Chapter 1

Introduction

The **S**tochastic **D**ynamic **D**istance Optimal **P**artitioning (SDDP) problem - an Operations Research problem - was the motivation for the investigations presented in this book.

As evident from the name of the problem, investigations in two different mathematical fields were necessary for its treatment, i.e. in stochastic dynamic programming and in combinatorics ("Partitioning").

This book therefore, apart from the introduction, covers the following three chapters

2 DA Stochastic Dynamic Programming with Random Disturbances,

3 The Problem of Stochastic Dynamic Distance Optimal Partitioning (SDDP),

4 Partitions-Requirements-Matrices (PRMs).

DA ("decision after") stochastic dynamic programming with random disturbances is characterized by the fact that these random disturbances are observed before the decision is made at each stage.

In the past only very moderate attention was given to problems with this characteristic (see also Section 1.1).

Examples of DA models are SDDP problems and certain inspection-replacement problems. (Also refer to connections with k-server problems and metric task systems at the end of Section 1.2.)

In Chapter 2 specific properties of DA stochastic dynamic programming problems are worked out for theoretical characterization and for more efficient solution strategies of such problems.

In order to understand this chapter, and the book altogether, previous knowledge about stochastic dynamic programming and Markov decision processes (MDPs) is useful, however not absolutely necessary since the concerned models are developed from scratch. (Basic knowledge can be found in [7], [12], [28], [16] or [31].)

In Chapter 3 we formulate and discuss in detail the problem of **S**tochastic **D**ynamic **D**istance Optimal **P**artitioning (**SDDP**).

SDDP problems are extremely complex.

Superordinately regarded, SDDP problems are DA stochastic dynamic programming problems (**S**tochastic **D**ynamic DP).

It requires a certain initial effort, however, in order to compute the real input data for the DA stochastic dynamic programming problem (SD **D**istance optimal P).

Furthermore, the problem shows combinatorial aspects (SDD **P**artitioning).

The understanding for the formulation of the problem and the basic methods of its solution requires knowledge from Section 2.1 (at least from the beginning of this section) and absolutely from Section 2.3.

However, an important statement concerning certain SDDP problems is proven at the end of Chapter 4, only after several combinatorial considerations.

Partitions-requirements-matrices (PRMs) (Chapter 4) are matrices of transition probabilities of SDDP problems which are formulated as Markov decision processes (MDPs).

PRMs "in the strict meaning" include optimal decisions of certain SDDP problems, as is shown toward the end of Chapter 4.

PRMs (in the strict meaning) themselves represent interesting (almost self-evident) combinatorial structures, which are not otherwise found in literature.

We therefore ensure that the treatise of Chapter 4 can essentially be understood independent of Chapters 2 and 3.

Relationships to Chapter 3 specifically marked and they can be omitted
if one is only interested in PRMs.

Retrospectively, in relation to the topic of "optimal dominant policies"
of MDPs, PRMs in the strict meaning include policies of certain SDDP
problems for which the "condition of dominance" is typically infringed on,
however only to a slight extent such that a generalization of the concept of
"dominant policies" seems possible.

We now discuss the contents of the chapters in more detail.

1.1   Chapter 2 Contents

In Section 2.1 we introduce the DA model of stochastic dynamic program-
ming with random disturbances and give the corresponding functional equa-
tion.

In Section 2.2 a "certainty equivalence principle" is formulated and also
proven in cases of DA models with linear dynamics and quadratic criteria.

Markov decision processes which result from DA models under appropriate
assumptions (DA MDPs) are investigated in Section 2.3.

In literature the state space, which is used for DA MDPs, is the cross
product set of the origin state space and the disturbance space.
However, such a state space is markedly larger than the original state
space.
Moreover, corresponding matrices of transition probabilities would have
many zeros, in general. An analogous situation is found in linear program-
ming: the classical transportation problem which can be solved by the Sim-
plex algorithm. Special solution methods for this transportation problem
have been developed (for example the "MODI-method", refer to [30], Sec-
tion 2.8.9).

In Section 2.3 we keep the origin state space when modelling DA models
as MDPs. In this way special structures of decisions follow.

Here, the corresponding decisions are characterized by a "simple" structure. The transition probability matrices differ by only two elements for corresponding "neighbouring" decisions.

An effect of this structure of decisions is that optimal decisions imply an "almost-partial order" of the states, if the underlying average one-step reward functions do not depend on the decisions.

Thus, the solution of a DA MDP by solving a corresponding parameterized DA MDP in terms of a continuation of the solutions of the parameterized problem arises as one variant for solving DA MDPs, for which the Howard algorithm (policy iteration) is adapted (Section 2.3.4). For this, the underlying internal costs and hence the average one-step reward functions are considered in dependence on one parameter such that these costs do not depend on the decisions for the initial parameter. Then, the adapted Howard algorithm yields a purposeful computation for the solution. Furthermore, under certain additional conditions, this solution method is a greedy algorithm.

Section 2.3.3 includes special considerations of DA MDPs with "distance properties" and "dominant policies".

"Distance properties" can also be found in flow problems, metric task system or k-server problems. In particular, we use the statements of this section for SDDP problems.

The "dominance of Markov chains" can be seen in Daley 68 (see [10]).
We can apply this concept to Markov chains which correspond to policies of MDPs. However, if we want to transfer this concept to the MDPs themselves then convenient properties are also required for the average one-step reward functions (and for the corresponding policies).
If dominant policies should also be optimal, further strong conditions (which contain comparisons of any feasible policies with the dominant policy) are required.
The question which follows is: can we find (useful) MDPs which fulfil all of these conditions?

A certain kind of equipment replacement models with dominant policies can be found in Puterman [31]. However, in these models only two different decisions are possible.

The chance of finding MDPs with more than two decisions which fulfil these conditions is better for MDPs which are based on DA models, due to their decision structures.

Some SDDP problems have optimal dominant policies (Section 4.6.2.2).
For other SDDP problems we will consider the above-mentioned interesting effect in which the conditions of dominance are infringed on, however only to a slight extent.

The state spaces of SDDP problems are inherently finite. Therefore, we will also concentrate our efforts on finite-state models in Chapter 2. Notes on countable-state models can be found in Puterman [31]; more information can be found here at the beginning of Section 2.3.

1.2 Chapter 3 Contents

In Chapter 3 the "Problem of Stochastic Dynamic Distance Optimal Partitioning (SDDP)" is described in detail. Possibilities and methods of its exact or approximate solution are discussed.

A problem in industry, which contains an optimal conversion of moulds, supplied the origin of investigations.
Essentially, SDDP problems include the following practical facts:

- A fixed number of machines is given. $(*)$
  (Moulds are also conceivable.)

- Different types of parts can be produced by these machines. For this purpose the machines have to be converted to states, which in accordance with the types of the parts. Costs are incurred. $(**)$

- The production takes place in successive stages (periods).

- In a single stage, one part (at most) can be produced by one machine.

- At each stage a requirement of parts (of several types) is to be met.

  Initially, probability functions of the requirements are given.

The realizations of the requirements are known at the beginning of the stages (before decisions about conversions of machines have to be made).

· The objective is to minimize the expected cost of the conversions over all stages (or the average expected cost per stage). (To accomplish this we must decide which machine should be converted into which state in each stage.)

Thus, SDDP problems are DA stochastic dynamic programming problems.

More specifically, from a mathematical view point, we could designate this practical problem as a stochastic dynamic transportation problem, since throughout the stages feasible solutions of transportation problems must be determined (see (∗∗)). (We have also used this designation in previous papers.)

Here, however designating this problem as a stochastic dynamic distance optimal partitioning problem (SDDP) seems more appropriate. Partitioning means partitions of the number of machines into numbers of machines which are in the same state. The number of machines is therefore constant (see (∗)).

We will thus use this designation in the future.

(In this way we also emphasize the conceptual distinguishment of the designation of our problem from the typical stochastic dynamic transportation problems, see Arnold [4].) [1]

In this mathematical model, partitions of integers are the "states" of the DA stochastic dynamic programming problems (ordered partitions in general and unordered partitions in the case of certain reduced SDDP problems).

Partitioning the integers as "states" involves the combinatorial aspects of SDDP problems, which can also be observed in "matrices of transition probabilities" and "average one-step reward functions" of SDDP problems, modelled as DA MDPs.

It can therefore, only in Chapter 4 by means of combinatorial consideration, be shown that decisions for feasible states with least square sums of

---

[1]Further comments in connection with transportation problems and corresponding references can be found in the preface of [22].

their parts are in every case optimal for special SDDP problems.

Partitions of integers as states of DA MDPs require an enormous amount of storage space for the corresponding computer programs.

Furthermore, many transportation problems have to be solved (see (∗∗)) in order to compute "average one-step reward functions" for the SDDP problems, modelled as DA MDPs.

Thus, investigations of inherent characteristic structures of SDDP problems are also important as a basis for heuristics.

Finally, we refer to connections of SDDP problems with other problems in operations research and informatics such as stochastic dynamic facility location problems (refer to [27]) or metric task systems and more specific k-server problems, see [8], Chapter 10 and [5], for instance.

Since the current request, which is to be fulfilled, is known (and without knowing the future requests) k-server problems can also be initially labeled as a certain kind of DA model. Furthermore, distance properties are also assumed for k-server problems. However, on-line algorithms are often the center of attraction for consideration of k-server problems.

In contrast, we assume probability functions for requirements of SDDP problems and consider SDDP problems as stochastic dynamic programming problems with the aim to minimize the expected cost or the average expected cost per stage. Typical characteristics of SDDP problems as stochastic dynamic programming problems, in particular Markov decision process, are worked out.

Furthermore, let us note that we consider a number of machines which are in the same state (in the terms of k-server problems, on the same point), in general, and many machines must convert at the beginning of each equidistant stage.

1.3   Chapter 4 Contents

Partitions-Requirements-Matrices (PRMs) are the main topic of Chapter 4.

If SDDP problems are modelled as DA MDPs, then the matrices of transition probabilities are called "general PRMs".

The strict meaning of PRMs assumes that the costs of converting the machines into different types are identical and the requirements are identically distributed. Then in every case decisions for feasible states with least square sums of their components lead to PRMs (in the strict meaning).

The definition of PRMs (in the strict meaning) includes that PRMs can be initially computed by means of simple enumeration, however a laborious method. In addition, there is a main difficulty to deal with: No formulas are known for most of the elements in PRMs. Due to this lack of formulas, PRMs themselves represent interesting (almost self-evident) combinatorial structures.

Properties which are associated with SDDP problems (modelled as DA MDPs), besides the search for effective methods to compute the elements of PRMs, are in the realm of investigation of PRMs (in the strict meaning) in this chapter.

Thus in Section 4.6 so-called "Poisson equations" are considered. That their solutions are "monotone" is shown in many cases. This means that, in every case, decisions for feasible states with least square sums of their components are optimal for the corresponding SDDP problems.

The above-mentioned SDDP problems, for which the "condition of dominance" is infringed on, however only to a slight extent, are also in this set of SDDP problems.

A more detailed specification of the content of Chapter 4 can be found at the beginning of this chapter.