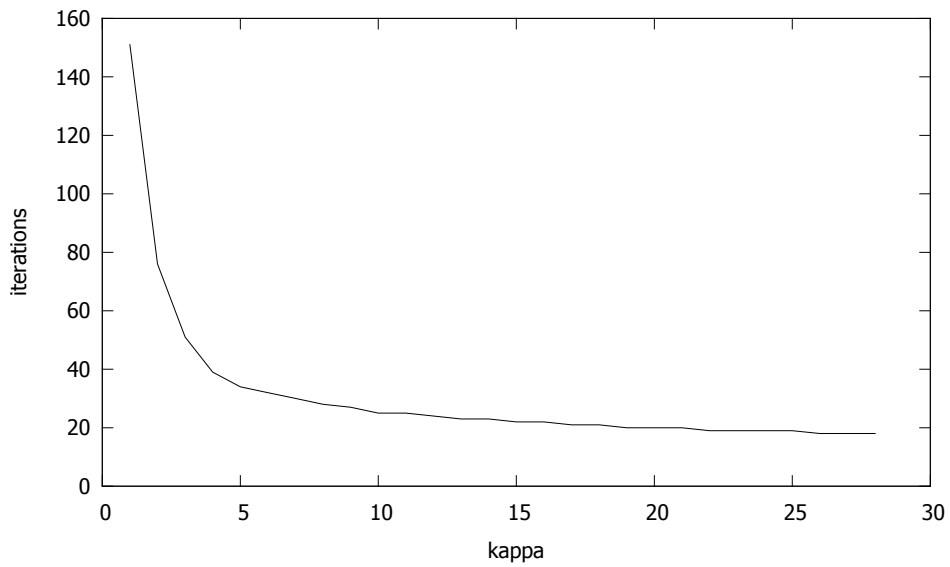
Runtime of the algorithm (in log-scale) plotted against kappa

The largest value of kappa we could examine was 28. At larger values the amount of space needed for storing the powers of the transition matrix explodes and exceeds the RAM-size of 1 GB. The algorithm slows significantly down due to hard disc accesses for reading the matrix powers.

We have a look at the time needed to calculate the powers of the transition matrix. There is a logarithmic dependency on kappa.

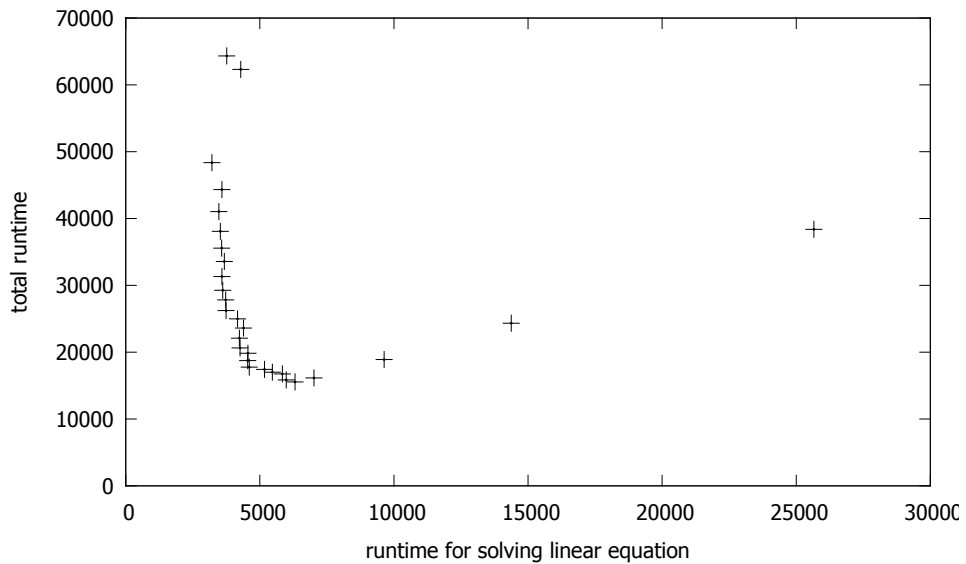


Runtime (in log-scale) for calculating powers of the transition matrix plotted against kappa.



Number of iterations plotted against kappa.

The number of iterations increases with kappa decreasing.



In each iteration we have to solve a linear equation independent of κ and hence independent of kappa. We can see in the example above a more or less linear dependency between the runtime for solving linear equation and the total runtime for the first five points when encountered from right to left. These points are related to kappa going from 1 to 5. The points on the left, more or less one above the other, are related to higher inefficient values of kappa. The number of iterations is strongly correlated with the runtime for solving linear equations and merely related to the exact problem.

We have also done some tests with more arbitrary $\kappa : \mathbb{N} \rightarrow \mathcal{Pot}(\mathbb{N})$. The time needed for doing the n th iteration does mostly depend on the size of $\kappa(n)$, the set itself does not really influence the runtime.

Consider the algorithm is doing n iterations. If $1 \notin \kappa(n)$ then in our tests the resulting stopping area was often larger than the optimal stopping area, since going one more step was better, but going more than one step was not, but this improvement could not be found by the algorithm, since $1 \notin \kappa(n)$. But given $1 \in \kappa(n)$ (and sufficiently high machine precision) in our tests the final result was always the optimal stopping area.

10 Monotone Case

This chapter examines the term “monotone case” which appears manifold in the literature. By FII Algorithm we can reprove the well known result that in the monotone case the myopic rules are optimal (under weak conditions) and we will enlarge this result to some more stopping rules appearing naturally within FII Algorithm, and adapt the result to the Markovian case.

In Definition 10.1 (page 90) we give an overview over common definitions in this field. We show in Remark 10.2 (page 91) that these terms naturally appear in our model. We enlarge the common terms by the concept of *m-monotone* and *m-stages look-ahead rules* in Definition 10.3 (page 91). Is an *m-monotone* problem also *p-monotone*? If $p \geq m$ this is true as shown in Lemma 10.4 (page 92), otherwise it is in general not true as shown in Lemma 10.6 (page 93). We show that *m-monotone* implies optimality of the *m-stages look-ahead rules* under weak conditions in Theorem 10.8 (page 94). In Part 10.10 (page 95) we give an adaption of this to the Markovian case. (In this chapter we concentrate on the infinite case, but the results can be easily transferred to the finite case.)

10.1 Definition

In [CRS71] and in [Irl79] a problem is called *monotone* iff

$$\{X_n \geq E(X_{n+1}|\mathcal{A}_n)\} \subseteq \{X_{n+1} \geq E(X_{n+2}|\mathcal{A}_{n+1})\} \text{ for all } n \in \mathbb{N}_0$$

and

$$\bigcup_{n \in \mathbb{N}_0} \{X_n \geq E(X_{n+1}|\mathcal{A}_n)\} = \Omega.$$

Further it is called *weakly monotone* in [Irl79] iff

$$\left\{ X_n \geq \sup_{k \geq 1} E(X_{n+k}|\mathcal{A}_n) \right\} \subseteq \left\{ X_{n+1} \geq \sup_{k \geq 1} E(X_{n+1+k}|\mathcal{A}_{n+1}) \right\} \text{ for all } n \in \mathbb{N}_0$$

and

$$\bigcup_{n \in \mathbb{N}_0} \left\{ X_n \geq \sup_{k \geq 1} E(X_{n+k}|\mathcal{A}_n) \right\} = \Omega.$$

The stopping rules

$$\min \{n \geq l; X_n \geq E(X_{n+1}|\mathcal{A}_n)\}, l \in \mathbb{N}_0$$

are called *one-stage look-ahead rules* or shortly *myopic rules*.

The stopping rules

$$\min \left\{ n \geq l; X_n \geq \sup_{k \geq 1} E(X_{n+k}|\mathcal{A}_n) \right\}, l \in \mathbb{N}_0$$

are called *one-time look-ahead rules*.

10.2 Remark: Appearance in the above Model

The definitions recapitulated in Definition 10.1 (page 90) can be found in the above model. Consider $C^{(0)} \equiv \Omega$ and $\tau^{(0)}$ and $C^{(0)}$ corresponding.

- $\tau^{(1)}$ depends on $\kappa(0)$. If $\kappa(0) = 1$, then it is a family of one-stage look-ahead rules, if $\kappa(0) = \infty$, then it is a family of one-time look-ahead rules.
- $C^{(1)}$ depends also on $\kappa(0)$.
Given $\kappa(0) = 1$ the problem is monotone iff

$$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega, \quad C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0.$$

Given $\kappa(0) = \infty$ the problem is weakly monotone iff

$$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega, \quad C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0.$$

10.3 Definition and Remark: *m*-monotone and *m*-stages look-ahead rules

The above remark gives reason to the following definition:

For all $m \in \mathbb{N} \cup \{\infty\}$ we will call a problem *m*-monotone iff

$$\left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} \subseteq \left\{ X_{n+1} \geq \sup_{m+1 > k \geq 1} E(X_{n+1+k} | \mathcal{A}_{n+1}) \right\} \text{ for all } n \in \mathbb{N}_0$$

and

$$\bigcup_{n \in \mathbb{N}_0} \left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} = \Omega.$$

With this definition the expressions “monotone” and “1-monotone” are equivalent as well as the definitions “weakly monotone” and “ ∞ -monotone”.

For all $m \in \mathbb{N} \cup \{\infty\}$ we will call the stopping rules

$$\min \left\{ n \geq l ; X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\}, l \in \mathbb{N}_0$$

m-stages look-ahead rules.

With this definition the expression “one-time look-ahead rules” becomes equivalent to “infinite-stages look-ahead rules”.

Consider $C^{(0)} \equiv \Omega$ and $\tau^{(0)}$ and $C^{(0)}$ corresponding.

Then the problem is $\kappa(0)$ -monotone iff

$$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega \quad \text{and} \quad C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0,$$

furthermore $\tau^{(1)}$ is a family of $\kappa(0)$ -stages look-ahead rules.

10.4 Lemma: m -monotone

Consider $m \in \mathbb{N}$.

If a problem is m -monotone, then for all $p \in \mathbb{N}_0 \cup \{\infty\}$ with $m \leq p$ we have

$$\left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} = \left\{ X_n \geq \sup_{p+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} \text{ for all } n \in \mathbb{N}_0$$

and hence the problem is p -monotone.

Proof: Consider an m -monotone problem, hence we have for all $n \in \mathbb{N}_0$

$$\left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} \subseteq \left\{ X_{n+1} \geq \sup_{m+1 > k \geq 1} E(X_{n+1+k} | \mathcal{A}_{n+1}) \right\}.$$

Consider $p \in \mathbb{N}_0 \cup \{\infty\}$ with $m \leq p$ and $n \in \mathbb{N}_0$.

The inclusion “ \supseteq ” is obvious.

Since the problem is m -monotone we have

$$X_{n+l} \geq E(X_{n+l+1} | \mathcal{A}_n) \text{ on } \left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} \text{ for all } l \in \mathbb{N}_0.$$

It follows by induction

$$X_n \geq E(X_{n+l} | \mathcal{A}_n) \text{ on } \left\{ X_n \geq \sup_{m+1 > k \geq 1} E(X_{n+k} | \mathcal{A}_n) \right\} \text{ for all } l \in \mathbb{N}_0.$$

Thus we have the inclusion “ \subseteq ”.

□

Hence a monotone problem (1-monotone) is weakly monotone (∞ -monotone).

10.5 Conclusion

Consider $C^{(0)} \equiv \Omega$ and $\tau^{(0)}$ and $C^{(0)}$ corresponding.

Consider $m \in \mathbb{N}$ and a m -monotone problem. Then we have for every $\kappa(0) \geq m$

$$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega, \quad C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0.$$

Proof: For every $\kappa(0) \geq m$ we know by Lemma 10.4 (page 92) that the problem is $\kappa(0)$ -monotone. The result follows by Definition 10.3 (page 91). □

10.6 Counter-Example: m -monotone

Consider $m, p \in \mathbb{N} \cup \{\infty\}$ with $p < m$.

Being m -monotone does in general not imply being p -monotone.

Proof: Set $l := m$ if $m < \infty$ and $l := p + 1$ otherwise.

Consider e.g. X with $X_1 = 0$, $X_{1+k} = -1$ for all $k \in \mathbb{N}$ with $k < l$, $X_{l+k} = \frac{1}{k}$ for all $k \in \mathbb{N}$. Hence X is pointwise constant and of the form $(0, -1, -1, \dots, -1, -1, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots)$.

For all $n \in \mathbb{N}$ with $n \leq l$ we have $n + 1 \leq l + 1 \leq m + 1 < n + m + 1$ and thus

$$\{0, -1\} \ni X_n < 1 = X_{l+1} \leq \sup_{m+1 > k \geq 1} X_{n+k} = \sup_{m+1 > k \geq 1} \mathbb{E}(X_{n+k} | \mathcal{A}_n),$$

hence

$$\left\{ X_n \geq \sup_{m+1 > k \geq 1} \mathbb{E}(X_{n+k} | \mathcal{A}_n) \right\} = \emptyset.$$

For all $n \in \mathbb{N}$ and $k \in \mathbb{N}$ with $k < m + 1$

$$X_{l+n} = \frac{1}{n} \geq \frac{1}{n+k} = X_{l+n+k},$$

thus

$$X_{l+n} \geq \sup_{m+1 > k \geq 1} X_{l+n+k} = \sup_{m+1 > k \geq 1} \mathbb{E}(X_{l+n+k} | \mathcal{A}_{l+n}),$$

hence

$$\left\{ X_{l+n} \geq \sup_{m+1 > k \geq 1} \mathbb{E}(X_{l+n+k} | \mathcal{A}_{l+n}) \right\} = \Omega.$$

Obviously the problem is m -monotone.

We have for all $k \in \mathbb{N}$ with $k < m + 1$

$$\{0, -1\} \ni X_{l-p} \geq -1 = X_{l-p+k},$$

hence

$$\sup_{p+1 > k \geq 1} X_{l-p+k} = \sup_{p+1 > k \geq 1} \mathbb{E}(X_{l-p+k} | \mathcal{A}_{l-p})$$

and

$$X_{l-p+1} = -1 < 1 = X_{l+1} = X_{l-p+1+p} = \sup_{p+1 > k \geq 1} X_{l-p+1+k} = \sup_{p+1 > k \geq 1} \mathbb{E}(X_{l-p+1+k} | \mathcal{A}_{l-p+1}),$$

hence

$$\Omega = \left\{ X_{l-p} \geq \sup_{p+1 > k \geq 1} \mathbb{E}(X_{l-p+k} | \mathcal{A}_{l-p}) \right\}$$

and

$$\emptyset = \left\{ X_{l-p+1} \geq \sup_{p+1 > k \geq 1} \mathbb{E}(X_{l-p+1+k} | \mathcal{A}_{l-p+1}) \right\}.$$

Since $\Omega \not\subseteq \emptyset$ the problem is not p -monotone. □

10.7 Lemma

Choose any $\kappa(0)$ and $\kappa(1)$. ($C^{(1)}$ depends on $\kappa(0)$ and $C^{(2)}$ on $\kappa(0)$ and $\kappa(1)$.)

If

$$C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0,$$

then we have

$$C^{(1)} = C^{(2)}.$$

Proof: Assume $C_n^{(1)} \subseteq C_{n+1}^{(1)}$ for all $n \in \mathbb{N}_0$. Obviously we have $C_n^{(1)} \subseteq C_{n+k}^{(1)}$ for all $n, k \in \mathbb{N}_0$. Consider $n \in \mathbb{N}_0$. Since $\tau^{(1)}$ and $C^{(1)}$ are corresponding we have by Theorem 2.3 (page 19)

$$C_n^{(1)} \subseteq C_{n+k}^{(1)} = \left\{ \tau_{n+k}^{(1)} = n+k \right\} \text{ for all } k \in \mathbb{N}.$$

Hence on $C_n^{(1)}$ we have

$$\mathbb{E} \left(X_{\tau_{n+k}^{(1)}} \middle| \mathcal{A}_n \right) = \mathbb{E} (X_{n+k} | \mathcal{A}_n) \text{ for all } k \in \mathbb{N}.$$

We have thus

$$\begin{aligned} C_n^{(2)} &= C_n^{(1)} \cap \bigcap \left\{ \left\{ \mathbb{E} \left(X_{\tau_p^{(1)}} \middle| \mathcal{A}_n \right) \leq X_n \right\} ; n \leq p < \min \{n + \kappa(1), N\} + 1 \right\} \\ &= C_n^{(1)} \cap \bigcap \left\{ \left\{ \mathbb{E} (X_p | \mathcal{A}_n) \leq X_n \right\} ; n \leq p < \min \{n + \kappa(1), N\} + 1 \right\} \\ &= C_n^{(1)} \cap \Omega \cap \bigcap \left\{ C_p^{(1)} ; n+1 \leq p < \min \{n + \kappa(1), N\} + 1 \right\} \\ &= C_n^{(1)}. \end{aligned} \quad \square$$

10.8 Theorem

Consider $m \in \mathbb{N} \cup \{\infty\}$, a m -monotone problem and Assumption 5.2 (page 50).

Then for each $p \in \mathbb{N} \cup \{\infty\}$ with $p \geq m$ the family of p -stages look-ahead rules is optimal.

Proof: Consider $C^{(0)} \equiv \Omega$ and $\tau^{(0)}$ and $C^{(0)}$ corresponding and $\kappa(0) := \kappa(1) := p$. By Conclusion 10.5 (page 92) we have

$$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega, \quad C_n^{(1)} \subseteq C_{n+1}^{(1)} \text{ for all } n \in \mathbb{N}_0.$$

Then we have by Lemma 10.7 (page 94)

$$C^{(1)} = C^{(2)}.$$

By Remark 5.14 (page 61) we have the optimality of $\tau^{(\infty)}$.

Further we have by Conclusion 5.12 (page 60)

$$\tau^{(1)} = \tau^{(\infty)}.$$

□

10.9 Lemma

$\bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega$ implies $\tau_n^{(1)}$ finite for all $n \in \mathbb{N}_0$.

Proof: Assume there is $n \in \mathbb{N}_0$ and $\omega \in \Omega$ with $\tau_n^{(1)}(\omega) = \infty$. We have

$$\infty = \tau_n^{(1)}(\omega) = \inf \left\{ p \in \mathbb{N}_0 ; \mathbf{1}_{C_p^{(1)}}(\omega) = 1 \right\} = \inf \left\{ p \in \mathbb{N}_0 ; \omega \in C_p^{(1)} \right\},$$

hence

$$\emptyset = \left\{ p \in \mathbb{N}_0 ; \omega \in C_p^{(1)} \right\}.$$

So we have $\omega \notin \bigcup_{n \in \mathbb{N}_0} C_n^{(1)} = \Omega$, a contradiction. □

10.10 Markov Case

In Section 6 (page 62) we discussed the Markovian case. As in that section let $(Z_n)_{n \in \mathbb{N}_0}$ be a time-homogeneous Markov process with respect to the underlying filtration with state space (S, \mathcal{S}) and the FII algorithm only has to work in the systems of subsets of the state space S .

Assume further that we have a payoff $(X_n)_{n \in \mathbb{N}_0}$ such that our improvement step can be done in the way of Setting 7.2 (page 68) for all $k \in \mathbb{N}$ by

$$S^{*k} \stackrel{\text{def.}}{=} \left\{ z \in S ; g(z) \geq \sup \{ E(X_l | Z_0 = z) ; l \in \mathbb{N}_{\leq k} \} \right\}.$$

This is the case in all examples within Section 6 (page 62).

For all $k \in \mathbb{N} \cup \{\infty\}$ we will call a problem k -monotone iff

$$P \left(Z_1 \in S^{*k} \mid Z_0 = z \right) = 1 \text{ for all } z \in S^{*k},$$

and the k -stages look-ahead rule is $\tau_0(S^{*k})$.

Consider $k \in \mathbb{N}$. The results shown above for the general case in Lemma 10.4 (page 92) and Lemma 10.6 (page 93) can be easily transferred to the Markovian case: A problem being k -monotone is l -monotone if $l \geq k$ and in general not l -monotone if $l < k$. If a problem is k -monotone then we have $P(\tau_0(S^{*k}) = 1 | Z_0 = z) = 1$ for all $z \in S^{*k}$ and hence $\tau_0(S^{*k})$ is optimal (given conditions (7.2.1) and (7.4.6)).

11 Example: No-Information Version of the Best Choice Problem with Random Population Size

In this chapter we consider the best choice problem, also called secretary, beauty contest, dowry or marriage problem. We consider the no-information version with a random number of observations. Given these constraints the problem was posed first by Presman and Sonin, see [PS73]. Here we use a setting based on and compatible with that in [Irl80, part 1]. Neither recall nor uncertainty of selection is allowed, one choice must be made. The relative ranks of the items are observed sequentially with the object of selecting the absolut best one.

First we give the setting in Part 11.1 (page 96) and Part 11.2 (page 97). In [Irl80, part 3] we find some examples of how the FII Algorithm works in this setting. *We found out that these examples have a common structure, which is general and independent of the examples.*

We will show in Theorem 11.13 (page 104) that for each $k \in \mathbb{N}_0$ the $C^{(k)} \in \mathcal{Pot}(\Omega)^{\mathbb{N}}$ of the algorithm can be expressed by some $B^{(k)} \subseteq \mathbb{N}$ (when starting with $C^{(0)} := \Omega$), a noteworthy simplification.

We will further show (in Lemma 11.10 (page 101)) that the algorithmic step of going from some C to C^{*k} by calculating certain conditional expectations (as defined in Lemma 5.10 (page 53)) can be done by going from B to B^{*k} by direct evaluation of a function given in an analytic form (defined in Setting 11.3 (page 97)).

11.1 General Setting

Consider some $p \in [0, 1]^{\mathbb{N}}$ with $\sum_{i=1}^{\infty} p_i = 1$.
For all $k \in \mathbb{N}$ define

$$\Omega_k := \{\sigma : \mathbb{N}_{\leq k} \longrightarrow \mathbb{N}_{\leq k} ; \sigma \text{ bijective}\}.$$

Define $\Omega := \bigcup_{k \in \mathbb{N}} \Omega_k$, $\mathcal{A} := \mathcal{Pot}(\Omega)$ and $P : \mathcal{A} \longrightarrow [0, 1]$ by

$$P(\{\omega\}) = \frac{pk}{k!}$$

for all $\omega \in \Omega_k$ and $k \in \mathbb{N}$.

Define $N : \Omega \longrightarrow \mathbb{N}$ as the mapping of each ω onto the unique $k \in \mathbb{N}$ with $\omega \in \Omega_k$.

The maximal possible size of the population is of importance for the following considerations, hence define

$$n_0 := \inf \{n \in \mathbb{N} ; P(\{N \leq n\}) = 1\}$$

and use the convention that the infimum of an empty set is infinity, hence

$$n_0 \in \mathbb{N} \cup \{\infty\}.$$

For all $i \in \mathbb{N}$ define

$$A_i : \Omega \longrightarrow \mathbb{N}, \omega \longmapsto \begin{cases} \omega_i & \text{if } k \in \mathbb{N} \text{ exists with } \omega \in \Omega_k \text{ and } i \leq k \\ \infty & \text{otherwise,} \end{cases}$$

and

$$R_i : \Omega \longrightarrow \mathbb{N}, \omega \longmapsto \begin{cases} |\{j \in \mathbb{N} ; j \leq i, \omega_j \leq \omega_i\}| & \text{if } k \in \mathbb{N} \text{ exists with } \omega \in \Omega_k \text{ and } i \leq k \\ \infty & \text{otherwise.} \end{cases}$$

For all $k \in \mathbb{N}$ define $\mathcal{A}_k := \sigma(\{R_i ; i \in \mathbb{N}_{\leq k}\})$.

For all $k \in \mathbb{N}$ let $q_k : \mathbb{N} \cup \{\infty\} \longrightarrow \mathbb{R}_{\geq 0}$ with $q_k(\infty) = 0$.

For all $k \in \mathbb{N}$ let $X_k := E(q_k(A_k) | \mathcal{A}_k)$

N is the number of available elements.

$q_k(i)$ is the payoff you get, if you stop at the k th element and this has the absolute rank i .

R_i is the relative rank and A_i the absolute rank of the i th element.

They are both infinite iff the number of available elements is strictly smaller than i .

11.2 Special Setting

Consider some $d \in]0, 1]$.

For all $k \in \mathbb{N}$ and $i \in \mathbb{N} \cup \{\infty\}$ let $q_k(i) := \delta_{i,1} d^k$.

Here the best choice is maximizing the possibility of having the best of all possible candidates.

11.3 Definitions and Remarks

- Define

$$\varphi : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{R}_{\geq 0}, (n, r) \longmapsto \begin{cases} \frac{n}{P(N \geq n)} \sum_{i=n}^{n_0} \frac{p_i}{i} & \text{if } P(N \geq n) > 0 \text{ and } r = 1 \\ 0 & \text{otherwise.} \end{cases}$$

We have by Lemma 11.4 (page 98) below

$$X_k = E(q_k(A_k) | \mathcal{A}_k) = d^k P(A_k = 1 | \mathcal{A}_k) = d^k \varphi(k, R_k) \text{ for all } k \in \mathbb{N}.$$

- Since $\bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} \{R_k = \infty\} = \Omega$ we have $0 = \lim_{n \rightarrow \infty} X_n$ pointwise, so by defining

$$X_\infty := 0$$

we have $X_\infty = \lim_{n \rightarrow \infty} \sup_{k \geq n} X_k$.

By Lemma 5.6 (page 51) the assumptions (5.2.1) and (5.2.2) are true.

- Define stopping areas for all $B \subseteq \mathbb{N}_{< n_0}$ and $n \in \mathbb{N} \cup \{\infty\}$ by

$$C_n(B) := \begin{cases} \{R_n \in \{1, \infty\}\} & \text{if } n \in B \\ \Omega & \text{if } n \geq n_0 \\ \{R_n = \infty\} & \text{otherwise.} \end{cases}$$

In the examples examined in [Irl80, part 3] the appearing stopping areas were all of this kind. We will show in the sequel that this was not due to the examples, it's a general result for this best choice problem.

Define a sequence of stopping areas for all $B \subseteq \mathbb{N}_{< n_0}$ by

$$C(B) := (C_n(B))_{n \in \mathbb{N} \cup \{\infty\}}.$$

Hence we have constructed for each $B \subseteq \mathbb{N}_{<n_0}$ some $C(B) \in \mathcal{C}$. In Theorem 11.13 (page 104) we show that for each $k \in \mathbb{N}$ there is some $B^{(k)} \subseteq \mathbb{N}_{<n_0}$ such that $C^{(k)} = C(B^{(k)})$ (when starting with $C^{(0)} := \Omega$).

- Using the same definitions as in Definition 5.8 (page 52) we have for all $B \subseteq \mathbb{N}_{<n_0}$ and $n \in \mathbb{N}_0$

$$\begin{aligned} \tau_n(C(B)) &= (n \vee n_0) \wedge \inf \{l \in \mathbb{N} \cap [n, n_0[; R_l = \infty\} \\ &\quad \wedge \inf \{l \in B \cap [n, n_0[; R_l = 1\}, \end{aligned}$$

and thus

$$\begin{aligned} \tau_1(C(B)) &= n_0 \wedge \inf \{l \in \mathbb{N}_{<n_0} ; R_l = \infty\} \\ &\quad \wedge \inf \{l \in B ; R_l = 1\}. \end{aligned}$$

- As mentioned above the algorithmic step can be done by calculating for some B a $B^{*\kappa}$ by using a deterministic function in analytical form as follows.

Define

(using the convention that an empty sum is zero and an empty product is one)

$$\begin{aligned} a : \mathbb{N}_{<n_0} \times [0, 1] \times \text{Pot}(\mathbb{N}_{<n_0}) \times \mathbb{N} &\longrightarrow \mathbb{R}, \\ (n, d, B, j) &\longmapsto \sum_{l=n}^{n_0} \frac{pl}{l} \left(nd^n - \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} \right). \end{aligned}$$

Hence for all $n \in \mathbb{N}_{<n_0}$ and $j \in \mathbb{N}$ we have

$$(11.3.1) \quad a(n, d, \mathbb{N}_{<n_0}, j) = \sum_{l=n}^{n_0} \frac{pl}{l} \left(nd^n - (n+j-1) \sum_{i=n+j}^l \frac{d^i}{i-1} \right).$$

Define¹ further for all $\kappa \in \mathbb{N} \cup \{\infty\}$

(using the convention that the infimum of an empty set is infinity)

$$\begin{aligned} a_\kappa : \mathbb{N}_{<n_0} \times [0, 1] \times \text{Pot}(\mathbb{N}_{<n_0}) &\longrightarrow \mathbb{R} \cup \{-\infty, \infty\}, \\ (n, d, B) &\longmapsto \inf \{a(n, d, B, j) ; j \in \mathbb{N} \text{ with } j < \kappa + 1\}. \end{aligned}$$

Define for all $\kappa \in \mathbb{N} \cup \{\infty\}$ and $B \subseteq \mathbb{N}_{<n_0}$

$$B^{*\kappa} := \{n \in B ; a_\kappa(n, d, B) \geq 0\}.$$

11.4 Lemma

For all $n \in \mathbb{N}$ we have

$$P(A_n = 1 | \mathcal{A}_n) = P(A_n = 1 | R_n = 1) \cdot \mathbf{1}_{\{R_n=1\}}$$

and if $P(N \geq n) > 0$ we also have

$$P(A_n = 1 | R_n = 1) = \frac{n}{P(N \geq n)} \sum_{i=n}^{n_0} \frac{p_i}{i}.$$

Proof: A well-known result of the theory of best-choice problems. □

¹ In [Ir180, part 3] this function had been used and defined only for $B = \mathbb{N}_{<n_0}$ and $\kappa = 1$. The expressions $a_1(n, d, \mathbb{N}_{<n_0}, 1)$ and $a(n, d, \mathbb{N}_{<n_0}, 1)$ herein are equal to $n \cdot c_n(d)$ therein.

11.5 Lemma

Let $n \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$, $j \in \mathbb{N}$. We have

$$\mathbb{E} \left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n \right) = \sum_{\substack{i \in B \cup \{n_0\} \\ i \geq n+j}} d^i \varphi(i, 1) \cdot P(R_i = 1, R_k \neq 1 \text{ for all } k \in B \cap [n+j, i] \middle| \mathcal{A}_n).$$

Proof: We have for all $i \in \mathbb{N}$ with $n+j \leq i \leq n_0$

$$\begin{aligned} & \mathbb{E} \left(X_i \mathbf{1}_{\{i=\tau_{n+j}(C(B))\}} \middle| \mathcal{A}_n \right) \\ &= \mathbb{E} \left(d^i \varphi(i, R_i) \mathbf{1}_{\{i=\tau_{n+j}(C(B))\}} \middle| \mathcal{A}_n \right) \\ &= \mathbb{E} \left(d^i \varphi(i, 1) \mathbf{1}_{\{R_i=1\}} \mathbf{1}_{\{i=\tau_{n+j}(C(B))\}} \middle| \mathcal{A}_n \right) \\ &= d^i \varphi(i, 1) \cdot P(i = \tau_{n+j}(C(B)), R_i = 1 \middle| \mathcal{A}_n) \\ &= d^i \varphi(i, 1) \cdot P(i = \inf \{p \in B \cup \{n_0\} ; p \geq n+j, R_p = 1\}, R_i = 1 \middle| \mathcal{A}_n) \\ &= \begin{cases} d^i \varphi(i, 1) \cdot P(R_i = 1, R_k \neq 1 \text{ for all } k \in B \cap [n+j, i] \middle| \mathcal{A}_n) & \text{if } i \in B \cup \{n_0\} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Hence we have

$$\begin{aligned} & \mathbb{E} \left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n \right) \\ &= \mathbb{E} \left(\sum_{i=n+j}^{\infty} X_i \mathbf{1}_{\{i=\tau_{n+j}(C(B))\}} \middle| \mathcal{A}_n \right) \\ &= \sum_{\substack{i \in B \cup \{n_0\} \\ i \geq n+j}} d^i \varphi(i, 1) \cdot P(R_i = 1, R_k \neq 1 \text{ for all } k \in B \cap [n+j, i] \middle| \mathcal{A}_n). \quad \square \end{aligned}$$

11.6 Lemma

For all finite sets $I \subseteq \mathbb{N}$ and $b \in \mathbb{N}^I$ with $b_i \leq i$ for all $i \in I$ we have

$$P(\forall i \in I \ R_i = b_i \mid N \geq \max I) = \prod_{i \in I} \frac{1}{i}.$$

11.7 Lemma

For all finite sets $I \subseteq \mathbb{N}$ and $D \in \text{Pot}(\mathbb{N})^I$ with $D_i \subseteq \mathbb{N}_{\leq i}$ for all $i \in I$ we have

$$P(\forall i \in I \ R_i \in D_i \mid N \geq \max I) = \prod_{i \in I} \frac{|D_i|}{i}.$$

Proof: By Lemma 11.6 (page 99). □

11.8 Lemma

Let $B \subseteq \mathbb{N}_{<n_0}$, $n, i \in \mathbb{N}$ with $n < i \leq n_0$, $j \in \mathbb{N}$. Then we have

$$P(R_i = 1 \wedge \forall k \in B \cap [n+j, i[\ R_k \neq 1 | R_n = 1) = \frac{P(N \geq i)}{i} \frac{1}{P(N \geq n)} \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k}.$$

Proof: We have

$$P(R_n = 1) = P(R_n < \infty) \cdot P(R_n = 1 | R_n < \infty) = \frac{1}{n} P(N \geq n)$$

and

$$\begin{aligned} & P(R_i = 1 \wedge \forall k \in B \cap [n+j, i[\ R_k \neq 1 \wedge R_n = 1) \\ &= P(R_i < \infty) \cdot P(R_i = 1 \wedge R_n = 1 \wedge \forall k \in B \cap [n+j, i[\ R_k \neq 1 | R_i < \infty) \\ &\stackrel{11.7}{=} P(R_i < \infty) \cdot \left(\frac{1}{i} \frac{1}{n} \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} \right) \\ &= \frac{1}{n} \frac{P(N \geq i)}{i} \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k}. \end{aligned} \quad \square$$

11.9 Lemma

Let $B \subseteq \mathbb{N}_{<n_0}$, $n \in B$, $j \in \mathbb{N}$, $\omega \in \{R_n = 1\}$. We have

$$\left(X_n - \mathbb{E} \left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n \right) \right) (\omega) \geq 0 \iff a(n, d, B, j) \geq 0.$$

Proof: For any sequence $(c_i)_{i \in \mathbb{N}}$ we have

$$\begin{aligned} & \sum_{i \in (B \cup \{n_0\}) \cap [n+j, n_0]} d^i \left(\varphi(i, 1) \cdot \frac{P(N \geq i)}{i} \right) c_i \\ (11.9.1) \quad & \stackrel{11.3}{=} \sum_{i \in (B \cup \{n_0\}) \cap [n+j, n_0]} d^i \left(\sum_{l=i}^{n_0} \frac{p_l}{l} \right) c_i \\ & = \sum_{l=n}^{n_0} \frac{p_l}{l} \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i c_i. \end{aligned}$$

Since $n < n_0$ we have $P(N \geq n) > 0$ and hence we have

$$\begin{aligned} & \left(X_n - \mathbb{E} \left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n \right) \right) (\omega) \\ & \stackrel{11.5}{=} d^n \varphi(n, 1) - \sum_{i \in (B \cup \{n_0\}) \cap [n+j, n_0]} d^i \varphi(i, 1) \cdot P(R_i = 1, \forall k \in B \cap [n+j, i[\ R_k \neq 1 | \mathcal{A}_n) (\omega) \\ & \stackrel{11.8}{=} d^n \varphi(n, 1) - \sum_{i \in (B \cup \{n_0\}) \cap [n+j, n_0]} d^i \varphi(i, 1) \cdot \frac{P(N \geq i)}{i} \frac{1}{P(N \geq n)} \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} \\ (11.9.1) \quad & \stackrel{(11.9.1)}{=} \frac{1}{P(N \geq n)} \sum_{l=n}^{n_0} \frac{p_l}{l} \left[n d^n - \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} \right] \\ (11.3) \quad & \stackrel{(11.3)}{=} \frac{1}{P(N \geq n)} a(n, d, B, j). \end{aligned} \quad \square$$

11.10 Theorem²

For all $B \subseteq \mathbb{N}_{<n_0}$, $\kappa \in \mathbb{N} \cup \{\infty\}$ we have

$$(C(B))^{*\kappa} = C(B^{*\kappa}),$$

where the first $^{*\kappa}$ is defined in Definition 5.8 (page 52), while the latter $^{*\kappa}$ as well as the mapping C is defined in Setting 11.3 (page 97).

Proof: Let $B \subseteq \mathbb{N}_{<n_0}$ and $n \in \mathbb{N}$.

On $\{R_{n+1} = \infty\}$ we have $X_k = 0$ for all $k \in \mathbb{N}$ with $k > n$, hence we have $\tau_{n+j}(C(B)) = n + j > n$ for all $j \in \mathbb{N}$ and thus

$$X_{\tau_{n+j}(C(B))} = 0 \text{ and } X_n - \mathbb{E}\left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n\right) \geq 0 \text{ for all } j \in \mathbb{N}.$$

This shows

$$(11.10.1) \quad \{R_n = \infty\} \subseteq \{R_{n+1} = \infty\} \subseteq (C(B))_n^{*\kappa}.$$

Case of $n \geq n_0$:

We have $C_n(B) = \Omega$ and $\{R_{n+1} = \infty\} = \Omega$. Hence

$$(C(B))_n^{*\kappa} = \Omega.$$

Case of $n \in B$:

We have $C_n(B) = \{R_n \in \{1, \infty\}\}$.

For all $\omega \in \{R_n = 1\}$ we have by Lemma 11.9 (page 100) the following equivalence:

$$\begin{aligned} & \omega \in (C(B))_n^{*\kappa} \\ \stackrel{\text{def.}}{\iff} & \left(X_n - \mathbb{E}\left(X_{\tau_{n+j}(C(B))} \middle| \mathcal{A}_n\right) \right) (\omega) \geq 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa + 1 \\ \stackrel{11.9}{\iff} & a(n, d, B, j) \geq 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa + 1 \\ \stackrel{\text{by def.}}{\iff} & a_\kappa(n, d, B) \geq 0 \\ \stackrel{\text{def.}}{\iff} & n \in B^{*\kappa}. \end{aligned}$$

Hence we have

$$\{R_n = 1\} \subseteq (C(B))_n^{*\kappa} \iff n \in B^{*\kappa}$$

and

$$\{R_n = \infty\} \stackrel{(11.10.1)}{\subseteq} (C(B))_n^{*\kappa} \stackrel{\text{by def.}}{\subseteq} C_n(B) \stackrel{\text{def.}}{=} \{R_n \in \{1, \infty\}\}.$$

Hence we have by

$$(C(B))_n^{*\kappa} = \begin{cases} \{R_n \in \{1, \infty\}\} & \text{if } n \in B^{*\kappa} \\ \{R_n = \infty\} & \text{otherwise.} \end{cases}$$

Case of $n < n_0$ and $n \notin B$:

We have

$$C_n(B) \stackrel{\text{def.}}{=} \{R_n = \infty\} \stackrel{(11.10.1)}{\subseteq} (C(B))_n^{*\kappa} \stackrel{\text{by def.}}{\subseteq} C_n(B),$$

hence

$$(C(B))_n^{*\kappa} = \{R_n = \infty\}. \quad \square$$

²For $\kappa = 1$ and $B = \mathbb{N}_{<n_0}$ this can be found in [Irl80, 3.2] with $n \cdot c_n(d) = a(n, d, \mathbb{N}_{<n_0}, 1) = a(n, d, \mathbb{N}_{<n_0})$.

11.11 Theorem³

Assume $C_n^{(0)} := \Omega$ and $\tau_n^{(0)} := n$ for all $n \in \mathbb{N} \cup \{\infty\}$.

Then we have

$$C^{(1)} = C(\mathbb{N}_{<n_0}),$$

where the left-hand side is defined in Algorithm 2.13 (page 24) and the mapping C on the right in Setting 11.3 (page 97).

Proof: It is enough to show that

- for all $n \in \mathbb{N}_{<n_0}$ we have $C_n^{(1)} = \{R_n \in \{1, \infty\}\}$ and
- for all $n \in \mathbb{N}_{\geq n_0}$ we have $C_n^{(1)} = \Omega$.

Define $B := \mathbb{N}_{<n_0}$. First let $n \in \mathbb{N}_0$ with $n_0 \leq n$.

Since $R_k \equiv \infty$ for all $k \in \mathbb{N}_0$ with $n_0 + 1 \leq k$ follows

$$\mathbb{E}\left(X_{\tau_{n+j}^{(0)}} \middle| \mathcal{A}_n\right) = \mathbb{E}(X_{n+j} | \mathcal{A}_n) = 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa(1) + 1,$$

hence

$$X_n - \mathbb{E}\left(X_{\tau_{n+j}^{(0)}} \middle| \mathcal{A}_n\right) = X_n \geq 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa(1) + 1.$$

Thus we have $C_n^{(1)} = \Omega$ for all $n \in \mathbb{N}_0$ with $n_0 \leq n$.

Next let $n \in \mathbb{N}$ with $n < n_0$.

We have $P(N \geq n) > 0$.

On $\{R_n = \infty\}$ we have $R_k = \infty$ for all $k \in \mathbb{N}_0$ with $n \leq k$.

Hence

$$X_n - \mathbb{E}\left(X_{\tau_{n+j}^{(0)}} \middle| \mathcal{A}_n\right) \geq 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa(k) + 1.$$

Thus we have $\{R_n = \infty\} \subseteq C_n^{(1)}$.

Now consider some $\omega \in \{R_n < \infty\}$ and some $j \in \mathbb{N}$ with $j < \kappa(k) + 1$.

If $n + j > n_0$, then we have $\mathbb{E}(X_{n+j} | \mathcal{A}_n)(\omega) = 0$, hence $(X_n - \mathbb{E}(X_{n+j} | \mathcal{A}_n))(\omega) \geq 0$.

So assume $n + j \leq n_0$. Then we have

$$\begin{aligned} & \mathbb{E}(\varphi(n+j, R_{n+j}) | \mathcal{A}_n)(\omega) \\ &= \varphi(n+j, 1)P(R_{n+j} = 1 | \mathcal{A}_n)(\omega) \\ &= \varphi(n+j, 1)P(R_{n+j} = 1 | R_n < \infty) \\ &= \varphi(n+j, 1)P(R_{n+j} = 1 | R_{n+j} < \infty) \frac{P(R_{n+j} < \infty)}{P(R_n < \infty)} \\ &= \left[\frac{n+j}{P(N \geq n+j)} \sum_{i=n+j}^{n_0} \frac{p_i}{i} \right] \frac{1}{n+j} \frac{P(N \geq n+j)}{P(N \geq n)} \\ &= \frac{1}{P(N \geq n)} \sum_{i=n+j}^{n_0} \frac{p_i}{i}, \end{aligned}$$

³For $\kappa = 1$ this can be found in [Irl80, 3.1].

hence

$$\begin{aligned}
(11.11.1) \quad \mathbb{E}(X_{n+j}|\mathcal{A}_n)(\omega) &= d^{n+j}\mathbb{E}(\varphi(n+j, R_{n+j})|\mathcal{A}_n)(\omega) \\
&= \frac{1}{P(N \geq n)} d^{n+j} \sum_{i=n+j}^{n_0} \frac{p_i}{i} \\
&= \frac{d^n}{P(N \geq n)} d^j \sum_{i=n+j}^{n_0} \frac{p_i}{i}.
\end{aligned}$$

If $\omega \in \{R_n \neq 1\} \cap \{R_n < \infty\}$ we have $X_n(\omega) \stackrel{11.3}{=} d^n \varphi(n, R_n)(\omega) = 0$ and $\mathbb{E}(X_{n+j}|\mathcal{A}_n)(\omega) > 0$, hence

$$\left(X_n - \mathbb{E}\left(X_{\tau_{n+j}^{(0)}} \middle| \mathcal{A}_n \right) \right) (\omega) = (X_n - \mathbb{E}(X_{n+j}|\mathcal{A}_n))(\omega) < 0.$$

On $\{R_n = 1\}$ we have

$$\begin{aligned}
X_n &\stackrel{11.3}{=} d^n \varphi(n, 1) \\
&= \frac{nd^n}{P(N \geq n)} \left(\sum_{i=n}^{n+j-1} \frac{p_i}{i} + \sum_{i=n+j}^{n_0} \frac{p_i}{i} \right) \\
&= \frac{d^n}{P(N \geq n)} \left(n \sum_{i=n}^{n+j-1} \frac{p_i}{i} + n \sum_{i=n+j}^{n_0} \frac{p_i}{i} \right)
\end{aligned}$$

and hence

$$\frac{P(N \geq n)}{d^n} (X_n - \mathbb{E}(X_{n+j}|\mathcal{A}_n)) \stackrel{11.11.1}{=} n \sum_{i=n}^{n+j-1} \frac{p_i}{i} + \underbrace{(n-d^j)}_{\geq n-1 \geq 0} \cdot \sum_{i=n+1+j}^{n_0} \frac{p_i}{i} \geq 0.$$

Thus it follows

$$\left(X_n - \mathbb{E}\left(X_{\tau_{n+j}^{(0)}} \middle| \mathcal{A}_n \right) \right) > 0 \text{ on } \{R_n = 1\}.$$

So we infer $C_n^{(1)} = \{R_n \in \{1, \infty\}\}$ for all $n \in \mathbb{N}$ with $n < n_0$.

Hence we have $C^{(1)} = C(\mathbb{N}_{<n_0})$. □

11.12 Setting Continued

Let $\kappa : \mathbb{N} \rightarrow \mathbb{N} \cup \{\infty\}$. Assume $C_n^{(0)} := \Omega$ and $\tau_n^{(0)} := n$ for all $n \in \mathbb{N} \cup \{\infty\}$. Define $B^1 := \mathbb{N}_{<n_0}$ and for all $k \in \mathbb{N}$ inductively with $*$ defined in Setting 11.3 (page 97)

$$B^{k+1} = (B^k)^{* \kappa(k)}.$$

Define further

$$B^\infty := \bigcap_{k \in \mathbb{N}} B^k.$$

11.13 Theorem

For all $k \in \mathbb{N}$ we have

$$C^{(k)} = C(B^k),$$

where the left-hand side is defined in Algorithm 2.13 (page 24) and the mapping C on the right in Setting 11.3 (page 97).

$\tau(C(B^\infty))$ is optimal.

Proof: We start the algorithm with $C^{(0)} \equiv \Omega$, but there is (except for the trivial case of $n_0 = 1$) no $A \subseteq \mathbb{N}$ with $C(A) \equiv \Omega$ (see Definition 11.3 (page 97)). Hence the induction cannot start with $k = 0$. The induction start with $k = 1$ is done in Theorem 11.11 (page 102). Thereby we have $C^{(1)} = C(\mathbb{N}_{<n_0}) = C(B^1)$. For the induction step⁴ consider $k \in \mathbb{N}$ with $C^{(k)} = C(B^k)$, then we have

$$C^{(k+1)} \stackrel{\text{def}}{=} (C^{(k)})^{*\kappa(k)} \stackrel{\text{l.v.}}{=} (C(B^k))^{*\kappa(k)} \stackrel{11.10}{=} C((B^k)^{*\kappa(k)}) \stackrel{\text{def}}{=} C(B^{k+1}).$$

The second statement follows by the first with Remark 5.14 (page 61). \square

11.14 Lemma⁵

For all $l \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$, $\kappa_1, \kappa_2 \in \mathbb{N} \cup \{\infty\}$ with $\kappa_2 \leq \kappa_1$ we have

$$\mathbb{N} \cap [l, n_0[\subseteq B^{*\kappa_1} \implies \mathbb{N} \cap [l, n_0[\subseteq (B^{*\kappa_1})^{*\kappa_2}.$$

Proof: By Definition 11.3 (page 97) we have obviously for all $n \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$, $j \in \mathbb{N}$

$$a(n, d, B, j) = a(n, d, B \cap [n + j, n_0[, j)$$

and hence for all $\kappa \in \mathbb{N} \cup \{\infty\}$

$$a_\kappa(n, d, B) = a_\kappa(n, d, B \cap [n, n_0[).$$

Let $l \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$, $\kappa_1, \kappa_2 \in \mathbb{N} \cup \{\infty\}$ with $\mathbb{N} \cap [l, n_0[\subseteq B^{*\kappa_1}$. Let $n \in \mathbb{N} \cap [l, n_0[$.

Then we have $B \cap [n, n_0[= \mathbb{N} \cap [n, n_0[= B^{*\kappa_1} \cap [n, n_0[$.

Since $n \in B^{*\kappa_1}$ we have

$$0 \leq a_{\kappa_1}(n, d, B) \leq a_{\kappa_2}(n, d, B) = a_{\kappa_2}(n, d, B^{*\kappa_1}),$$

hence we have $n \in (B^{*\kappa_1})^{*\kappa_2}$. \square

11.15 Lemma: Optimality⁶

Consider $B \subseteq \mathbb{N}_{<n_0}$, $k \in \mathbb{N}_{\geq 2}$ and some $l \in \mathbb{N}$ with $B^k = \mathbb{N} \cap [l, n_0[$. Then we have

$$B^k = B^{k+1}.$$

Proof: By Lemma 11.14 (page 104) we have

$$B^k \stackrel{11.12}{=} (B^{k-1})^{*\kappa(k)} \subseteq ((B^{k-1})^{*\kappa(k)})^{*\kappa(k)}.$$

The other inclusion is always true by definition. Using Lemma 11.14 (page 104), Theorem 11.13 (page 104), Conclusion 5.12 (page 60) and Theorem 5.11 (page 57) we infer $\tau(C(B^k))$ is optimal. Thus it follows by Lemma 5.9 (page 52) that $C^{(k)} = C^{(k+1)}$ and hence $B^k = B^{k+1}$. \square

⁴The general induction step from $k=1$ to $k=2$ can be found in [Irl80, 3.2] with $n \cdot c_n(d) = a(n, d, 1, \mathbb{N}_{<n_0})$.

Further induction steps are done in [Irl80] for several examples individually.

⁵Generalisation of the ideas of [Irl80, Lemma 3.4], therein $B = \mathbb{N}_{<n_0}$.

⁶For $B = \mathbb{N}_{<n_0}$ this can be found in [Irl80, Theorem 3.4].

11.16 Lemma: Problem-Independent Positive Values of a^7

Let $n \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$.

(11.16.1) For all $j \in \mathbb{N}$ with $n_0 - 2n \leq j \vee n(1-d) \geq d^j$ we have $a(n, d, B, j) \geq 0$.

Consider $2 \leq n_0 < \infty$.

Then we have

(11.16.2)
$$a(n, d, B, n_0 - n - 1) \geq 0$$

and for all $\kappa \in \mathbb{N} \cup \{\infty\}$

(11.16.3)
$$a_\kappa(n_0 - 1, d, B) \geq 0.$$

Proof: To prove (11.16.1) let $j \in \mathbb{N}$.

- If $n(1-d) \geq d^j$, then we have $d \neq 1$ and

$$\begin{aligned} \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} &\leq \sum_{i=n+j}^l d^i \cdot 1 \\ &\leq d^{n+j} \sum_{i=0}^{\infty} d^i = d^n d^j \frac{1}{1-d} \leq n d^n. \end{aligned}$$

- If $n_0 - n \leq j$, then

$$a(n, d, B, j) = \sum_{l=n}^{n_0} \frac{p_l}{l} (n d^n - 0) = n d^n \sum_{l=n}^{n_0} \frac{p_l}{l} \geq 0.$$

- If $n_0 - 2n \leq j < n_0 - n$, then we have

$$n \geq n_0 - n - j = \sum_{i=n+j}^{n_0} 1 \geq \sum_{i=n+j}^{n_0} d^{i-n} = \frac{1}{d^n} \sum_{i=n+j}^{n_0} d^i,$$

hence for all $l \in \mathbb{N}$ with $n \leq l \leq n_0$

$$\sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+j, i[} \frac{k-1}{k} \leq \sum_{i=n+j}^{n_0} d^i \cdot 1 \leq n d^n.$$

Now consider $2 \leq n_0 < \infty$.

Then we have $-n \leq -1$, hence $n_0 - 2n \leq n_0 - n - 1$.

And thus follows (11.16.2) by (11.16.1).

For all $j \in \mathbb{N}$ we have $n_0 - 2(n_0 - 1) = 2 - n_0 \leq 0 \leq j$,

hence by (11.16.1) $a(n_0 - 1, d, B, j) \geq 0$.

Consider $\kappa \in \mathbb{N} \cup \{\infty\}$. By the definition of a_κ we have (11.16.3). □

11.17 Conclusion about a and a_κ

Let $B \subseteq \mathbb{N}_{n_0}$, $n \in B$, $\kappa \in \mathbb{N} \cup \{\infty\}$. We have

$$a_\kappa(n, d, B) \geq 0$$

\iff

$$a(n, d, B, j) \geq 0 \text{ for all } j \in \mathbb{N} \text{ with } j < \kappa \wedge (n_0 - 2n) \text{ and } n(1-d) < d^j.$$

Proof: This is a consequence of Lemma 11.16 (page 105). □

⁷This is a generalisation of the text behind the proof of [Irl80, 3.6], therein only for $j = 1$, and the ideas of [Irl80, Lemma 3.3], therein only for $B = \mathbb{N}_{<n_0}$.

11.18 Remark

Let $n \in \mathbb{N}$, $j \in \mathbb{N}$ with $n(1-d) \geq d^j$. Then we have

$$\forall n' \in \mathbb{N}, j \in \mathbb{N}_0 \quad n \leq n', j \leq j' \implies n'(1-d) \geq d^{j'}.$$

11.19 Lemma⁸

Let $d < 1$, $B \subseteq \mathbb{N}_{<n_0}$, $\kappa \in \mathbb{N} \cup \{\infty\}$. We have

$$B \cap \left[\frac{d}{1-d}, n_0 \right] \subseteq B^{*\kappa}.$$

Proof: Let $n \in B$ with $\frac{d}{1-d} \leq n < n_0$. Then for all $j \in \mathbb{N}$ we have $(1-d)n \geq d^1 \geq d^j$, hence $a(n, d, B, j) \geq 0$. Thus we have $a_\kappa(n, d, B) \geq 0$, hence $n \in B^{*\kappa}$. \square

11.20 Conclusion: Reasonable Minimal Value for d

If $d \leq \frac{1}{2}$ we have $B = B^{*\kappa}$ for all $B \subseteq \mathbb{N}_{<n_0}$, $\kappa \in \mathbb{N} \cup \{\infty\}$.

Proof: Let $d \leq \frac{1}{2}$. Then $2d \leq 1$, hence $d \leq 1-d$, $\frac{d}{1-d} \leq 1$ and thus $\mathbb{N}_{<n_0} \cap \left[\frac{d}{1-d}, n_0 \right] = \mathbb{N}_{<n_0}$. The rest follows by Lemma 11.19 (page 106). \square

11.21 Lemma: Bounded Values

Let $n \in \mathbb{N}_{<n_0}$, $B \subseteq \mathbb{N}_{<n_0}$, $j \in \mathbb{N}$. Then we have

$$|a(n, d, B, j)| \leq \sum_{l=n}^{n_0} p_l \leq 1.$$

Proof: For all $i \in [n+j, n_0]$ we have $d^i \leq d^n$ since $d \leq 1$.

Hence we have

$$(11.21.1) \quad \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+j, i]} \frac{k-1}{k} \leq \sum_{i=n+j}^l d^i \leq \sum_{i=n+j}^l d^n \leq l \cdot d^n.$$

Define for all $l \in [n, n_0] \cap \mathbb{N}$

$$c_l := \frac{1}{l} \left(nd^n - \sum_{i \in [n+j, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+j, i]} \frac{k-1}{k} \right).$$

Obviously we have

$$-1 \leq -d^n = \frac{1}{l} (-l \cdot d^n \cdot 1) \stackrel{(11.21.1)}{\leq} c_l \leq \frac{1}{l} nd^n \leq d^n \leq 1,$$

hence $|c_l| \leq 1$. So it follows that

$$|a(n, d, B, j)| = \left| \sum_{l=n}^{n_0} p_l c_l \right| \leq \sum_{l=n}^{n_0} p_l |c_l| \leq \sum_{l=n}^{n_0} p_l \leq \sum_{l=1}^{n_0} p_l = 1. \quad \square$$

⁸Generalisation of [Irl80, text behind the proof of 3.6], therein $j = 1$

11.22 Conclusion: Bounded Values

The image of the function a is a subset of $[-1, 1]$ and hence for all $\kappa \in \mathbb{N} \cup \{\infty\}$ the image of a_κ is a subset of $[-1, 1] \cup \{\infty\}$.

Proof: This is a consequence of Lemma 11.21 (page 106). □

11.23 Lemma: Limit Value Zero⁹

Consider $n_0 = \infty$. Let $B \subseteq \mathbb{N}_{<n_0}$. We have

$$\lim_{n \rightarrow \infty} a(n, d, B, j) = 0 \text{ for all } j \in \mathbb{N}$$

and hence for all $\kappa : \mathbb{N} \rightarrow \mathbb{N} \cup \{\infty\}$

$$\lim_{n \rightarrow \infty} a_{\kappa(n)}(n, d, B) = 0.$$

Proof: Since $\lim_{n \rightarrow \infty} \sum_{l=n}^{\infty} p_l = 0$, this follows by Lemma 11.21 (page 106). □

11.24 Theorem¹⁰

Let $B \subseteq \mathbb{N}_{<n_0}$, $\kappa \in \mathbb{N} \cup \{\infty\}$.

Assume for all $n \in \mathbb{N}_{<n_0-2}$ with $n(1-d) < d$

$$\frac{a_\kappa(n, d, B)}{n} \geq \frac{a_\kappa(n+1, d, B)}{n+1} \implies \frac{a_\kappa(n+1, d, B)}{n+1} \geq \frac{a_\kappa(n+2, d, B)}{n+2}.$$

Then we have

$$(11.24.1) \quad B^{*\kappa} = \emptyset$$

or

$$(11.24.2) \quad \text{there is } l \in B \text{ with } B^{*\kappa} = B \cap [l, n_0[.$$

Proof: Since this is obvious for $n_0 \leq 2$, consider $n_0 > 2$.

We will prove now

$$(11.24.3) \quad \forall p \in B \setminus B^{*\kappa} \quad \forall k \in B^{*\kappa} \quad k \notin]p, n_0[.$$

Suppose $p \in B \setminus B^{*\kappa}$. Since $p \notin B^{*\kappa}$, we have $a_\kappa(p, d, B) < 0$.

Assume the existence of some $k \in B^{*\kappa}$ with $k < p$, hence $k \in B$ and $a_\kappa(k, d, B) \geq 0$.

Then there is $m \in \mathbb{N}$ with $k \leq m \leq p$ and $a_\kappa(m, d, B) \geq 0 > a_\kappa(m+1, d, B)$.

If $2 < n_0 < \infty$, it follows by induction over n that $a_\kappa(n_0 - 1, d, B) < 0$, contradicting (11.16.3).

If $n_0 = \infty$, it follows that $\lim_{n \rightarrow \infty} a_\kappa(n, d, B) < 0$, contradicting Lemma 11.23 (page 107).

Now consider the situation of $B^{*\kappa} \neq \emptyset$ and let l be the minimal element of $B^{*\kappa}$. Assume the existence of some $p \in B \setminus B^{*\kappa}$ with $p > l$, then we have by (11.24.3) that $B^{*\kappa} \subseteq]p, n_0[$ which is a contradiction to the minimality of l . Hence we have $B^{*\kappa} = B \cap [l, n_0[$. □

⁹Generalisation of [Irl80, Remark behind Lemma 3.3], therein for $B = \mathbb{N}_{<n_0}$

¹⁰Generalisation of [Irl80, Proposition 3.6.a], therein for $B = \mathbb{N}_{<n_0}$.

11.25 Theorem¹¹

For all $k \in \mathbb{N}$ we have $B^k \neq \emptyset$.

Proof: We will show this by induction.

By definition we have $B^1 = \mathbb{N}_{<n_0} \neq \emptyset$.

Let us first consider the case $n_0 < \infty$. By equation (11.16.3) we have $a_{\kappa(k)}(n_0 - 1, d, B^k) \geq 0$ for all $k \in \mathbb{N}$, hence it follows (due to $n_0 - 1 \in \mathbb{N}_{<n_0} = B^1$) inductively that $n_0 - 1 \in B^k$ for all $k \in \mathbb{N}$.

Now let us consider the case $n_0 = \infty$. Let $k \in \mathbb{N}$ with $B^k \neq \emptyset$. Assume for all $n \in B^k$ that $a_{\kappa(k)}(n, d, B^k) < 0$. Then we have $\emptyset = B^{k+1}$, hence $B^{k+2} = \emptyset$ and thus $B^{k+1} = B^{k+2}$. The assumptions made in Conclusion 5.12 (page 60) and Theorem 5.11 (page 57) are obviously fulfilled.¹² By the mentioned theorem and conclusion and Lemma 11.10 (page 101) $\inf \{p \in \mathbb{N}; R_p = \infty\} = \tau_0(C(\emptyset))$ is optimal with $E(X_{\tau_0(C(\emptyset))}) = 0$. We have $E(X_1) > 0$, a contradiction. Hence $B^{k+1} \neq \emptyset$. \square

11.26 Lemma

Let $n \in \mathbb{N}_{<n_0}$ and $B \subseteq \mathbb{N}_{<n_0}$ with $n+1 \in B \cup \{n_0\}$. Then we have

$$\frac{a(n, d, B, 1)}{n} - \frac{a(n+1, d, B, 1)}{n+1} = \frac{d^n}{n} \left(p_n + (n(1-d) - d) \sum_{l=n+1}^{n_0} \frac{p_l}{l} \right),$$

hence

$$\frac{a(n, d, B, 1)}{n} \geq \frac{a(n+1, d, B, 1)}{n+1} \quad \text{iff} \quad p_n \geq (d - n(1-d)) \sum_{l=n+1}^{n_0} \frac{p_l}{l}.$$

Proof: We have

$$\begin{aligned} \frac{a(n, d, B, 1)}{n} &= \frac{1}{n} \sum_{l=n}^{n_0} \frac{p_l}{l} \left(nd^n - \sum_{i \in [n+1, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+1, i[} \frac{k-1}{k} \right) \\ &= \frac{d^n}{n} \sum_{l=n}^{n_0} \frac{p_l}{l} \left(n - \sum_{i \in [n+1, l] \cap (B \cup \{n_0\})} d^{i-n} \cdot \prod_{k \in B \cap [n+1, i[} \frac{k-1}{k} \right) \\ &= \frac{d^n}{n} \left(p_n + \sum_{l=n+1}^{n_0} \frac{p_l}{l} \left(n - d - \underbrace{\sum_{i \in [n+2, l] \cap (B \cup \{n_0\})} d^{i-n} \cdot \frac{n}{n+1} \prod_{k \in B \cap [n+2, i[} \frac{k-1}{k}}_{C:=} \right) \right) \end{aligned}$$

and

$$\begin{aligned} \frac{a(n+1, d, B, 0)}{n+1} &= \frac{1}{n+1} \sum_{l=n+1}^{n_0} \frac{p_l}{l} \left((n+1)d^{n+1} - \sum_{i \in [n+2, l] \cap (B \cup \{n_0\})} d^i \cdot \prod_{k \in B \cap [n+2, i[} \frac{k-1}{k} \right) \\ &= \frac{d^n}{n} \left(\sum_{l=n+1}^{n_0} \frac{p_l}{l} \left(nd - \frac{n}{n+1} \underbrace{\sum_{i \in [n+2, l] \cap (B \cup \{n_0\})} d^{i-n} \cdot \prod_{k \in B \cap [n+2, i[} \frac{k-1}{k}}_{C:=} \right) \right). \end{aligned}$$

¹¹Generalisation of [Irl80, Lemma 3.3], therein $k = 2$ and $\kappa = 1$, hence it was shown that $(\mathbb{N}_{<n_0})^* \neq \emptyset$.

¹²See Setting 11.3 (page 97) and [Irl80, part 3].

Hence it follows

$$\begin{aligned} \frac{a(n, d, B, 0)}{n} - \frac{a(n+1, d, B, 0)}{n+1} &= \frac{d^n}{n} \left(p_n + \sum_{l=n+1}^{n_0} \frac{pl}{l} (n-d-C - (nd-C)) \right) \\ &= \frac{d^n}{n} \left(p_n + (n(1-d) - d) \sum_{l=n+1}^{n_0} \frac{pl}{l} \right). \end{aligned} \quad \square$$

11.27 Theorem: Deterministic Population Size¹³

If $n_0 < \infty$ and $p_{n_0} = 1$, then

$$\text{there is } l \in \mathbb{N}_{<n_0} \text{ with } B^2 = B^\infty = \mathbb{N} \cap [l, n_0[.$$

Proof: By Lemma 11.15 (page 104) it is enough to show that there is $l \in \mathbb{N}$ with

$$(\mathbb{N}_{<n_0})^{*\kappa(1)} = \mathbb{N} \cap [l, n_0[.$$

Therefore it suffices to show that we have

$$a_{\kappa(1)}(n, d, \mathbb{N}_{<n_0}) \geq 0 \implies a_{\kappa(1)}(n+1, d, \mathbb{N}_{<n_0}) \geq 0 \quad \text{for all } n \in \mathbb{N} \text{ with } n+1 < n_0.$$

For all $n \in \mathbb{N}_{<n_0}$ and $j \in \mathbb{N}$ we have

$$\begin{aligned} a(n, d, \mathbb{N}_{<n_0}, j) &\stackrel{(11.3.1)}{=} \sum_{l=n}^{n_0} \frac{pl}{l} \left(nd^n - (n+j-1) \sum_{i=n+j}^l \frac{d^i}{i-1} \right) \\ &= \frac{pn_0}{n_0} (n+j-1) \left(\frac{n}{n+j-1} d^n - \sum_{i=n+j}^{n_0} \frac{d^i}{i-1} \right) \end{aligned}$$

and hence

$$(11.27.1) \quad a(n, d, \mathbb{N}_{<n_0}, j) \geq 0 \iff \sum_{i=n+j}^{n_0} \frac{d^i}{i-1} \leq \frac{n}{n+j-1} d^n.$$

Let $n \in \mathbb{N}$ with $n+1 < n_0$ and assume $a_{\kappa(1)}(n, d, \mathbb{N}_{<n_0}) \geq 0$.

Let $j \in \mathbb{N}$ with $j < n_0 - (n+1)$ and $n(1-d) < d^j$. Then we have

$$(11.27.2) \quad n - d^j < n - n(1-d) = nd$$

and by Conclusion 11.17 (page 105)

$$(11.27.3) \quad a(n, d, \mathbb{N}_{<n_0}, j) \geq 0.$$

Then we have

$$\begin{aligned} \sum_{i=(n+1)+j}^{n_0} \frac{d^i}{i-1} &= \sum_{i=n+j}^{n_0} \frac{d^i}{i-1} - \frac{d^{n+j}}{n+j-1} \\ &\stackrel{(11.27.3)}{\leq} \stackrel{(11.27.1)}{=} \frac{n}{n+j-1} d^n - \frac{d^{n+j}}{n+j-1} \\ &= \frac{n-d^j}{n+j-1} d^n \\ &\stackrel{(11.27.2)}{<} \frac{nd}{n+j-1} d^n < \frac{(n+1)}{(n+1)+j-1} d^{n+1}. \end{aligned}$$

Hence by (11.27.1) we have $a(n+1, d, \mathbb{N}_{<n_0}, j) \geq 0$.

By Conclusion 11.17 (page 105) we have

$$a_{\kappa(1)}(n+1, d, \mathbb{N}_{<n_0}) \geq 0. \quad \square$$

¹³See also [Irl80, 3.5].

11.28 Conclusion¹⁴

If

$$(11.28.1) \quad \frac{p_{n+1} \cdot p_{n+k}}{n+k} \geq \frac{p_n \cdot p_{n+k+1}}{n+k+1} \text{ for all } k \in \mathbb{N}_{<n_0-n} \text{ and } n \in \mathbb{N}_{<n_0-2}$$

or

$$(11.28.2) \quad p_n \leq p_{n+1} \text{ for all } n \in \mathbb{N}_{<n_0-2},$$

then $B^2 = B^\infty$ and hence $C^{(2)} = C^{(\infty)}$.

Proof: If we have (11.28.2), then Lemma 11.26 (page 108) easily implies the assertion of Theorem 11.24 (page 107) for $B = B^1 = \mathbb{N}_{<n_0}$. Next we will show that (11.28.1) also implies that assertion.

Hence assume (11.28.1) and let $n \in \mathbb{N}_{<n_0-2}$ with $n(1-d) < d$.

If $p_n = 0$ and $\frac{a(n,d,B,0)}{n} \geq \frac{a(n+1,d,B,0)}{n+1}$, then

$$p_{n+1} \geq 0 = p_n \geq (d - n(1-d)) \sum_{l=n+1}^{n_0} \frac{p_l}{l} \geq (d - n(1-d)) \sum_{l=n+2}^{n_0} \frac{p_l}{l}$$

hence we have $\frac{a(n+1,d,B,0)}{n} \geq \frac{a(n+2,d,B,0)}{n+1}$ by Lemma 11.26 (page 108).

By (11.28.1) with $k = n_0 - n - 1$ and $p_{n+k+1} = p_{n_0} > 0$ we have

$$p_n > 0 \text{ implies } p_{n+1} > 0.$$

Assume

$$\frac{a(n,d,B,0)}{n} \geq \frac{a(n+1,d,B,0)}{n+1}.$$

Then we have by Lemma 11.26 (page 108)

$$p_n \geq (d - n(1-d)) \sum_{l=n+1}^{n_0} \frac{p_l}{l}.$$

We have

$$\sum_{l=n+1}^{n_0} \frac{p_l}{l \cdot p_n} \geq \sum_{k=1}^{n_0-n} \frac{p_{n+k}}{(n+k)p_n} \stackrel{(11.28.1)}{\geq} \sum_{k=1}^{n_0-n-1} \frac{p_{n+k+1}}{(n+k+1)p_{n+1}} \geq \sum_{l=n+1+1}^{n_0} \frac{p_l}{l \cdot p_{n+1}}$$

and $d - n(1-d) > d - (n+1)(1-d)$. Hence we have

$$1 \geq (d - n(1-d)) \sum_{l=n+1}^{n_0} \frac{p_l}{l} \geq (d - (n+1)(1-d)) \sum_{l=n+1+1}^{n_0} \frac{p_l}{l \cdot p_{n+1}},$$

hence

$$p_{n+1} \geq (d - (n+1)(1-d)) \sum_{l=n+1+1}^{n_0} \frac{p_l}{l}$$

and thus

$$\frac{a(n+1,d,B,0)}{n+1} \geq \frac{a(n+2,d,B,0)}{n+2}.$$

Hence the assertion of Theorem 11.24 (page 107) is fulfilled for $B = B^1 = \mathbb{N}_{<n_0}$.

Since Theorem 11.25 (page 108) says that $B^2 \neq \emptyset$, we do not have (11.24.1).

Hence we have (11.24.2).

We infer by Lemma 11.15 (page 104) that $B^2 = B^3$, hence $B^2 = B^\infty$. □

¹⁴See [Irl80, 3.6.b.1]

11.29 Remark¹⁵

(11.28.1) is fulfilled for geometric, Poisson, binomial, hypergeometric, and Laplace distributions.

11.30 Conclusion¹⁶

Consider $l, m, s \in \mathbb{N}$ with $l < m < s$ and $B^2 = \{n \in \mathbb{N}; l < n < m \text{ or } s < n < n_0\}$.

If

$$a_{\kappa(2)}(n, d, B^2) \geq 0 \text{ for all } n \in \mathbb{N} \text{ with } l < n < m,$$

we have $B^2 = B^\infty$ and hence $C^{(2)} = C^{(\infty)}$.

If

$$a_{\kappa(2)}(n, d, B^2) < 0 \text{ for all } n \in \mathbb{N} \text{ with } l < n < m,$$

we have $B^3 = B^\infty = \{n \in \mathbb{N}; s < n < n_0\}$ and hence $C^{(3)} = C^{(\infty)}$.

¹⁵See [Irl80, text between 3.6 and 3.7]

¹⁶See [Irl80, 3.8]

Bibliography

- [Asm03] Søren Asmussen. *Applied Probability and Queues*. Applications of mathematics 51. Springer, 2003.
- [BKS06] Christian Bender, Anastasia Kolodko, and John Schoenmakers. Policy iteration for American options: overview. *Monte Carlo Methods Appl.*, 12(5-6):347–362, 2006.
- [BKS08] Christian Bender, Anastasia Kolodko, and John Schoenmakers. Enhanced policy iteration for American options via scenario selection. *Quantitative Finance*, 8(2):135–146, 2008.
- [BS06] Christian Bender and John Schoenmakers. An iterative method for multiple stopping: convergence and stability. *Advances in Applied Probability*, 38(3):729–749, 2006.
- [CRS71] Yuan Shih Chow, Herbert Robbins, and David Siegmund. *Great expectations: the theory of optimal stopping*. Houghton Mifflin Company Boston, 1971.
- [Gla03] Paul Glasserman. *Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability)*. Springer, 2003.
- [How60] Ronald A. Howard. *Dynamic Programming and Markov Processes*. MIT Press, Cambridge, MA, 1960.
- [Irl79] Albrecht Irl. Monotone stopping problems and continuous time process. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 48:49–56, 1979.
- [Irl80] Albrecht Irl. On the best choice problem with random population size. *Zeitschrift für Operations Research. Serie A. Serie B*, 24(5):177–190, 1980.
- [Irl02] Albrecht Irl. *Finanzmathematik*. Teubner, Wiesbaden, 2002.
- [Irl05] Albrecht Irl. *Wahrscheinlichkeitstheorie und Statistik: Grundlagen - Resultate - Anwendungen*. Teubner Verlag, 2005.
- [Irl06] Albrecht Irl. A forward algorithm for solving optimal stopping problems. *Journal of Applied Probability*, 43(1):102–113, 2006.
- [Irl09] Albrecht Irl. On forward improvement iteration for stopping problems. In *Proceeding of the International Workshop of Sequential Methodologies*, Troyes, 2009.
- [KS06] Anastasia Kolodko and John Schoenmakers. Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics*, 10(1):27–49, 2006.
- [Nev75] Jacques Neveu. *Discrete-parameter martingales*. Elsevier, 1975.
- [Pre10] E. L. Presman. A new approach to the solution of optimal stopping problem. In *Conference Proceedings*, Turku, 2010.
- [PS73] E. L. Presman and I. M. Sonin. The best choice problem for a random number of objects. *Theory of Probability and its Applications*, 17:657–668, 9 1973.
- [Put94] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience, 1994.
- [Sne52] James Laurie Snell. Applications of martingale system theorems. *Transactions of the American Mathematical Society*, 73:293–312, 1952.

Eidesstattliche Erklärung

Ich habe die vorliegende Arbeit, abgesehen von der Beratung durch den Betreuer meiner Promotion, unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel verwendet; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Diese Arbeit ist bisher weder veröffentlicht noch zur Veröffentlichung eingereicht worden und hat weder ganz noch zum Teil an anderer Stelle im Rahmen eines Prüfungsverfahrens vorgelegen. Des Weiteren habe ich noch keinen Promotionsversuch unternommen.

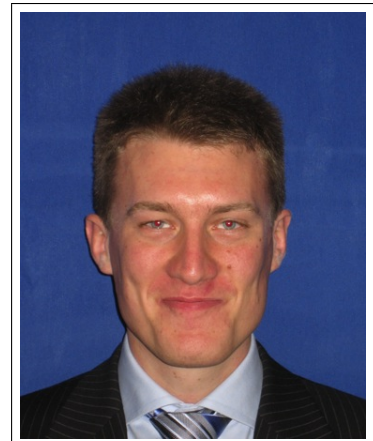
Kiel, Dezember 2010

(Julian Peter Lemburg)

Lebenslauf

Persönliche Daten

Name **Julian Peter Lemburg**
Geburtsdatum **18. Februar 1983**
Geburtsort **Eckernförde, Deutschland**
Nationalität **deutsch**



Ausbildung

10/2002 - **Studium der Mathematik mit Nebenfach Volkswirtschafts-**
05/2008 **lehre an der CAU Kiel, Deutschland, Diplom 05/2008, Note sehr**
gut, Vordiplom 04/2005, Note sehr gut.

08/1989 - **Schulbesuch in Dänischenhagen, Altenholz-Stift und Kiel-**
06/2002 **Südfriedhof, Deutschland, Abitur 06/2002, Note 1.3, Fachhoch-**
schulreife 07/2001, Note 1.2, Leistungskurse Mathematik und Daten-
verarbeitungstechnik.

Berufserfahrung

05/2008 - **Wissenschaftlicher Angestellter am Mathematischen Seminar**
05/2011 **der CAU Kiel, Deutschland, Verteidigung 01/2011, noch keine No-**
te.

Kiel, Dezember 2010

Curriculum Vitae

Personal Data

Name **Julian Peter Lemburg**
Date of birth **February 18, 1983**
Place of birth **Eckernförde, Germany**
Citizenship **German**



Education

10/2002 - **Studies of mathematics with political economics at the University of Kiel, Germany, Diploma in 05/2008, graded sehr gut, Vordiplom, graded sehr gut in 04/2005.**
05/2008
08/1989 - **Schools in Dänischenhagen, Altenholz-Stift and Kiel-Südfriedhof, Germany, Abitur in 06/2002, graded 1.3, Fachhochschulreife in 07/2001, graded 1.2, major fields of study: mathematics and information technology.**
06/2002

Professional Experience

05/2008 - **Research associate of the department of mathematics at the University of Kiel, Germany, Defence 01/2011, not yet graded.**
05/2011

Kiel, December 2010