Contamination for multistage stochastic programs

Jitka Dupačová

Abstract: Contamination technique will be examined as a possible approach to robustness analysis of results obtained for multistage stochastic linear programs with respect to changes of their structure or of the input data. We shall focus on the case when the already selected scenario tree gets extended for additional (stress or out-of-sample) scenarios and/or additional stages.

MSC 2000: 90C15, 90C31
Key words: Scenario-based stochastic linear programs, stages, contamination

1 Multiperiod and multistage stochastic programs

When formulating multistage stochastic programs it is common to fix first the horizon and the sequence of times at which decisions will be made. An important requirement is that the decisions must be nonanticipative, i.e. in any stage of the decision process they are allowed to depend only on the past observations and decisions. Discrete approximations of the data process may be available at much finer timestep than the intervals between these decision points, see e.g. [2, 11]. The crucial task is then to relate the time instants and stages.

In the general *T-stage stochastic program* we think of a stochastic data process $\omega = (\omega_1, \ldots, \omega_{T-1})$ and a decision process $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots \boldsymbol{x}_T)$. The components $\omega_1, \ldots, \omega_{T-1}$ of ω and the decisions $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_T$ are assumed to be random vectors, not necessarily of the same dimension, defined on some probability space $(Z, \mathcal{F}, \mu)$, while $\boldsymbol{x}_1$ is a nonrandom vector-valued variable.

The decision process is *nonanticipative* which means that decisions taken at any stage of the process do neither depend on future *realizations* of stochastic data nor on future decisions, whereas the past information as well as the knowledge of the probability distribution of the data process are exploited. This can be expressed as follows: Let $\mathcal{F}_{t-1} \subseteq \mathcal{F}$ be the σ-field generated by the part $\omega^{t-1,\bullet} := (\omega_1, \ldots, \omega_{t-1})$ of the stochastic data process ω that precedes stage t. The dependence of the t-th stage decision $\boldsymbol{x}_t$ only on the past means that $\boldsymbol{x}_t$ is $\mathcal{F}_{t-1}$-measurable. We denote $\boldsymbol{x}^{t-1,\bullet} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1})$ the sequence of decisions at stages $1, \ldots, t-1$, P the distribution function of ω, P_t denotes the marginal probability distribution of ω_t, and $P_t(\cdot|\omega^{t-1,\bullet})$, $t = 2, \ldots, T-1$, its conditional probability distribution.

The first-stage decisions consist of all decisions that have to be selected before further information is revealed whereas the second-stage decisions are allowed to adapt to this information, etc. In each of the stages, the decisions are limited by

This research is supported by the project "Methods of modern mathematics and their applications" – MSM 0021620839 and by Grant Agency of the Czech Republic (grants 201/05/2340, 402/05/0115). Helpful discussions with Werner Römisch are gratefully acknowledged.

constraints that may depend on the previous decisions and observations. *Stages do not necessarily refer to time periods, they correspond to steps in the decision process.*

An example is the *nested form* of the multistage stochastic linear program (MSLP) which resembles the backward recursion of stochastic dynamic programming with an additive overall cost function:

$$\min_{\boldsymbol{x}_1 \in \mathcal{X}_1} [\boldsymbol{c}_1^\top \boldsymbol{x}_1 + E_P \{\varphi_1(\boldsymbol{x}_1, \omega_1)\}] \text{ with } \mathcal{X}_1 := \{\boldsymbol{x}_1 \,|\, \boldsymbol{A}_1 \boldsymbol{x}_1 = \boldsymbol{b}_1,\, \boldsymbol{l}_1 \leq \ \boldsymbol{x}_1 \leq \ \boldsymbol{u}_1\}\,, \quad (1)$$

and $\varphi_{t-1}(\cdot,\cdot)$, $\ t = 2, \ldots, T$, defined recursively as

$$\varphi_{t-1}(\boldsymbol{x}^{t-1,\bullet}, \omega^{t-1,\bullet}) = \min_{\boldsymbol{x}_t} [\boldsymbol{c}_t(\omega^{t-1,\bullet})^\top \boldsymbol{x}_t + E_{P_t(\cdot|\omega^{t-1,\bullet})} \{\varphi_t(\boldsymbol{x}^{t-1,\bullet}, \boldsymbol{x}_t, \omega^{t-1,\bullet}, \omega_t)\}] \quad (2)$$

subject to constraints

$$\boldsymbol{B}_t(\omega^{t-1,\bullet})\, \boldsymbol{x}_{t-1} + \boldsymbol{A}_t(\omega^{t-1,\bullet})\, \boldsymbol{x}_t = \boldsymbol{b}_t(\omega^{t-1,\bullet}), \quad \boldsymbol{l}_t(\omega^{t-1,\bullet}) \leq \boldsymbol{x}_t \leq \boldsymbol{u}_t(\omega^{t-1\bullet}) \text{ a.s.}$$

and φ_T is explicitly given, e.g. $\varphi_T \equiv 0$.

Matrices $\boldsymbol{A}_t$ are of a fixed (m_t, n_t) type and the remaining vectors and matrices are of consistent dimensions. For the first stage, known values of all elements of $\boldsymbol{b}_1, \boldsymbol{c}_1, \boldsymbol{A}_1, \boldsymbol{l}_1, \boldsymbol{u}_1$ are assumed and the main decision variable is $\boldsymbol{x}_1$ that corresponds to the first stage. The first-stage problem (1) has the form of the expectation-type stochastic program with the set of feasible decisions *independent* of P. Constraints involving random elements hold almost surely. For simplicity we will assume that *all infima are attained*, which is related with the *relatively complete recourse*, and that *all expectations exist.* See recent books [16] or [18] for more general cases.

For applications one mostly approximates the true probability distribution P of ω by a discrete probability distribution carried by a finite number of atoms, say, $\omega^1, \ldots, \omega^K$; see [10] for a survey of scenario generation techniques. Accordingly, the supports of marginal and conditional probability distributions P_t, $P_t(\cdot|\omega^{t-1,\bullet})\ \forall t$ are finite sets. For disjoint sets of indices $\mathcal{K}_t$, $t = 2, \ldots, T$, let us list as $\tilde{\omega}_{k_t}$, $k_t \in \mathcal{K}_t$ all possible realizations of $\omega^{t-1,\bullet}$ and denote by the same subscripts the corresponding values of the t-th stage coefficients. The total number of scenarios K equals the number of elements of $\mathcal{K}_T$. Each scenario $\omega^k = \{\omega_1^k, \ldots, \omega_{T-1}^k\}$ thus generates a sequence of coefficients $\{\boldsymbol{c}_{k_2}, \ldots, \boldsymbol{c}_{k_T}\}$, $\{\boldsymbol{A}_{k_2}, \ldots, \boldsymbol{A}_{k_T}\}$, $\{\boldsymbol{B}_{k_2}, \ldots, \boldsymbol{B}_{k_T}\}$, $\{\boldsymbol{b}_{k_2}, \ldots, \boldsymbol{b}_{k_T}\}$, $\{\boldsymbol{l}_{k_2}, \ldots, \boldsymbol{l}_{k_T}\}$, $\{\boldsymbol{u}_{k_2}, \ldots, \boldsymbol{u}_{k_T}\}$. The data are organized in the form of the scenario tree: Its nodes are determined by all considered realizations $\tilde{\omega}_{k_t}$, $k_t \in \mathcal{K}_t, t = 2, \ldots, T$, and by the root indexed as $k_1 = 1$; each realization $\tilde{\omega}_{k_{t+1}}$ of $\omega^{t\bullet}, t = 1, \ldots, T$, has a unique ancestor $\tilde{\omega}_{k_t}$ (a realization of $\omega^{t-1,\bullet}$), we denote it by subscript $a(k_{t+1})$, and a finite number of descendants — realizations of $\omega^{t+1,\bullet}$.

This allows to rewrite the T-stage stochastic linear program (1)–(2) in the following *arborescent form*:

$$\min \left[\boldsymbol{c}_1^\top \boldsymbol{x}_1 + \sum_{k_2 \in \mathcal{K}_2} p_{k_2} \boldsymbol{c}_{k_2}^\top \boldsymbol{x}_{k_2} + \sum_{k_3 \in \mathcal{K}_3} p_{k_3} \boldsymbol{c}_{k_3}^\top \boldsymbol{x}_{k_3} + \ldots + \sum_{k_T \in \mathcal{K}_T} p_{k_T} \boldsymbol{c}_{k_T}^\top \boldsymbol{x}_{k_T} \right] \quad (3)$$

subject to

$$\begin{array}{lllll}
\boldsymbol{A}_1\boldsymbol{x}_1 & & & = \boldsymbol{b}_1 & \\
\boldsymbol{B}_{k_2}\boldsymbol{x}_1 + \boldsymbol{A}_{k_2}\boldsymbol{x}_{k_2} & & & = \boldsymbol{b}_{k_2}, & k_2 \in \mathcal{K}_2 \\
\boldsymbol{B}_{k_3}\boldsymbol{x}_{a(k_3)} + \boldsymbol{A}_{k_3}\boldsymbol{x}_{k_3} & & & = \boldsymbol{b}_{k_3}, & k_3 \in \mathcal{K}_3 \\
\ddots \quad \ddots & & & \vdots & \\
\boldsymbol{B}_{k_T}\boldsymbol{x}_{a(k_T)} + \boldsymbol{A}_{k_T}\boldsymbol{x}_{k_T} & & & = \boldsymbol{b}_{k_T}, & k_T \in \mathcal{K}_T
\end{array}$$

$$\boldsymbol{l}_1 \le \boldsymbol{x}_1 \le \boldsymbol{u}_1,\ \boldsymbol{l}_{k_t} \le \boldsymbol{x}_{k_t} \le \boldsymbol{u}_{k_t},\ k_t \in \mathcal{K}_t,\ t = 2, \dots, T. \tag{4}$$

The *path probabilities* $p_{k_t} > 0\,\forall k_t$, $\sum_{k_t \in \mathcal{K}_t} p_{k_t} = 1$, $t = 2, \dots T$, of partial sequences of coefficients are probabilities of realizations of $\omega^{t-1,\bullet}\,\forall t$. They may be obtained by stepwise multiplication of the marginal probabilities p_{k_2} by the conditional arc (transition) probabilities, say, $\pi_{k_{\tau-1}k_\tau}$, $\tau = 3, \dots, t$. Probabilities p^k of individual scenarios $\omega^k, k = 1, \dots, K$, are equal to the corresponding path probabilities p_{k_T}.

The nonanticipativity constraints are included in an implicit way. Notice, that (3)–(4) may correspond also to a T-period two-stage stochastic program based on the same scenarios: Except for the root, there is only one descendant $d(k_t)$ of each of t-th stage nodes, that is, the transition probabilities $\pi_{k_t,d(k_t)} = 1\,\forall k_t \in \mathcal{K}_t, t = 2, \dots, T-1$. Scenarios are identified by sequences $\{k_2, \dots, k_T\}$ such that $k_t \in \mathcal{K}_T$, $k_{t+1} = d(k_t)\,\forall t$ and the objective function (3) may be simplified to

$$\boldsymbol{c}_1^\top \boldsymbol{x}_1 + \sum_{k_T \in \mathcal{K}_T} p_{k_T}[\boldsymbol{c}_{k_2}^\top \boldsymbol{x}_{k_2} + \boldsymbol{c}_{k_3}^\top \boldsymbol{x}_{k_3} + \ldots + \boldsymbol{c}_{k_T}^\top \boldsymbol{x}_{k_T}]. \tag{5}$$

Problem (5), (4) is called the *two-stage relaxation* of MSLP (3)–(4).

With explicit inclusion of nonanticipativity constraints, the scenario-based multiperiod or multistage stochastic programs with linear constraints can be again written as a large-scale deterministic program: Given scenario ω^k denote by $\boldsymbol{c}(\omega^k)$ the vector composed of all corresponding coefficients, say, $\boldsymbol{c}_1, \boldsymbol{c}_{k_t}, t = 2, \dots, T$, in the objective function, by $\boldsymbol{A}(\omega^k)$ the matrix of all coefficients of system of constraints (4) for scenario ω^k, and, similarly, by $\boldsymbol{b}(\omega^k)$, $\boldsymbol{l}(\omega^k)$, $\boldsymbol{u}(\omega^k)$ the vectors composed of right-hand sides in (4) and bounds of the box constraints for scenario ω^k. The *scenario-splitted* form of the T-stage stochastic linear program is

$$\min_{\mathcal{X} \cap \mathcal{C}} \left\{ \sum_{k=1}^{K} p^k \boldsymbol{c}(\omega^k)^\top \boldsymbol{x}^k \mid \boldsymbol{A}(\omega^k)\boldsymbol{x}(\omega^k) = \boldsymbol{b}(\omega^k),\ \boldsymbol{l}(\omega^k) \le \boldsymbol{x}(\omega^k) \le \boldsymbol{u}(\omega^k)\,\forall k \right\}. \tag{6}$$

Set $\mathcal{X}$ is defined by deterministic constraints on $\boldsymbol{x}_t(\omega^k)\,\forall t, k$, $\mathcal{C}$ by the nonanticipativity conditions, and $\boldsymbol{x}(\omega^k)$ is the corresponding decision vector composed of stage related subvectors $\boldsymbol{x}_t(\omega^k) \forall t$. For two-stage stochastic programs the nonanticipativity constraints boil down to the requirement that the first-stage decisions must be scenario independent, i.e. $\boldsymbol{x}_1(\omega^k) = \boldsymbol{x}_1(\omega^{k'})\,\forall k, k'$. Similar constraints guarantee that the t-th stage decisions based on the same history are equal. Such constraints

can be expressed as $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{x}$ where $\boldsymbol{x}$ contains carefully grouped components of all decision vectors $\boldsymbol{x}(\omega^k)$ and $\boldsymbol{U}$ is a 0-1 matrix.

The choice of stages, of the branching scheme, of scenarios and their probabilities influence the optimal first-stage decision and the overall optimal value. To use multiperiod two-stage model or to assign one stage to each of possible discretization points are two extreme cases. Requirements of various applications may lead to different topologies of the decision points: With a fixed time discretization of the data process $\omega = (\omega_t,\, t = 1, \ldots, T-1)$ the stages may be allocated to selected time points, say $\tau_1 < \ldots < \tau_S < T$. The decisions are made at τ_s, $s = 1, \ldots, S$, using the past information $\omega^{\tau_s - 1, \bullet}$ and the probabilistic specification. Similarly as the second stage decisions for multiperiod two-stage stochastic programs, all decisions at time points t *between* stages τ_s, τ_{s+1} are made at $t = \tau_s$ using the past information up to τ_s. The formulation exploits then a fixed suitable coarser structure (filtration) $\{\mathcal{F}_s,\, s = 1, \ldots, S\}$, $\mathcal{F}_s \subseteq \mathcal{F}$ defined by the data available at time τ_s which corresponds to stage s. The whole procedure has been developed in [2] for a specific application, see also [16] for the corresponding scenario tree construction.

It has been observed that various theoretical results valid for two-stage stochastic programs do not carry over to the multistage case (e.g. [7, 15, 20]). At the same time, input generation (i.e. generation of a scenario tree instead of a fan of scenarios) and the numerical solution of multistage programs is substantially more complicated. Hence, a natural question is how many stages and what topology of stages should be used, why to use multistage stochastic programs at all and how much we loose when simplifying them to their multiperiod two-stage variant by relaxation of nonanticipativity constraints. As the set of feasible decisions gets enlarged, the optimal value of the two-stage relaxation *based on identical data* provides a lower bound of the optimal value of the original multistage problem.

The impact of including additional scenarios and/or stages on the results is a problem of stress testing or output analysis. There exist several numerical studies in the context of multistage stochastic programs and their applications, cf. [3, 17]. We shall approach these problems via the contamination technique.

2 Stress testing via contamination

2.1 Contamination technique

Contamination approach was initiated in mathematical statistics as one of the tools for analysis of robustness of estimators with respect to deviations from the assumed probability distribution and/or its parameters. It goes back to von Mises and the concepts are briefly described e.g. in [19]. In stochastic programming, it was developed in a series of papers up to results applicable for two-stage stochastic linear programs, e.g. [4, 6], and to the first ideas dealing with the multistage case [5]. For construction of contamination bounds, it is important that the stochastic

program is reformulated as

$$\min_{\boldsymbol{x}\in\mathcal{X}} F(\boldsymbol{x}, P) := \int_{\Omega} f(\boldsymbol{x}, \omega) P(d\omega) \tag{7}$$

with $\mathcal{X}$ independent of P.

Via contamination, robustness analysis with respect to changes in P gets reduced to a much simple analysis with respect to a scalar parameter λ : Assume that (7) was solved for an already constructed scenario tree corresponding to the discrete probability distribution P. Denote $\varphi(P)$ the optimal value and $\mathcal{X}^*(P)$ the set of optimal solutions. Possible changes in probability distribution P are modeled using contaminated distributions P_λ,

$$P_\lambda := (1-\lambda)P + \lambda Q,\ \lambda \in [0,1]$$

with Q another *fixed* probability distribution. Limiting the analysis to a selected direction only, the results are directly applicable but they are less general than quantitative stability results with respect to arbitrary (but small) changes in P summarized e.g. by Römisch in Chapter 8 of [18].

The objective function in (7) is linear in P, hence

$$F(\boldsymbol{x}, \lambda) := \int_{\Omega} f(\boldsymbol{x}, \omega) P_\lambda(d\omega) = (1-\lambda)F(\boldsymbol{x}, P) + \lambda F(\boldsymbol{x}, Q)$$

is linear in λ. Suppose that the stochastic program (7) has an optimal solution for all considered distributions P_λ, $0 \le \lambda \le 1$. Then the optimal value function

$$\varphi(\lambda) := \min_{\boldsymbol{x}\in\mathcal{X}} F(\boldsymbol{x}, \lambda)$$

is concave on $[0,1]$ which implies its continuity and existence of directional derivatives in $(0,1)$. Continuity at the point $\lambda = 0$ is a property related with stability results for the stochastic program in question. In general, one needs a nonempty, bounded set of optimal solutions $\mathcal{X}^*(P)$ of the initial stochastic program (7). This assumption together with stationarity of derivatives $\frac{dF(\boldsymbol{x},\lambda)}{d\lambda} = F(\boldsymbol{x}, Q) - F(\boldsymbol{x}, P)$ are used to derive the form of the directional derivative

$$\varphi'(0^+) = \min_{\boldsymbol{x}\in\mathcal{X}^*(P)} F(\boldsymbol{x}, Q) - \varphi(0) \tag{8}$$

which enters the upper bound for the optimal value function $\varphi(\lambda)$

$$\varphi(0) + \lambda\varphi'(0^+) \ge \varphi(\lambda) \ge (1-\lambda)\varphi(0) + \lambda\varphi(1),\ \lambda \in [0,1]; \tag{9}$$

see [4, 6] and references therein.

If $\boldsymbol{x}^*(P)$ is the *unique* optimal solution of (7), $\varphi'(0^+) = F(\boldsymbol{x}^*(P), Q) - \varphi(0)$, i.e. the *local change of the optimal value function caused by a small change of P in direction $Q - P$ is the same as that of the objective function at $\boldsymbol{x}^*(P)$*. If there

are multiple optimal solutions, each of them leads to an upper bound $\varphi'(0^+) \leq F(\boldsymbol{x}(P), Q) - \varphi(0)$, $\boldsymbol{x}(P) \in \mathcal{X}^*(P)$. Contamination bounds can be then written as

$$(1-\lambda)\varphi(P) + \lambda F(\boldsymbol{x}(P), Q) \geq \varphi(P_\lambda) \geq (1-\lambda)\varphi(P) + \lambda\varphi(Q) \tag{10}$$

valid for an arbitrary $\boldsymbol{x}(P) \in \mathcal{X}^*(P)$ and $\lambda \in [0,1]$. If $\boldsymbol{x}(P)$ is an ε-optimal solution of (7) for probability distribution Q then the difference of the upper and lower bound in (10) is less or equal to $\lambda\varepsilon$.

Contamination bounds (9), (10) help to *quantify* the change in the optimal value due to the considered perturbations of (7). They were applied in [8] and [1, 9], to stress test of CVaR and of multiperiod two-stage bond portfolio management problems, respectively.

2.2 Contamination for multistage stochastic linear programs

Also multistage stochastic programs can be formulated as (7), with $\mathcal{X}$ the set of feasible first-stage decisions, cf. (1). Still, a note of warning is needed: In (7), the random objective $f(\cdot,\cdot)$ is a given function whereas the random objective $\varphi_1(\cdot,\cdot)$ in (1) changes when the topology of stages, i.e. the filtration, get changed. This indicates that for a fixed topology of stages contamination with respect to additional scenarios goes its usual way. Indeed, the corresponding contamination bounds were derived in [5] for MSLP with respect to additional out-of-sample scenarios, which increase the branching number of selected nodes of the scenario tree but do not change the topology of stages. The results were applied to multistage problems with a *fixed* topology of stages in [1, 13].

Example 2.1 (Out-of-sample scenarios) *Consider for simplicity a 3-stage SLP with random right-hand sides written in the arborescent form; scenarios correspond here to sequences* $\{\boldsymbol{b}_{k_2}, \boldsymbol{b}_{k_3}\}$:

$$\min\,[\boldsymbol{c}_1^\top \boldsymbol{x}_1 + \sum_{k_2 \in \mathcal{K}_2} p_{k_2} \boldsymbol{c}_2^\top \boldsymbol{x}_{k_2} + \sum_{k_3 \in \mathcal{K}_3} p_{k_3} \boldsymbol{c}_3^\top \boldsymbol{x}_{k_3}]$$

subject to nonnegativity conditions and

$$\boldsymbol{A}_1\boldsymbol{x}_1 = \boldsymbol{b}_1,\ \boldsymbol{B}_2\boldsymbol{x}_1 + \boldsymbol{A}_2\boldsymbol{x}_{k_2} = \boldsymbol{b}_{k_2},\ k_2 \in \mathcal{K}_2,\ \boldsymbol{B}_3\boldsymbol{x}_{a(k_3)} + \boldsymbol{A}_3\boldsymbol{x}_{k_3} = \boldsymbol{b}_{k_3},\ k_3 \in \mathcal{K}_3.$$

Let $\boldsymbol{x}_1^*(P)$, $\boldsymbol{x}_{k_t}^*(P)\,\forall t$ *be its optimal solution.*

There are three basic types of additional scenarios:

a. *One new scenario* $\omega^* := \{\boldsymbol{b}_2^*, \boldsymbol{b}_3^*\}$ *is included, contaminating probability distribution* $Q = \delta\{\omega^*\}$ *is degenerated. To get the directional derivative we have to compute the value*

$$F(\boldsymbol{x}^*(P), \delta\{\omega^*\}) = \boldsymbol{c}_1^\top \boldsymbol{x}_1^*(P) + \min\left[\boldsymbol{c}_2^\top \boldsymbol{x}_2 + \boldsymbol{c}_3^\top \boldsymbol{x}_3\right]$$

with minimization over nonnegative $\boldsymbol{x}_2, \boldsymbol{x}_3$ *such that*

$$\boldsymbol{B}_2\boldsymbol{x}_1^*(P) + \boldsymbol{A}_2\boldsymbol{x}_2 = \boldsymbol{b}_2^*, \quad \boldsymbol{B}_3\boldsymbol{x}_2 + \boldsymbol{A}_3\boldsymbol{x}_3 = \boldsymbol{b}_3^*.$$

This type of contamination means that all original scenarios are kept but with path probabilities reduced by factor $1-\lambda$, *new scenario* ω^* *enters with probability* λ.

b. *In the 3rd stage, an additional realization* $\boldsymbol{b}_3^*$ *of* $\boldsymbol{b}_3$ *is considered. Hence, only the conditional probability distribution* $P_2(\boldsymbol{b}_3|\boldsymbol{b}_{a(k_*)})$ *in the 3rd stage gets contaminated which is determined by the considered ancestor* $a(k_*)$ *of the new terminal node. The scenario tree is extended just for one new branch that emanates from* $\boldsymbol{b}_{a(k_*)}$, *its arc probability equals* λ, *arc probabilities of all original branches emanating from the same node* $\boldsymbol{b}_{a(k_*)}$ *at the 2nd stage level are multiplied by* $1-\lambda$ *and the remaining probabilities do not change. For the corresponding degenerated contaminating distribution* Q,

$$F(\boldsymbol{x}^*(P),Q) = \boldsymbol{c}_1^\top \boldsymbol{x}_1^*(P) + \boldsymbol{c}_2^\top \boldsymbol{x}_{a(k_*)}^*(P) + \min \boldsymbol{c}_3^\top \boldsymbol{x}_3$$

with minimization over $\boldsymbol{x}_3 \geq 0$ *that fulfils condition* $\boldsymbol{B}_3\boldsymbol{x}_{a(k_*)}^*(P) + \boldsymbol{A}_3\boldsymbol{x}_3 = \boldsymbol{b}_3^*$.

c. *A new realization* $\boldsymbol{b}_3^*$ *in the 3rd stage is included independently of the past history, i.e. for all realizations* $\boldsymbol{b}_{k_2}$ *of* $\boldsymbol{b}_2$. *It corresponds to contamination of all conditional distributions* $P_2(\boldsymbol{b}_3|\boldsymbol{b}_2 = \boldsymbol{b}_{k_2})$ *by* $\delta\{\boldsymbol{b}_3^*\}$. *Arc probabilities of all original branches of the 3rd stage are multiplied by* $1-\lambda$ *and each bundle of branches emanating from the same ancestor at the 2nd stage level is enlarged for one branch corresponding to* $\boldsymbol{b}_3^*$ *with arc probability* λ. *Hence, in the 3rd stage the problem gets extended for* $\#\mathcal{K}_2$ *systems of equations and subvectors of variables. With the new descendant scenarios and variables indexed by* $d(k_2), k_2 \in \mathcal{K}_2$ *we get*

$$F(\boldsymbol{x}^*(P),Q) = \boldsymbol{c}_1^\top \boldsymbol{x}_1^*(P) + \sum_{k_2 \in \mathcal{K}_2} p_{k_2} \left[\boldsymbol{c}_2^\top \boldsymbol{x}_{k_2}^*(P) + \min \boldsymbol{c}_3^\top \boldsymbol{x}_{d(k_2)}\right]$$

with minimizations over $\boldsymbol{x}_{d(k_2)} \geq 0$ *that fulfil equations* $\boldsymbol{B}_3\boldsymbol{x}_{k_2}^*(P) + \boldsymbol{A}_3\boldsymbol{x}_{d(k_2)} = \boldsymbol{b}_3^*$.

Similarly, the directional derivatives (8) and the contamination bounds (9) can be derived for general scenario-based MSLP provided that both the original problem and the alternative one are related to the same scheme of decision points for distributions P, Q; degenerated distributions Q applied in the above example are a special case. On the other hand, an application to MSLP with a varying topology of stages is not straightforward. Let's introduce first a motivating example.

Example 2.2 *Consider a stochastic dedicated bond portfolio selection problem modeled as two-stage multiperiod stochastic linear program, see e.g. [12, 21]. There are many scenarios of interest rates which enter the coefficients. The problem is solved over* T *time periods for a given portfolio of bonds and for probability distribution* P *carried by the fan of selected scenarios, say* $\omega^k = (\omega_t^k,\ t = 1,\dots,T-1)$, $k = 1,\dots,K$, *with probabilities* p^k. *The first-stage decision* $\boldsymbol{x}_1$ *must be scenario independent whereas the second-stage decisions* $\boldsymbol{x}_t^k$ *depend on scenarios* ω^k *and are constructed at once for all subsequent periods* $t = 2,\dots,T$.

The alternative probability distribution Q *carried by scenarios* ω^h, $h = 1,\dots,H$, *with probabilities* q_h *corresponds to a possible call option at* $t = t_1 > 2$ *for certain*

bond, under some of scenarios. It provides an optimal 1st stage decision if the call option is exercised. The interest rate scenarios are identical with those for P but the related cash flows of the bond with call option differ — the full nominal value plus coupon and premium get paid at t_1 and zero cash flows follow in subsequent periods. Of course, the investment decisions for Q will be different.

Contaminated probability distribution P_λ takes into account both possibilities and the contamination parameter λ reflects the belief that the call option will be exercised. To get a 3-stage SLP one keeps the time horizon T and the time discretization and includes an additional decision point at t_1. This means, inter alia, that the system of linear constraints written for the pooled set of scenarios corresponding to P_λ must be extended for the nonanticipativity condition: Decisions $\boldsymbol{x}_t^k$ at $t < t_1$ cannot count upon the outcome of the option at time t_1.

To detail the idea from Example 2.2, think of including an additional stage (not additional time discretization point!) at $t = t_1$. This means to reflect in the arborescent form of MSLP (3)–(4) for contaminated probability distribution P_λ the additional nonanticipativity conditions: For $t < t_1$, all coefficients and decision variables for P and Q are equal. The corresponding subsystem of constraints in (4) will be called *common constraints*. For $t \geq t_1$ constraints for scenarios ω^k are kept, called P-system, and another Q-system of constraints for scenarios ω^h will be attached. The ancestors $a(h_{t_1})$ and the corresponding decision variables $\boldsymbol{x}_{a(h_{t_1})}$ in the Q-system come from the common constraints. Thus using the pooled set of scenarios from P and Q we get a *fixed* system of linear constraints and the contaminated stochastic program is a *linear parametric program* with parameter λ only in the objective function:

$F(\boldsymbol{x}, \lambda) :=$

$$\boldsymbol{c}_1^\top \boldsymbol{x}_1 + \sum_{t=2}^{t_1-1} \sum_{k_t \in \mathcal{K}_t} p_{k_t} \boldsymbol{c}_{k_t}^\top \boldsymbol{x}_{k_t} + (1-\lambda) \sum_{t=t_1}^{T} \sum_{k_t \in \mathcal{K}_t} p_{k_t} \boldsymbol{c}_{k_t}^\top \boldsymbol{x}_{k_t} + \lambda \sum_{t=t_1}^{T} \sum_{h_t \in \mathcal{H}_t} q_{h_t} \boldsymbol{c}_{h_t}^\top \boldsymbol{x}_{h_t} \tag{11}$$

is minimized with respect to (4) and the Q-system

$$\boldsymbol{B}_{h_t} \boldsymbol{x}_{a(h_t)} + \boldsymbol{A}_{h_t} \boldsymbol{x}_{h_t} = \boldsymbol{b}_{h_t}, \quad \boldsymbol{l}_{h_t} \leq \boldsymbol{x}_{h_t} \leq \boldsymbol{u}_{h_t}, \quad h_t \in \mathcal{H}_t,\, t \geq t_1. \tag{12}$$

The optimal value of (11), (4), (12) is denoted $\varphi(\lambda)$ and $\mathcal{X}^*(\lambda)$ is the set of optimal solutions. The symbols $\varphi(P), \mathcal{X}^*(P), \varphi(Q), \mathcal{X}^*(Q)$ are kept for optimal values and sets of optimal solutions of the two MSLP obtained for P and Q separately; notice that $\varphi(0) = \varphi(P)$ and $\varphi(1) = \varphi(Q)$.

Proposition 2.3 *Assume that the sets $\mathcal{X}^*(\lambda)$ are nonempty for all $\lambda \in [0,1]$ and $\mathcal{X}^*(0)$ is bounded. Then the optimal value function $\varphi(\lambda)$ is concave on $[0,1]$ and contamination bounds (9) follow with $\varphi'(0^+) = \min_{\boldsymbol{x} \in \mathcal{X}^*(0)} F(\boldsymbol{x}, 1) - \varphi(0)$.*

The proof is an adaptation of results on existence and form of directional derivatives of the optimal value function of perturbed linear programs, cf. Chapter 3.5 of [14], to the parametric linear program (4), (11), (12).

To get an upper bound for the derivative means to evaluate $F(\boldsymbol{x}(0), 1)$ for an optimal solution $\boldsymbol{x}(0)$ of the contaminated problem (11), (4), (12) with $\lambda = 0$. These optimal solutions consist of $\boldsymbol{x}(P)$ — an optimal solution of (3)–(4), complemented by an arbitrary feasible solution of the related Q-system (12). Taking an optimal feasible solution $\boldsymbol{x}^*_{h_t}$, $t \geq t_1$ of (12) we get

$$\varphi'(0^+) \leq \sum_{t=t_1}^{T} \left[\sum_{h_t \in \mathcal{H}_t} q_{h_t} \boldsymbol{c}^\top_{h_t} \boldsymbol{x}^*_{h_t} - \sum_{k_t \in \mathcal{K}_t} p_{k_t} \boldsymbol{c}^\top_{k_t} \boldsymbol{x}_{k_t}(P) \right].$$

2.3 Comments

A similar theorem holds true for other instances of scenario-based MSLP, for more complex changes of their structure and it can be also extended to scenario-based nonlinear problems.

Notice that working with the scenario-splitted form (6) would mean to accept changes of the system $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{x}$ if the topology of stages varies which is a substantial change of the resulting deterministic program.

References

[1] M. Bertocchi, J. Dupačová and V. Moriggia. Bond portfolio management via stochastic programming. To appear in: W. T. Ziemba and S. A. Zenios (eds.) *Handbook of Asset & Liability Management*. Elsevier, 2006.

[2] M. A. H. Dempster et al. Planning logistics operations in the oil industry. *J. of ORS*, 51:1271–1288, 2000.

[3] M. A. H. Dempster and G. W. P. Thompson. Dynamic portfolio replication using stochastic programming. In: M. A. H. Dempster (ed.) *Risk Management: Value at Risk and Beyond*. Cambridge Univ. Press 2002, pp. 100–128.

[4] J. Dupačová. Stability in stochastic programming with recourse – contaminated distributions. *Math. Programing Study*, 27:133–144, 1986.

[5] J. Dupačová. Postoptimality for multistage stochastic linear programs. *Annals of Oper. Res.*, 56:65–78, 1995.

[6] J. Dupačová. Scenario-based stochastic programs: Resistence with respect to sample. *Annals of Oper. Res.*, 64:21–38, 1996.

[7] J. Dupačová. Reflections on output analysis for multistage stochastic linear programs. In: K. Marti, Y. Ermoliev and G. Pflug (eds.) *Dynamic Stochastic Optimization*, LNEMS 532, Springer Verlag, Berlin, 2004, pp. 3–20.

[8] J. Dupačová. Stress testing via contamination. To appear in LNEMS (Proceedings of the CwU Workshop of IIASA 2004). Springer Verlag, Berlin, 2006.

[9] J. Dupačová, M. Bertocchi and V. Moriggia. Postoptimality for scenario based financial models with an application to bond portfolio management. In: W. T. Ziemba and J. Mulvey (eds.) *World Wide Asset and Liability Modeling*. Cambridge Univ. Press, 1998, pp. 263–285.

[10] J. Dupačová, G. Consigli and S. W. Wallace. Scenarios for multistage stochastic programs. *Annals of Oper. Res.*, 100:25–53, 2000.

[11] J. Dupačová, N. Gröwe-Kuska and W. Römisch. Scenario reduction in stochastic programming: An approach using probability metrics. *Math. Programming A*, 95:493-511, 2003.

[12] J. Dupačová, J. Hurt and J. Štěpán. *Stochastic Modeling in Economics and Finance*. Kluwer Acad. Publ., Dordrecht, 2002.

[13] J. Dupačová and J. Polívka. Asset-liability management for Czech pension funds using stochastic programming. SPEPS 2004-01, 2004.

[14] E. G. Gol'štejn and D. B. Yudin. *New Directions in Linear Programming* [in Russian]. Sovetskoe Radio, Moscow, 1966.

[15] H. Heitsch, W. Römisch and C. Strugarek. Stability of multistage stochastic programs. To appear in *SIAM J. Optimization*.

[16] P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Theory and Computation*. Springer International Series, New York, 2005.

[17] S. S. Nielsen and R. Poulsen. A two-factor, stochastic programming model of Danisch mortgage-backed securities. *J. Econ. Dynamics and Control*, 28:1267–1289, 2004.

[18] A. Ruszczyński and A. Shapiro (eds.). *Stochastic Programming*. Handbooks in Operations Research and Management Science 10, Elsevier, Amsterdam, 2003.

[19] R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. J. Wiley, New York, 1980.

[20] A. Shapiro. Statistical inference of multistage stochastic programming problems. *Math. Methods of Oper. Res.*, 58:57–68, 2003.

[21] J. F. Shapiro. Stochastic programming models for dedicated portfolio selection. In: B. Mitra (ed.) *Mathematical Models for Decision Support*. NATO ASI Series F48, Springer Verlag, Berlin, 1988, pp. 587–611.

Jitka Dupačová: Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, Praha, 186 75, Czech Republic, dupacova@karlin.mff.cuni.cz