Probability Theory and Stochastic Modelling 75

Pierre Carpentier
Jean-Philippe Chancelier
Guy Cohen
Michel De Lara

# Stochastic Multi-Stage Optimization

## At the Crossroads between Discrete Time Stochastic Control and Stochastic Programming

Springer

# Probability Theory and Stochastic Modelling

Volume 75

**The Probability Theory and Stochastic Modelling** series is a merger and continuation of Springer's two well established series Stochastic Modelling and Applied Probability and Probability and Its Applications series. It publishes research monographs that make a significant contribution to probability theory or an applications domain in which advanced probability methods are fundamental. Books in this series are expected to follow rigorous mathematical standards, while also displaying the expository quality necessary to make them useful and accessible to advanced students as well as researchers. The series covers all aspects of modern probability theory including

- Gaussian processes
- Markov processes
- Random Fields, point processes and random sets
- Random matrices
- Statistical mechanics and random media
- Stochastic analysis

as well as applications that include (but are not restricted to):

- Branching processes and other models of population growth
- Communications and processing networks
- Computational methods in probability and stochastic processes, including simulation
- Genetics and other stochastic models in biology and the life sciences
- Information theory, signal processing, and image synthesis
- Mathematical economics and finance
- Statistical methods (e.g. empirical processes, MCMC)
- Statistics for stochastic processes
- Stochastic control
- Stochastic models in operations research and stochastic optimization
- Stochastic models in the physical sciences

More information about this series at http://www.springer.com/series/13205

Pierre Carpentier · Jean-Philippe Chancelier
Guy Cohen · Michel De Lara

# Stochastic Multi-Stage Optimization

## At the Crossroads between Discrete Time Stochastic Control and Stochastic Programming

Springer

Pierre Carpentier
ENSTA ParisTech
Palaiseau
France

Jean-Philippe Chancelier
CERMICS
École Nationale des Ponts
et Chaussées ParisTech
Marne la Vallée
France

Guy Cohen
CERMICS
École Nationale des Ponts
et Chaussées ParisTech
Marne la Vallée
France

Michel De Lara
CERMICS
École Nationale des Ponts
et Chaussées ParisTech
Marne la Vallée
France

ISSN 2199-3130 ISSN 2199-3149 (electronic)
Probability Theory and Stochastic Modelling
ISBN 978-3-319-18137-0 ISBN 978-3-319-18138-7 (eBook)
DOI 10.1007/978-3-319-18138-7

Library of Congress Control Number: 2015937212

Mathematics Subject Classification (2010): 93C15, 93C39, 49–XX, 60–XX

Springer Cham Heidelberg New York Dordrecht London


Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

# Preface

This book can be considered as the result of a ten-year cooperation (starting in 2000) of the four authors within the so-called Stochastic Optimization Working Group (SOWG), a research team of the CERMICS (Applied Mathematics Laboratory) of École Nationale des Ponts et Chaussées (ENPC-ParisTech). Among the topics addressed in this working group, a major concern was to devise numerical methods to effectively solve stochastic optimization problems, particularly in a dynamic context, as this was the context of most real-life applications also tackled by the group.

The background of the four authors is system theory and control but the 2000s have seen the emergence of the Stochastic Programming stream, a stochastic expansion of Mathematical Programming, so the group was interested in bridging the gap between these two communities.

Of course, several Ph.D. students took part in the activities of this group, and among them were Kengy Barty, Laetitia Andrieu, Babacar Seck, Cyrille Strugarek, Anes Dallagi, Pierre Girardeau. Their contributions are gratefully acknowledged. We hope this book can help future students to get familiar with the field.

The book comprises five parts and two appendices. The first part provides an introduction to the main issues discussed later in the book, plus a chapter on the stochastic gradient algorithm which addresses the so-called open-loop optimization problems in which on-line information is absent. Part Two introduces the theoretical tools and notions needed to mathematically formalize and handle the topic of information which plays a major part in stochastic dynamic problems. It also discusses optimality conditions for such problems, such as the dynamic programming equation, and a variational approach which will lead to numerical methods in the next part. Part Three is precisely about discretization and numerical approaches. A simple benchmark illustrates the contribution of the particle method proposed in Chap. 7. Convergence issues of all those techniques are discussed in Part Four. Part Five is devoted to more advanced topics that are more or less out of reach of the numerical methods previously discussed, namely multi-agent problems and the presence of the so-called dual effect. Appendix A recalls some basic facts on

Optimization, while Appendix B provides a brief description of essential tools of Probability theory.

Although the four authors share the responsibility of the whole book contents, the reader may be interested in knowing who was the primary writer of each chapter. Here is the list:

**Pierre Carpentier**: Chapter 2, Appendix A;
**Jean-Philippe Chancelier**: Chapter 8, Appendix B;
**Guy Cohen**: Notation (in preliminary pages), Chapters 1, 5, 6, 7;
**Michel De Lara**: Chapters 3, 4, 9, 10.

# Contents

# Notation

Here we explain some notation and typographical conventions that we have used throughout this book. We conclude with a short list of symbols, abbreviations and acronyms to which the reader may refer. In this discussion about notation, we raise a tricky point that stems from some divergence between conventional mathematical concepts on the one hand, and a long-standing practice and terminology used in Probability Theory on the other. Most of the time, this divergence causes no problem in understanding what is meant, but we point out a few circumstances when some confusion may arise.

## Some General Principles

This book is about stochastic optimization. As such, random variables are among the main mathematical notions involved. Unless specific reasons prevent us from doing so, we denote random variables by *capital bold* letters, e.g. $\boldsymbol{U}$. As taught in any elementary course in Probability Theory (see Appendix B in this book), random *variables* are indeed *functions* or *mappings* from a set generally called $\Omega$ to some other set, say $\mathbb{U}$.[1]

The space in which random variables, and more generally functions, take their values are denoted with the $\mathbb{BLACKBOARD}$ font. However, as is expected, symbols such as $\mathbb{R}$ and $\mathbb{N}$ have a special meaning, namely the set of real and integer numbers, respectively (they are included in the list below with additional variations such as $\overline{\mathbb{R}}$). Also, $\mathbb{P}$ denotes a probability measure and $\mathbb{E}$ denotes mathematical expectation (or conditional mathematical expectation). Functional spaces are generally denoted with the *calligraphic* font; for example, a mapping $\boldsymbol{U} : \Omega \to \mathbb{U}$ belongs to the set $\mathcal{U}$.

[1]Additional ingredients are also required ($\sigma$-fields over $\Omega$ and $\mathbb{U}$, a measurability requirement about the mapping, probability measure $\mathbb{P}$, etc.) but it is not our purpose to dwell on that here.

The *script* font is generally used to denote $\sigma$-fields (e.g. $\mathcal{F}$). We now refer the reader to the list of symbols and abbreviations at the end of this introduction.

## A Tricky Point

Here we make a few remarks about the effects of calling (random) "variables" objects which are indeed "functions", and the consequences of this abuse of language on notation. This abuse of language is customary in the world of Probability Theory but may cause substantial confusion for less aware readers. We discuss this issue by referring to several puzzling situations that arise in this book.

For example, consider the expression $\mathbb{E}(f(\boldsymbol{U}))$. With no hesitation, one understands that $f$ is a mapping from $\mathbb{U}$ to some other set (say $\mathbb{R}$ to fix ideas), that $f(\boldsymbol{U})$ must be interpreted as a new random variable, namely $f \circ \boldsymbol{U} : \Omega \to \mathbb{R}$, and that its expectation—that is, the integral of this function over $\Omega$ against the probability measure $\mathbb{P}$—is then evaluated. Hence, while in $f(\boldsymbol{U})$, $\boldsymbol{U}$ seems to play the part of a "variable", namely an argument of function $f$ according to its position within parentheses, it must indeed be remembered that this is a mapping to be composed with $f$ in order to produce a new mapping of the argument $\omega \in \Omega$ whose integral is then to be evaluated. Thus, there is no real difficulty.

The evaluation of the considered expression would change if $\boldsymbol{U}$ was replaced by another random variable $\boldsymbol{V}$. Because of the dependence of this expression upon this random variable, one would naturally consider the result as a function of the dummy argument $\boldsymbol{U}$. If $g$ denotes this function, we may write

$$g(\boldsymbol{U}) = \mathbb{E}(f(\boldsymbol{U})) = \int_{\Omega} f \circ \boldsymbol{U}(\omega)\, \mathbb{P}(\mathrm{d}\omega)\ ,$$

and we may even replace the first sign $=$ by $:=$ (which means that the left-hand side is defined by the expression on the right-hand side). Observe that the parts played by $\boldsymbol{U}$ in $g(\boldsymbol{U})$ and in $f(\boldsymbol{U})$ are quite different, despite the similarity in notation. Strictly speaking, $g(\boldsymbol{U})$ is a correct mathematical expression since $g$ is indeed a function of the random variable $\boldsymbol{U}$, whereas $f(\boldsymbol{U})$ is an ambiguous shortcut that experienced readers are able to interpret. However, a problem may arise when both expressions appear on both sides of an equality as in the first of the two equalities above. Notice that if the intermediate expression in these two equalities is cancelled and only the two extreme members of the equalities are kept, no question arises since, now, everywhere $\boldsymbol{U}$ is interpreted as a *function* (and $g$ is generally called a "*functional*" as a function of a function).

Therefore, the correct notation would be $g(\boldsymbol{U})$, whereas $f(\boldsymbol{U})$ is a shortcut that requires some appropriate interpretation, but in order to conform with a long-standing tradition in Probability Theory, we sometimes change $g(\boldsymbol{U})$ to $g([\boldsymbol{U}])$ in order to emphasize the fact that the "argument" $\boldsymbol{U}$ must rigorously be interpreted as the "global function" object and not only be used for the collection of its values $\boldsymbol{U}(\omega)$ as in $f(\boldsymbol{U})$.

Let us give other instances when such a distinction is necessary. In this book, stochastic optimization problems of the following generic form are considered:

$$\min_{\boldsymbol{U}} \mathbb{E}(j(\boldsymbol{U}, \boldsymbol{W})) \ ,$$

in which

- $\boldsymbol{U}$ is a random variable taking values in $\mathbb{U}$ and plays the part of the decision variable;
- $\boldsymbol{W}$ is another random variable taking values in $\mathbb{W}$ and plays the part of the "noise";
- $j$ is a real-valued mapping defined over $\mathbb{U} \times \mathbb{W}$ playing the part of the cost function.

A decision $\boldsymbol{U}$ is thus a random variable possibly subject to various constraints that are described in this book, and whose performance is evaluated by computing the expectation of the cost function which also involves an exogenous disturbance $\boldsymbol{W}$. The expression behind the min operator in the above formulation must be interpreted as we did for $f(\boldsymbol{U})$ in the previous discussion. Namely, a real-valued random variable $j(\boldsymbol{U}(\cdot), \boldsymbol{W}(\cdot))$ must be considered and its expectation must be evaluated. On the contrary, the minimization operation involves the random variable $\boldsymbol{U}$ "as a whole"; in particular, as we shall see later in this book, some constraints (so-called informational or measurability constraints) may prevent independent consideration of the individual values $\boldsymbol{U}(\omega)$ and force us to globally consider the whole function $\boldsymbol{U}$ in this minimization operation. Thus, according to our notational convention, we should instead write the previous stochastic optimization problem as

$$\min_{[\boldsymbol{U}]} \mathbb{E}(j(\boldsymbol{U}, \boldsymbol{W})) \ .$$

Nevertheless, for the sake of simplicity, we keep the former notation since the particular position of the decision in the min should prevent any ambiguity.

Finally, a third instance when this notation $[\boldsymbol{X}]$ proves useful is the following. The reader may refer to Appendix B to find definitions of conditional expectations $\mathbb{E}(\boldsymbol{X}|\boldsymbol{Y}))$ where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are two random variables with values in $\mathbb{X}$ and $\mathbb{Y}$, respectively. This conditional expectation is also a random variable with values in $\mathbb{X}$. Sometimes, we are also led to manipulate the function $\Psi : \mathbb{Y} \to \mathbb{X}$ which, whenever the event $\{\boldsymbol{Y} = y\}$ (that is the subset $\boldsymbol{Y}^{-1}(y) = \{\omega \,|\, \boldsymbol{Y}(\omega) = y\}$) has a positive probability for a given value $y \in \mathbb{Y}$, may be interpreted as the "expectation of $\boldsymbol{X}$ conditioned by the event $\{\boldsymbol{Y} = y\}$". It is explained in the appendix that $\Psi$ is a function of $y \in \mathbb{Y}$, that is, of the values taken by the random variable $\boldsymbol{Y}$, but that this function also depends on the "whole" function $\boldsymbol{Y}$ (and of course also on the function $\boldsymbol{X}$ as does the expectation $\mathbb{E}(\boldsymbol{X})$ itself). To emphasize this fact, we could write

$\Psi_{[\boldsymbol{Y}]}(y)$ instead of merely $\Psi(y)$. In the former expression, both the values $y$ taken by $\boldsymbol{Y}$ and the "global" random variable $\boldsymbol{Y}$ appear to play a part.

We will occasionally refer back to this discussion in the rest of this book.

## Symbols and Abbreviations

| | |
|---|---|
| $\mathbb{N}$ | Set of integer (natural) numbers |
| $\mathbb{R}$ | Set of real numbers |
| $\overline{\mathbb{R}}$ | $\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ |
| $\mathbb{E}$ | Mathematical expectation |
| $\mathbb{P}$ | Probability measure |
| Var | Variance (of a random variable) |
| $\mathrm{I}_{\mathbb{A}}$ | Identity function over set $\mathbb{A}$ |
| $\chi_A$ | Characteristic function of subset $A$ |
| $1_A$ | Indicator function of subset $A$ |
| $\lvert\cdot\rvert$ | Absolute value |
| $\langle\cdot\,,\cdot\rangle$ | Scalar product |
| $\lVert\cdot\rVert$ | Norm |
| $\nabla$ | Gradient |
| $\nabla_x$ | Partial gradient (with respect to $x$) |
| $\partial$ | Subdifferential |
| $\partial\cdot/\partial x$ | Partial derivative (with respect to $x$) |
| $\mathrm{proj}_A$ | Projection onto subset $A$ |
| $\boldsymbol{U} \preceq \boldsymbol{V}$ | Random variable $\boldsymbol{U}$ measurable with respect to $\boldsymbol{V}$ (same as $\boldsymbol{V} \succeq \boldsymbol{U}$; used also with functions, $\sigma$-fields, partitions, etc.) |
| $x^{\mathrm{T}}$ | Transposition of vector $x$ |
| dom | Domain (of a function) |
| $\mathrm{co}\,A$ | Convex hull of subset $A$ |
| $\overline{\mathrm{co}}\,A$ | Closed convex hull of subset $A$ |
| $\xrightarrow{\mathcal{D}}$ | Convergence in distribution |
| $\xrightarrow{\mathbb{P}}$ | Convergence in probability |
| $\xrightarrow{\text{a.s.}}$ | Almost sure convergence |
| l.s.c. | Lower semicontinuous |
| u.s.c. | Upper semicontinuous |
| i.i.d. | Independently identically distributed |
| iff | If and only if |
| w.r.t. | With respect to |
| s.t. | Subject to |
| a.s. | Almost surely (or almost sure) |
| $\mathbb{P}$-a.s. | Almost surely (or almost sure) w.r.t. to the probability measure $\mathbb{P}$ |

## Acronyms

| | |
|---|---|
| ADP | Approximate Dynamic Programming |
| APP | Auxiliary Problem Principle |
| DIS | Dynamic Information Structure |
| DP | Dynamic Programming |
| LBG | Linearly Bounded Gradient |
| LQG | Linear-Quadratic-Gaussian |
| MASIOS | Multi-Agent Stochastic Input–Output System |
| MQE | Mean Quadratic Error |
| NOLDE | No Open-Loop Dual Effect |
| SA | Stochastic Approximation |
| SAA | Sample Average Approximation |
| SDDP | Stochastic Dual Dynamic Programming |
| SIS | Static Information Structure |
| SOC | Stochastic Optimal Control |
| SP | Stochastic Programming |

# Part I
# Preliminaries

# Chapter 1
# Issues and Problems in Decision Making Under Uncertainty

## 1.1 Introduction

The future cannot be predicted exactly, but one may learn from past observations. Past decisions can also improve future predictability. This is the context in which decisions are generally made. Herein, we discuss some mathematical issues pertaining to this topic.

### *1.1.1 Decision Making as Constrained Optimization Problems*

Making decisions in a rational way is a problem which can be mathematically formulated as an *optimization* problem. Generally, several conflicting goals must be taken into account simultaneously. A choice must be made about which goals are formulated as constraints to be satisfied at a certain "level" (apart from constraints which are imposed by physical limitations), and which goals are reflected by (and aggregated within) a *cost function*.[1] Duality theory for *constrained optimization* problems provides a way to analyze, afterwards, the sensitivity of the best achievable cost as a function of constraint levels which were fixed a priori, and, possibly, to tune those levels to achieve a better trade-off between conflicting goals.

Problems that involve systems evolving in time enter the realm of *Optimal Control*. In a deterministic setting, Optimal Control has a long history dating back to the fifties with famous names such as Pontryagin [124] and Bellman [15]. The former, with his *Maximum Principle*, was more in the line of a *variational* approach of such problems, whereas the latter introduced the *Dynamic Programming* (DP) technique in connection with the state space approach.

[1]Throughout this book, without loss of generality, optimization problems are formulated as *minimization* problems, hence the objective function to be minimized is called a *cost*.



P. Carpentier et al., *Stochastic Multi-Stage Optimization*, Probability Theory and Stochastic Modelling 75, DOI 10.1007/978-3-319-18138-7_1

### 1.1.2 Facing Uncertainty

In general, when making decisions, one is faced with *uncertainties* which affect the cost function and, generally, the constraints. There are several possible attitudes associated with uncertainties, and consequently, several possible mathematical formulations of decision making problems under uncertainty. Let us mention two main possibilities.

**Worst Case Design**
The assumption here is that uncertainties lie in particular bounded subsets and, that one must consider the *worst situation* to be faced and try to make it as good as possible. In more mathematical terms, and considering the cost only for the time being (see hereafter for constraints), since one would like to minimize that cost, one must minimize the *maximal* possible value Nature can give to that cost by playing with uncertainties within the assumed bounded subsets. That is, a *min-max* (game like) problem is formulated and a *guaranteed* performance can be evaluated (as long as assumptions on uncertainties hold true).

The treatment of constraints in such an approach should normally follow the same lines of thought (one must fight against the worst possible uncertainty outcomes from the point of view of constraint satisfaction). Sometimes the terminology of *robust* decision making (or control) is used for approaches along those lines [16].

**Stochastic Programming or Stochastic Control**
Here, uncertainties are viewed as random variables following *a priori* probability laws. We shall call them "primitive" random variables as opposed to other "secondary" random variables involved in the problem and which are derived from the primitive ones by applying functions such as dynamic equations, feedback laws (see hereafter), etc. Then the cost to be minimized is the mathematical expectation of some performance index depending on those random variables and on decisions.

For this mathematical expectation to make sense, the decisions must also become random variables defined on the same underlying probability space. A trivial case is when those decisions are indeed *deterministic*: we shall call them *open-loop* decisions or "controls" later on. But they may also be true random variables because they are produced by applying functions to either primitive or secondary random variables. Here, we enter the domain of *feedback* or *closed-loop* control which plays a prominent part in decision making under uncertainty.

Let us now say a few words about constraint satisfaction. Constraints may be imposed as *almost sure* (a.s.) constraints. This is generally the case of equality or inequality constraints expressing physical laws or limitations. Other constraints may be formulated with mathematical expectations, although it is generally difficult to give a sound practical meaning to this approach. If a.s. requirements may sometimes be either unfeasible or not economically viable, one may appeal to "constraints in probability": the satisfaction of those constraints is required only "sufficiently often", that is, with a certain prescribed probability. We do not pursue this discussion here, as we mostly consider a.s. constraints in this book.

In the title of this section, we have used the words "Stochastic Programming" and "Stochastic Control". Stochastic Control, or rather Stochastic Optimal Control (SOC), is the extension of the theory of Deterministic Optimal Control to the situation when uncertainties are present and modeled by random variables, or stochastic processes since control theory mostly addresses dynamic problems. SOC problems were introduced not long after their deterministic counterparts, and the DP approach has been readily extended (under specific assumptions) to the stochastic framework. "Pontryagin like" or "variational" approaches appeared much later in the literature [25] and we shall come back to explanations for this fact. SOC is used to deal with *dynamic* problems. The notion of *feedback*, as naturally delivered by the DP approach, plays a central part in this area.

Stochastic Programming (SP), which can be traced back to such early contributors as Dantzig [50], is the extension of Mathematical Programming to the stochastic framework. As such, the initial emphasis is on optimization, possibly in a *static* setting, and numerical resolution methods are based on variational techniques; randomness is generally addressed by appealing to the Monte Carlo technique which, roughly speaking, amounts to representing this uncertainty through the consideration of several "samples" or "scenarios". This is why, historically, the notions of *feedback* and *information* were less present in SP than they were in SOC.

However, the SP community[2] has progressively considered two-stage, and then multi-stage problems. Inevitably, the question of *information structures* popped up in the field, at least to handle the elementary constraint of *nonanticipativeness*: one should not assume that the exact realizations of random variables at and after stage $t+1$ are known when making decisions at stage $t$; only a probabilistic description of future occurrences can be taken into account.

It is therefore natural that the two communities of SOC and SP tend to merge and borrow ideas from each other. The concepts of information and feedback are more developed in the former, and the variational and Monte Carlo approaches are more widespread in the latter. Getting closer to each other for the two communities should perhaps begin with unifying the terminology: as far as we understand, *recourse* in the SP community is used as a substitute for *feedback*. This book is an attempt to close the gap. The comparison between SOC and SP approaches is already addressed by Varaiya and Wets in this interesting paper [148].

### *1.1.3 The Role of Information in the Presence of Uncertainty*

In Deterministic Optimal Control, as mentioned previously, there are two main approaches in connection with Pontryagin's and Bellman's contributions. The former

[2]The official web page of the SP community http://www.stoprog.org/ offers links to several tutorials and examples of applications of SP.

focuses on open-loop controls, whereas the latter provides closed-loop solutions. By open-loop controls, we mean that the decisions are given as a function of *time* only, whereas closed-loop strategies compute the control to be implemented at each time instant as a function of both *time and observations*; the observations may be the state itself.

In fact, there are no discrepancies in the performance achieved by both approaches because, in a deterministic situation, everything is uniquely determined by the decision maker. Therefore, if closed-loop strategies are implemented, one can simulate the closed-loop dynamic system, record the trajectories of state, control and observations variables, substitute those trajectories in the control strategy, and compute an open-loop control history that would generate exactly the same trajectories.

The situation is quite different in an uncertain environment, since trajectories are not predictable in advance (off-line) because they depend on on-line realizations of random variables. Available observations reveal some information about those realizations, at least on *past* realizations (because of *causality*). By using this on-line information, one can do better than simply apply a *blind* open-loop control which has been determined only on the basis of a priori probability laws followed by the random "noises".

This means that the achievable performance is dependent on what we call the *information pattern* or *information structure* of the problem: a decision making problem under uncertainty is not well-posed until the exact amount of information available prior to making every decision has been defined. Open-loop problems are problems in which no actual realization can be observed, and thus, the optimal decisions solely depend on a priori probability laws. In dynamic situations, every decision may depend on certain on-line observations that must be specified. Of course, the optimal decisions also depend on a priori probability laws since, generally, not all random realizations can be observed prior to making decisions, if only because of causality or nonanticipativeness.

Because of these considerations, one must keep in mind that solving stochastic optimization problems, especially in dynamic situations when on-line observations are made available, is not just a matter of optimization, of dealing with conventional constraints, or even of computing or evaluating mathematical expectations (which is generally a difficult task by itself); it is also the question of properly handling specific constraints that we shall call *informational constraints*. Indeed, as this book illustrates, there are essentially two ways of dealing with such constraints. That used by the DP approach is a *functional* way: decisions are searched for as *functions* of observations (feedback laws). But another way, which is more adapted to variational approaches in stochastic optimization, may also be considered: all variables of the problem, including decisions, are considered as random variables or stochastic processes; then the dependency of decisions upon observations must go through notions of *measurability* as used by Measure Theory. We shall call this alternative approach an *algebraic* handling of informational constraints (this terminology stems from the fact that information may be mathematically captured by $\sigma$-algebras, also called $\sigma$-fields, another important notion introduced by Measure Theory). A difficult aspect of numerical resolution schemes is precisely the practical translation of those measurability or algebraic constraints into the numerical problem.

An even more difficult aspect of *dynamic* information patterns is that future information may be affected by past decisions. Such situations are called situations with *dual effect*, a terminology which tries to convey the idea that present decisions have two, very often conflicting, effects or objectives: directly contributing to optimizing the cost function on the one hand, modifying the informational constraints to which future decisions are subject, on the other. Problems with dual effect are generally among the most difficult decision making problems (see again [148] about this topic).

## 1.2 Problem Formulations and Information Structures

In this section, two formulations of stochastic optimization problems are proposed: they pertain to the two schools of SOC and SP alluded to above. The important issue of *information structures* is also discussed.

### *1.2.1 Stochastic Optimal Control (SOC)*

**General Formulation**

We consider the following formulation of a stochastic optimal control (SOC) problem in discrete time: for every time instant $t$, $\boldsymbol{X}_t$ (“state”[3]), $\boldsymbol{U}_t$ (control) and $\boldsymbol{W}_t$ (noise) are all random variables over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. They are related to each other by the *dynamics*

$$\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) \tag{1.1a}$$

which is satisfied $\mathbb{P}$-almost surely for $t = 0, \ldots, T-1$. Here, to keep things simple, $T$, the *time horizon*, should be a given deterministic integer value, but it may be a random variable in more general formulations. The variable $\boldsymbol{X}_0$ is a given random variable. It is convenient to view $\boldsymbol{X}_0$ as a given function of some other random variable called $\boldsymbol{W}_0$, in such a way that all primitive random variables are denoted $\boldsymbol{W}_s$, $s = 0, \ldots, T$, whereas $\boldsymbol{W}$ denotes the corresponding stochastic process $\{\boldsymbol{W}_s\}_{s=0,\ldots,T}$. The purpose is to minimize a cost function

$$\mathbb{E}\left(\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\right) \tag{1.1b}$$

[3]Those quotes around the word *state* become clearer when discussing the *Markovian case* by the end of this subsection.

in which $K$ is the *final* cost whereas $L_t$ is called the *instantaneous* cost. The symbol $\mathbb{E}(\cdot)$ denotes *expectation* w.r.t. $\mathbb{P}$ (assuming of course that the functions involved are measurable and integrable). The minimization is achieved by choosing the control variable $\boldsymbol{U}_t$ at each time instant $t$, but as previously mentioned, this is done after some *on-line* information has been collected (in addition to the *off-line* information composed of the model—dynamics and cost—and the a priori distribution of $\{\boldsymbol{W}_s\}_{s=0,\dots,T}$). This on-line information is supposed to be at least *causal* or *nonanticipative*, that is, the largest possible amount of information available at time instant $t$ is equivalent to the observation of the realizations of the random variables $\boldsymbol{W}_s$ for $s = 0, \dots, t$ (but not beyond $t$ ). In the language of Probability Theory, this amounts to saying that $\boldsymbol{U}_t$, as a random variable, is *measurable* w.r.t. the $\sigma$-field generated by $\{\boldsymbol{W}_s\}_{s=0,\dots,t}$ which is denoted $\mathcal{F}_t$:

$$\mathcal{F}_t = \sigma\big(\{\boldsymbol{W}_s\}_{s=0,\dots,t}\big) \tag{1.1c}$$

(the reader may refer to Appendix B for all those standard notions.) Of course, this $\sigma$-field increases as time passes, that is, $\mathcal{F}_t \subset \mathcal{F}_{t+1}$: it is then called a *filtration*.

*Remark 1.1* Observe that in the right-hand side of (1.1a), $\boldsymbol{U}_t$ must be chosen *before* $\boldsymbol{W}_{t+1}$ is observed: this is called the *decision-hazard* framework, as opposed to the *hazard-decision* framework in which the decision maker plays after "nature" at each time stage. This is why we put $\boldsymbol{W}_{t+1}$ rather than $\boldsymbol{W}_t$ in the right-hand side of (1.1a). ◊

It may be that $\boldsymbol{U}_t$ is constrained to be measurable w.r.t. some $\sigma$-field $\mathcal{G}_t$ smaller than $\mathcal{F}_t$:

$$\boldsymbol{U}_t \text{ is } \mathcal{G}_t\text{-measurable, } \mathcal{G}_t \subset \mathcal{F}_t, \quad t = 0, \dots, T-1. \tag{1.1d}$$

Unlike $\mathcal{F}_t$, the $\sigma$-field $\mathcal{G}_t$ is not necessarily increasing with $t$ (see hereafter).

**Information Structure**

Very often, $\mathcal{G}_t$ itself is a $\sigma$-field generated by some random variable $\boldsymbol{Y}_t$ called *observation*. Actually, $\boldsymbol{Y}_t$ should be considered as the collection of *all* observations available at $t$. That is, if $\boldsymbol{Z}_t$ denotes a new observation made available at $t$, but if the decision maker has *perfect memory* of all observations made so far, then $\boldsymbol{Y}_t = \{\boldsymbol{Z}_s\}_{s=0,\dots,t}$. In this case, as for $\mathcal{F}_t$, the $\sigma$-field $\mathcal{G}_t$ is increasing with $t$, but this is not necessarily always true.

The $\sigma$-fields $\mathcal{F}_t$, generated by $\{\boldsymbol{W}_s\}_{s=0,\dots,t}$, are of course only dependent upon the data of the problem, and this is also the case of the $\mathcal{G}_t$ if the observations $\boldsymbol{Y}_t$ are solely dependent on the primitive random variables $\boldsymbol{W}_s$. But if the observations depend also on the controls $\boldsymbol{U}_s$ (for example, if $\boldsymbol{Z}_t$ is a function of the "state" $\boldsymbol{X}_t$, possibly a function corrupted by noise), it is likely that the $\sigma$-field $\mathcal{G}_t$ depends on controls too, and therefore, the measurability constraint (1.1d) is an implicit constraint in that control is subject to constraints depending on controls! Fortunately, thanks to causality, this implicit character is only apparent, that is, the constraint on $\boldsymbol{U}_t$ depends on controls $\boldsymbol{U}_s$ with $s$ strictly less than $t$.

Nevertheless, this is generally a source of huge complexity in SOC problems which is known under the name of the *dual effect* of control. This terminology tries to convey the fact that when making decisions at every time instant $s$, the decision maker has to take care of the following double effect: on the one hand, his decision affects cost (directly, at the same time instant, and in the future time instants, through the "state" variables); but, on the other hand, it makes the next decisions $\boldsymbol{U}_t$, $t > s$ more or less constrained through (1.1d).

*Example 1.2* Let us give an example of this double or dual effect in the real life: the decision of investing in research in any industrial activity. On the one hand, investing in research costs money. On the other hand, an improved knowledge of the field of activity may help save money in the future by allowing better decisions to be made. This example shows that this future effect is very often contradictory with immediate cost considerations and thus the matter of a trade-off to be achieved. △

We now return to our general discussion of information structure in SOC problems. Even if the observations $\boldsymbol{Y}_t$ depend on past controls, it may happen than the $\sigma$-fields $\mathcal{G}_t$ they generate *do not* depend on those controls. This tricky phenomenon is discussed in Chap. 10. Apart from this rather exceptional situation, there are other circumstances when things turn out to be less complex than it may have seemed a priori.

The most classical such case is the *Markovian case*. Suppose the stochastic process $\boldsymbol{W}$ is a "white noise", that is, the random variables $\{\boldsymbol{W}_s\}_{s=0,\dots,T}$, are all mutually independent. Then, $\boldsymbol{X}_t$ truly deserves the name of the *state* variable at time $t$ (this is why, until now, we put the word "state" between quotes—see Footnote 3). Indeed, because of this assumption of white noise, the past realizations of the noise process $\boldsymbol{W}$ provide no additional information about the likelihood of future realizations. Hence, remembering $\boldsymbol{X}_t$ is sufficient information to keep to predict the future evolution of the system after $t$. That is, $\boldsymbol{X}_t$ "summarizes" the past and additional observations are therefore useless. The *Markovian case* is defined as the situation when $\boldsymbol{W}$ is a white noise stochastic process and $\mathcal{G}_t$ is generated at each time $t$ by the variable $\boldsymbol{X}_t$. Otherwise stated, the available observation $\boldsymbol{Y}_t$ at time $t$ is simply $\boldsymbol{X}_t$. This is a *perfect (noiseless) and full size* observation of the state vector. If the observation is *partial* (a non injective function of $\boldsymbol{X}_t$) and/or a *noisy* such function, then the Markovian situation is broken.

In the Markovian case, $\mathcal{G}_t$ *does depend*, in general, upon past controls $\boldsymbol{U}_s$, $s < t$, but we *would not do better* with $\mathcal{F}_t$ replacing $\mathcal{G}_t$. This is why the Markovian case, although potentially falling into the most difficult category of problems with a dual effect, is not so complex as more general problems in this category. The Markovian feature is exploited by the Dynamic Programming (DP) approach (see Sect. 4.4) which is conceptually simple, but quickly becomes numerically difficult, and, indeed, impossible when the dimension of the state vector $\boldsymbol{X}_t$ becomes large.

### *1.2.2 Stochastic Programming (SP)*

**Formulation**

Here we consider another formulation of stochastic optimization problems which ignores "intermediate" variables (such as the "state" $\boldsymbol{X}$ in the previous SOC formulation) and which concentrates on the essential items, namely, the

**control** or **decision** $\boldsymbol{U}$: a random variable over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in a measurable space $(\mathbb{U}, \mathcal{U})$;

**noise** $\boldsymbol{W}$: another random variable with values in a measurable space $(\mathbb{W}, \mathcal{W})$;

**cost function**: a measurable mapping $j : \mathbb{U} \times \mathbb{W} \to \mathbb{R}$;

$\sigma$**-fields**: $\mathcal{F}$ denotes the $\sigma$-field generated by $\boldsymbol{W}$ whereas $\mathcal{G}$ denotes the one w.r.t. which $\boldsymbol{U}$ is constrained to be measurable; generally, $\mathcal{G}$ is generated by an

**observation** $\boldsymbol{Y}$: another random variable with values in a measurable space $(\mathbb{Y}, \mathcal{Y})$; in this case, we use the notation

$$\boldsymbol{U} \preceq \boldsymbol{Y} \tag{1.2}$$

to mean that $\boldsymbol{U}$ is measurable w.r.t. (the $\sigma$-field generated by) $\boldsymbol{Y}$. As we see in Chap. 3, this relation between random variables corresponds to an order relation. We also use this notation in constraints as $\boldsymbol{U} \preceq \mathcal{G}$ to mean that the random variable $\boldsymbol{U}$ is measurable w.r.t. the $\sigma$-field $\mathcal{G}$.

With these ingredients at hand, the problem under consideration is set as follows:

$$\min_{\boldsymbol{U} \preceq \mathcal{G}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big) \quad \text{or} \quad \min_{\boldsymbol{U} \preceq \boldsymbol{Y}} \mathbb{E}\big(j(\boldsymbol{U}, \boldsymbol{W})\big). \tag{1.3}$$

Without going into detailed technical assumptions, we assume that expectations do exist, and that infima are reached (hence the use of the min symbol).

**Typology of Information Structures**

According to the nature of $\mathcal{G}$ or $\boldsymbol{Y}$, we distinguish the following three cases.

**Open-loop optimization**: this is the case when $\mathcal{G}$ is the trivial $\sigma$-field $\{\emptyset, \Omega\}$, or equivalently, $\boldsymbol{Y}$ is any deterministic variable (that is, a constant map over $\Omega$). In this case, an optimal decision is based solely on the a priori (off-line) knowledge of the model, and not on any on-line observation. Therefore, the decision itself is a deterministic variable $u \in \mathbb{U}$ which must minimize a cost function $J(u)$ defined as an expectation of $j(u, \boldsymbol{W})$. The numerical resolution of such problems is considered in Chap. 2.

**Static Information Structure (SIS)**: this is the case when $\mathcal{G}$ or $\boldsymbol{Y}$ are non trivial but *fixed*, that is, a priori given, independently of the decision $\boldsymbol{U}$. The terminology "static" does not imply that no dynamics such as (1.1a) are involved in the problem formulation. It just expresses that the $\sigma$-field $\mathcal{G}$ constraining the decision is a priori given at the problem formulation stage. If time $t$ is involved, one must rewrite the measurability constraint as prescribed at each time stage $t$ as "$\boldsymbol{U}_t$ is $\mathcal{G}_t$-measurable" as in (1.1d), and this *does* leave room for information made available on-line as time evolves. "Static" just says that this on-line information cannot be manipulated by past controls.

*Remark 1.3* When the collection $\{\boldsymbol{U}_s\}_{s=0,\dots,T-1}$ of random variables is interpreted as a random vector over the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, then its measurability is characterized by the $\sigma$-field $\sigma(\{\boldsymbol{U}_s\}_{s=0,\dots,T-1})$ on $(\Omega, \mathcal{A})$. However, with this interpretation, the collection of constraints (1.1d) cannot in general be reduced to a single "vector" constraint $\boldsymbol{U} \preceq \mathcal{G}$ where $\boldsymbol{U}$ would be the "vector" $\{\boldsymbol{U}_s\}_{s=0,\dots,T-1}$ and $\mathcal{G}$ a $\sigma$-field on $(\Omega, \mathcal{A})$, like $\sigma(\{\boldsymbol{U}_s\}_{s=0,\dots,T-1})$ is. For example, over a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, with $T = 2$, $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and $\mathcal{G}_1 = \mathcal{A}$, consider a random variable $\boldsymbol{U}_1$ such that $\sigma(\boldsymbol{U}_1) = \mathcal{A}$. Writing $\boldsymbol{U} \preceq \mathcal{G}$ implies that $\mathcal{G}$ would be the $\sigma$-field $\mathcal{A}$, which does not translate that $\boldsymbol{U}_0$ must be a *constant* (deterministic) variable as implied by $\boldsymbol{U}_0 \preceq \mathcal{G}_0$. ♢

*Remark 1.4* If $\mathcal{G}$ is generated by an observation $\boldsymbol{Y}$, either $\boldsymbol{Y}$ does not depend on $\boldsymbol{U}$, or the $\sigma$-field it generates is fixed despite $\boldsymbol{Y}$ does depend on $\boldsymbol{U}$ (as already mentioned, this may also happen in some special situations addressed in Chap. 10). One may also wonder whether $\boldsymbol{Y}$ has any relation with $\boldsymbol{W}$, for example, whether $\boldsymbol{Y}$ is given as a function $h(\boldsymbol{W})$, in which case $\mathcal{G}$ would be a sub-$\sigma$-field of $\mathcal{F}$, the $\sigma$-field generated by $\boldsymbol{W}$. For example, in the SOC problem (1.1), $\boldsymbol{Y}_t$ may be the complete or partial observation of past noises $\boldsymbol{W}_s$, $s = 0, \dots, t$, so that $\mathcal{G}_t \subseteq \mathcal{F}_t \subset \mathcal{F}_T$. Nevertheless, the fact that $\boldsymbol{Y}$ does or does not have a connection with $\boldsymbol{W}$ is not fundamental. Indeed, by manipulating notation, one can consider that this connection does exist. As a matter of fact, one can redefine the noise variable as the couple $\boldsymbol{W}' = (\boldsymbol{W}, \boldsymbol{Y})$ so that $\boldsymbol{Y}$ *is* a function of $\boldsymbol{W}'$. That the cost function $j$ does not depend on the "full" $\boldsymbol{W}'$ does not matter. ♢

**Dynamic Information Structure (DIS)**: this is the situation when $\mathcal{G}$ or $\boldsymbol{Y}$ depends on $\boldsymbol{U}$, which yields a seemingly implicit measurability constraint. Actually, it is difficult to imagine such problems without explicitly introducing several stages at which decisions must be taken based on observations which may depend on decisions at other stages.

Those stages may be a priori ordered, and the order may be a total order. This is the case of SOC problems (1.1); but other examples are considered hereafter in which those stages are not directly interpreted as "time instants" but rather as "agents" acting one after the other. As soon as such a total order of stages can be defined a priori, the notion of *causality* (who is "upstream" and who is "downstream") is natural and helps untangling the implicit character of the measurability constraint. Nevertheless, the difficulty of such problems with DIS still remains sometimes tremendous as it is shown with help of an example in Sect. 1.3.3.

More general problems may arise in which the order of stages or agent actions is only partial, and the situation may be even worse if this order itself depend on outcomes of the decisions and/or of hazard. At least in the case of a fixed but partial order, it turns out that two notions are paramount for the level of difficulty of the problem resolution:

- Who influences the available observations of whom?
- Who knows more than whom?

We shall not pursue the discussion of this difficult topic here. It is more thoroughly examined in Chap. 9. The forthcoming examples help us scratch the surface.

## 1.3 Examples

This section introduces a few simple examples in order to illustrate the impact of information structures on the formulation of stochastic optimization problems. The stress is more on this aspect than on being fussy about mathematical details (in particular, we assume that all expectations make sense without going into more precise assumptions).

### *1.3.1 A Basic Example in Static Information*

Consider two given scalar random variables, $\boldsymbol{W}$ and $\boldsymbol{Y}$, plus the decision $\boldsymbol{U}$, and finally the following problem of type (1.3):

$$\min_{\boldsymbol{U} \preceq \boldsymbol{Y}} \mathbb{E}\big((\boldsymbol{W} - \boldsymbol{U})^2\big). \tag{1.4}$$

It is well known that the solution of this problem, which consists in finding the best approximation of $\boldsymbol{W}$ which is $\boldsymbol{Y}$-measurable (that is, the projection of $\boldsymbol{W}$ onto the subspace of $\boldsymbol{Y}$-measurable random variables), is given by $\boldsymbol{U}^\sharp = \mathbb{E}(\boldsymbol{W} \mid \boldsymbol{Y})$, that is, the conditional expectation of $\boldsymbol{W}$ knowing $\boldsymbol{Y}$ (see Sect. 3.5.3 and Definition B.5).

Generally speaking, as we see it later on in Sects. 3.5.2 and 8.3.5, Problem (1.3) can be reformulated as follows:

$$\mathbb{E}\Big(\min_{u \in \mathbb{U}} \mathbb{E}\big(j(u, \boldsymbol{W}) \bigm| \boldsymbol{Y}\big)\Big). \tag{1.5}$$

In this form, since the conditional expectation subject to minimization is indeed a $\boldsymbol{Y}$-measurable random variable, it should be understood that the minimization operates parametrically for every realization driven by $\omega$ and this yields an arg min also parametrized by $\omega$, that is, in fact, a random variable which is also $\boldsymbol{Y}$-measurable. When using this new formulation for Problem (1.4), the solution is readily derived (Hint: expand the square in the cost function and observe that $\boldsymbol{Y}$-measurable random variables "get out" of the inner conditional expectation).

## 1.3.2 The Communication Channel

### Description of the Problem

This is the story of two agents trying to communicate through a noisy channel. This story is depicted in Fig. 1.1. The first agent (called the "encoder") gets a "message", here simply a random variable $\boldsymbol{W}_0$ supposed to be centered ($\mathbb{E}(\boldsymbol{W}_0) = 0$), and he wants to communicate it to the other agent. We may consider that the encoder's observation $\boldsymbol{Y}_0$ is precisely this $\boldsymbol{W}_0$. He knows that the channel adds a noise, say a centered random variable $\boldsymbol{W}_1$, to the message he sends, and so he must choose which "best" message to send. He has to "encode" the original signal $\boldsymbol{Y}_0$ into another variable $\boldsymbol{U}_0$ (what he decides to send through the channel), but the other agent (the "decoder") receives a noisy message $\boldsymbol{U}_0 + \boldsymbol{W}_1$. Finally, the decoder has to make his decision $\boldsymbol{U}_1$ about what was the original message $\boldsymbol{W}_0$, based on his observation, namely $\boldsymbol{Y}_1 = \boldsymbol{U}_0 + \boldsymbol{W}_1$, the message he received. That is, he has to "decode", in an "optimal" manner, the signal $\boldsymbol{Y}_1$ which is his observation.

This game is *cooperative* in that the encoder and the decoder try to help each other so as to reduce the error of communication as much as possible (a problem in "team theory" [104], which deals with decision problems involving several agents or decision makers with a common objective function but possibly different observations). Mathematically, this can be expressed by saying that they seek to minimize the expected square error $\mathbb{E}\big((\boldsymbol{U}_1 - \boldsymbol{W}_0)^2\big)$. However, without any other limitation or penalty, such a problem turns out to be rather trivial. For example, if the encoder sends an amplified signal $\boldsymbol{U}_0 = k\boldsymbol{Y}_0$ where $k$ is an arbitrarily large constant, then the noise $\boldsymbol{W}_1$ added by the channel is negligible in front of this very large signal, and the decoder can then decode it by dividing it by the same constant $k$. For the game to be interesting and realistic, one must put a penalty on the "power" $\mathbb{E}(\boldsymbol{U}_0^2)$ sent over the channel, either with help of a constraint limiting this power to a maximum level, or by introducing an additional term proportional to this power into the cost. To stay closer to the generic formulation (1.3), we choose the latter option. Finally, the problem under consideration is the following:

$$\min_{\boldsymbol{U}_0, \boldsymbol{U}_1} \mathbb{E}\big(\alpha \boldsymbol{U}_0^2 + (\boldsymbol{U}_1 - \boldsymbol{W}_0)^2\big) \tag{1.6a}$$

$$\text{s.t.} \quad \boldsymbol{U}_0 \preceq \boldsymbol{Y}_0 \,, \quad \boldsymbol{U}_1 \preceq \boldsymbol{Y}_1 \,. \tag{1.6b}$$

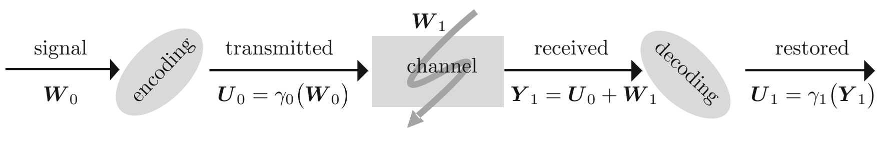


**Fig. 1.1** Communication through a noisy channel