RESEARCH ARTICLE

Uncertainties in minimax stochastic programs

Jitka Dupačová

Department of Probability and Mathematical Statistics, Charles University in Prague, Faculty of Mathematics and Physics, Sokolovská 83, Praha 186 75, Czech Republic

When using the minimax approach one tries to hedge against the worst possible distribution belonging to a specified class $\mathcal{P}$. A suitable stability analysis of results with respect to the choice of this class is an important issue. It has to be tailored to the type of the minimax problem, to the considered class of probability distributions and to the anticipated input perturbations. We shall be mainly concerned with the class of probability distributions whose supports belong to a given set and which fulfil certain moment conditions. We shall utilize results of parametric programming and of asymptotic statistics to analyze the effect of changes in input information. Among others, consistency of minimax solutions obtained for consistently estimated moments will be proved.

1. Introduction

50 years ago, stochastic programming was introduced to deal with uncertain values of coefficients which were observed in applications of linear programming. These uncertainties were modeled as random and the assumption of complete knowledge of the probability distribution P of random parameters became a standard. In practice however, complete knowledge of the probability distribution is rare. Using a hypothetical, ad hoc probability distribution P may lead to bad, costly decisions. It pays to include the existing, possibly limited information into the model. The incomplete knowledge of P is modeled by assuming that P belongs to a specified class $\mathcal{P}$ of probability distributions, the *ambiguity set*, and the *minimax approach* with respect to distributions belonging to the class $\mathcal{P}$ is applied.

The minimax approach has been developed for special types of stochastic programs and special choices of the ambiguity set $\mathcal{P}$. To illustrate the basic ideas let us consider stochastic programs of the form

$$\text{minimize } F(\boldsymbol{x}, P) := E_P f(\boldsymbol{x}, \omega) \text{ for } \boldsymbol{x} \in \mathcal{X} \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ and P is the probability distribution of the m-dimensional random vector ω. The minimax approach is applied when the probability distribution P of ω is only known to belong to a specified class $\mathcal{P}$ of probability distributions while $\mathcal{X}$ is assumed to be a fixed set, independent of P. To use the minimax approach

means to hedge against the worst possible distribution belonging to the class $\mathcal{P}$ by solving the *minimax problem*

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{P \in \mathcal{P}} F(\boldsymbol{x}, P). \qquad (2)$$

The optimal value of (2) called the (upper) *minimax bound* together with the lower bound $\min_{\boldsymbol{x} \in \mathcal{X}} \min_{P \in \mathcal{P}} F(\boldsymbol{x}, P)$ have been employed in approximation schemes used in algorithmic procedures for solving (1). See [13, 17, 25] for an introduction and for a survey of various results. Optimal solutions of (2), called *minimax solutions*, serve as a basis for decision making. They reflect the risk aversion inherent in specific applied problems. In the energy sector, for example, social and financial consequences of blackouts are hardly tolerated and it is important to know at least the character of decisions designed to hedge against the worst possible circumstances.

Minimax solutions and minimax bounds depend on specification of the class $\mathcal{P}$. Hence, we face an *additional level of uncertainty* which influences the results. Their robustness with respect to (small) changes of $\mathcal{P}$ is welcome and an output analysis is important. It has to be tailored to the type of $\mathcal{P}$, to the structure of the solved minimax problem and to the considered input perturbations, see e.g. [26, 33]. Refinement of minimax bounds by using additional information has been studied from the very beginning, e.g. [2, 9, 10]. Stability and sensitivity analysis of the minimax solutions with respect to perturbations of $\mathcal{P}$ is a more demanding task.

To construct the class $\mathcal{P}$ one often compromises between the wish to exploit existing, available information and the need to keep the minimax problem numerically tractable. One may rely on *sample information* to get sample moments or the empirical distribution, may use *experience* to get expert scenarios and some information about their probabilities, or to select a finite number of relevant probability distributions. Using both sample information and experience one can make a qualified guess about the "carrier" set which contains supports of all considered probability distributions. One also can incorporate qualitative information, like symmetry or unimodality. Compact and convex classes $\mathcal{P}$ play a key role and mathematical reasoning (& experience) may lead to a minimax problem of a manageable form, e.g. introducing bounds on probability distributions [32].

Let us mention some popular classes $\mathcal{P}$ from [13] and complete the list by introducing selected recent quotations.

- $\mathcal{P}$ consists of probability distributions supported on set $\Omega \subseteq R^m$ which fulfill certain *generalized moment conditions*, e.g.,

$$\mathcal{P} = \{P : E_P g_k(\omega) = y_k, k = 1, \dots, K\} \qquad (3)$$

  for given functions $g_1, \dots, g_K$ and prescribed values $y_k \, \forall k$. Mostly the first and second order moments appear in (3); for a brief exposition see Section 2.
- $\mathcal{P}$ defined as above with some or all equalities replaced by inequalities. An interesting idea [6] is to identify $\mathcal{P}$ by bounds on expectations (μ) and bounds on the covariance matrix, such as

$$E_P[(\omega - \mu)(\omega - \mu)^\top] \preceq \gamma \Sigma_0 \text{ for all } P \in \mathcal{P}, \qquad (4)$$

  and to apply approaches of semi-definite programming.

- P is defined as above with additional information, such as unimodality or symmetry of P taken into account [9, 11, 24, 31];
- $\mathcal{P}$ consists of probability distributions P supported on a fixed *finite* set Ω, i.e., to specify elements $P \in \mathcal{P}$ means to fix the probabilities of the considered atoms (scenarios) taking into account prior knowledge about their partial ordering [4] or their pertinence to an uncertainty set [34], etc.;
- $\mathcal{P}$ is a neighborhood of a (hypothetical, nominal or empirical) probability distribution P_0. This means that

$$\mathcal{P} := \{P \, : \, d(P, P_0) \leq \varepsilon\} \tag{5}$$

where $\varepsilon > 0$ and d is a suitable distance of probability measures. Naturally, its choice influences substantially the results. See [5] for the Kullback-Leibler distance, [23, 36] for the Kantorovich distance.
- $\mathcal{P}$ consists of a finite number of probability distributions $P_1, \ldots, P_k$ and the problem is

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{i=1,\ldots,k} F(\boldsymbol{x}, P_i); \tag{6}$$

see e.g. [33].

The listed classes are not strictly separated. For example, some moment problems lead to extremal distributions supported on a finite fixed set of scenarios which does not depend on the objective function. Thus they can be linked with the class of distributions supported on finitely many prescribed scenarios; see Example 2.2. Also the ε-neighborhood classes (5) in [5, 23] assume discrete probability distributions concentrated at finitely many a priori fixed scenarios, or at finitely many plausible scenarios to be constructed [36]. Moreover, depending on the choice of distance d in (5), $d(P, P_0) \leq \varepsilon$ can be treated as a generalized moment constraint. Moment conditions (3) are one of the ingredients for solving minimax problems under the unimodality assumption or, in general, for dealing with transformed moment problems, see Section 2.2.3.

We shall mainly deal with the class $\mathcal{P}$ of probability distributions identified by (generalized) moment conditions (3) and a given set Ω. In output analysis one may then apply results of parametric programming as done in Section 2.2 and some of asymptotic statistics; see e.g. [7, 12, 18] for early attempts in this direction. In Section 3 we shall study properties of minimax solutions based on estimated moment values and prove a consistency result. A different approach to output analysis is needed for nonparametric types of $\mathcal{P}$ such as (5) with empirical distribution function P_0 or for analysis of the sample counterpart of (6): Using asymptotic statistics, it is possible to construct nonparametric confidence sets [23] and to prove consistency of results under mild assumptions [33].

2. $\mathcal{P}$ defined by moment conditions

Theoretically the so called moment problems, e.g. [1], provide bounds for the expectation function $E_P f(\boldsymbol{x}, \omega)$ under rather general assumptions about the function $f(\boldsymbol{x}, \bullet)$ and about the considered set $\mathcal{P}$ of probability distributions on Ω defined by generalized moment conditions, such as (3). For a convex compact (in the weak topology) set $\mathcal{P}$ (continuous) expectation functions attain their maximal and minimal value at extremal points of $\mathcal{P}$. The corresponding extremal distributions

have finite supports, however, extremal distributions *independent of the form of f* appear only exceptionally.

In the case of incomplete knowledge of the probability distribution P in (1) the primal interest is in estimating the difference between the maximal and minimal expectation and in evaluation of bounds $L = \min_{\boldsymbol{x}\in\mathcal{X}} \inf_{P\in\mathcal{P}} E_P f(\boldsymbol{x},\omega)$ and $U = \min_{\boldsymbol{x}\in\mathcal{X}} \sup_{P\in\mathcal{P}} E_P f(\boldsymbol{x},\omega)$ for the optimal value of (1) that can be used in approximations. The thorough *worst case analysis* means computing minimax solutions as well.

There is a host of papers devoted to application of moment bounds in the context of stochastic programing, e.g. [3, 17], to their refinement [2] and to inclusion of qualitative information such as unimodality and/or symmetry of P by solving transformed moment problems, cf. [9–11, 24, 31].

2.1. *Basic assumptions and selected results*

Let $\mathcal{X} \subseteq \mathbb{R}^n$, $\Omega \subseteq \mathbb{R}^m$ be (Borel) measurable sets, $g_k : \Omega \to \mathbb{R}$, $k = 1\ldots, K$, and $f : \mathcal{X} \times \Omega \to \mathbb{R}$ be given measurable functions. Define $\boldsymbol{g}(\bullet) = (g_1(\bullet), \ldots, g_K(\bullet))$, denote $\mathcal{Y} := \mathrm{conv}\,\{\boldsymbol{g}(\Omega)\}$ and assume that $\mathrm{int}\mathcal{Y} \neq \emptyset$.

For $\boldsymbol{y} \in \mathcal{Y}$ let $\mathcal{P}_{\boldsymbol{y}}$ denote the class of probability distributions P of random vector ω supported on Ω such that for all $P \in \mathcal{P}_{\boldsymbol{y}}$, functions $g_k, k = 1, \ldots, K$, and $f(\boldsymbol{x}, \bullet)$ for all $\boldsymbol{x} \in \mathcal{X}$ are integrable and the moment conditions (3)

$$E_P g_k(\omega) = y_k, k = 1, \ldots, K$$

are fulfilled. The class $\mathcal{P}_{\boldsymbol{y}}$ is convex and the problem is to find

$$U(\boldsymbol{x},\boldsymbol{y}) = \sup_{P\in\mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x},\omega) \text{ and } L(\boldsymbol{x},\boldsymbol{y}) = \inf_{P\in\mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x},\omega), \qquad (7)$$

the bounds for the optimal value of (1)

$$U(\boldsymbol{y}) = \inf_{\boldsymbol{x}\in\mathcal{X}} \sup_{P\in\mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x},\omega) \text{ and } L(\boldsymbol{y}) = \inf_{\boldsymbol{x}\in\mathcal{X}} \inf_{P\in\mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x},\omega) \qquad (8)$$

and to compute the minimax solutions, elements of $\mathcal{X}^*(\boldsymbol{y}) := \arg\min_{\boldsymbol{x}\in\mathcal{X}} U(\boldsymbol{x},\boldsymbol{y})$.

An *important case* is when Ω is compact, $g_1, \ldots, g_K$ continuous with $\mathrm{int}\mathcal{Y} \neq \emptyset$ and $f(\boldsymbol{x}, \bullet)$ upper semicontinuous. Then the class $\mathcal{P}_{\boldsymbol{y}}$, for $\boldsymbol{y} \in \mathcal{Y}$, is nonempty, compact (in weak topology) and the supremum in (7) is achieved; see e.g. [19]. It is expedient to analyze the dual program to the upper bound in (7) which reads

$$\text{minimize } d_0 + \sum_{k=1}^{K} d_k y_k$$

subject to

$$\boldsymbol{d} \in \mathcal{D} := \{\boldsymbol{d} \in \mathbb{R}^{K+1} \,:\, d_0 + \sum_{k=1}^{K} d_k g_k(\boldsymbol{z}) \geq f(\boldsymbol{x}, \boldsymbol{z}) \,\forall \boldsymbol{z} \in \Omega\}.$$

If $\boldsymbol{y} \in \mathrm{int}\mathcal{Y}$, there is no duality gap between problem (7) and its dual problem, and the dual problem has an optimal solution; see e.g. [20]. It means that there exists

a probability distribution $P^* \in \mathcal{P}_{\boldsymbol{y}}$ and a vector $\boldsymbol{d}^* \in \mathcal{D}$ such that

$$U(\boldsymbol{x}, \boldsymbol{y}) = E_{P^*} f(\boldsymbol{x}, \omega) = d_0^* + \sum_{k=1}^{K} d_k^* y_k. \tag{9}$$

Hence, for the given $\boldsymbol{y} \in \text{int} \mathcal{Y}$, the bound can be obtained as the optimal value of the corresponding semiinfinite program

$$U(\boldsymbol{x}, \boldsymbol{y}) = \inf_{\boldsymbol{d}} \{ d_0 + \sum_{k=1}^{K} d_k y_k \ : \ d_0 + \sum_{k=1}^{K} d_k g_k(\boldsymbol{z}) \geq f(\boldsymbol{x}, \boldsymbol{z}) \, \forall \, \boldsymbol{z} \in \Omega \}. \tag{10}$$

Evidently, as a function of the parameter $\boldsymbol{y}$, $U(\boldsymbol{x}, \boldsymbol{y})$ is concave.

From the point of view of computation it is important that P^* in (9) is in fact a *discrete* distribution. For these and other related results see e.g. [1, 17, 20, 25]. Similar statements hold true also for the case of inequality constraints in (3), see [3, 25].

Under *additional assumptions,* e.g. for Ω a bounded polyhedron and

$$h(\boldsymbol{x}, \boldsymbol{z}) := f(\boldsymbol{x}, \boldsymbol{z}) - \sum_{k=1}^{K} d_k g_k(\boldsymbol{z})$$

a piecewise convex or quasi-convex function on Ω for an arbitrary $\boldsymbol{x} \in \mathcal{X}$, (10) can be reduced to a *finite-dimensional* linear program; cf. Theorem 2 of [8] or Theorem 8.1.1 of [25]. For example, if $h(\boldsymbol{x}, \boldsymbol{z})$ is convex in $\boldsymbol{z}$ and $\Omega = \text{conv} \{ \boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(H)} \}$, we get the set of feasible solutions of (10) in the form

$$\mathcal{D} = \{ \boldsymbol{d} \in \mathbb{R}^{K+1} \ : \ d_0 + \sum_{k=1}^{K} d_k g_k(\boldsymbol{z}^{(h)}) \geq f(\boldsymbol{x}, \boldsymbol{z}^{(h)}), h = 1, \ldots, H \}. \tag{11}$$

2.1.1. *Special convex case*

Assumption 2.1 Assume that for all $\boldsymbol{x} \in \mathcal{X}$, $f(\boldsymbol{x}, \bullet)$ is convex, $g_k(\boldsymbol{z}) = z_k, k = 1, \ldots, m$, Ω is a bounded polyhedron $\text{conv} \{ \boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(H)} \}$ and $\boldsymbol{y} \in \text{int} \mathcal{Y} (= \text{int} \Omega)$.

This is the favorite class $\mathcal{P}_{\boldsymbol{y}}$ defined by prescribed values of the first moments. The upper bound $U(\boldsymbol{x}, \boldsymbol{y})$ reduces to the Edmundson–Madansky bound [22] and $U(\boldsymbol{x}, \boldsymbol{y})$ is the optimal value of the linear program dual to (10):

$$U(\boldsymbol{x}, \boldsymbol{y}) = \max_{\boldsymbol{p}} \{ \sum_{h=1}^{H} p_h f(\boldsymbol{x}, \boldsymbol{z}^{(h)}) \ : \ \sum_{h=1}^{H} p_h \boldsymbol{z}^{(h)} = \boldsymbol{y}, \sum_{h=1}^{H} p_h = 1, p_h \geq 0 \, \forall h \}. \tag{12}$$

According to Jensen's inequality $L(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{x}, \boldsymbol{y})$. For an extension to piecewise linear functions g_k and an unbounded convex closed set Ω see [3].

In a similar way, it is possible to formulate the moment problem for probability distributions supported on a *known finite set of scenarios*, i.e. for $\Omega = \{ \omega^1, \ldots, \omega^I \}$. The probability distributions are then fully identified by probabilities $p_i, i = 1, \ldots, I$, of these scenarios and by moment conditions. The problem to

solve is

$$U(\boldsymbol{x}, \boldsymbol{y}) = \max_{\boldsymbol{p}} \{\sum_{i=1}^{I} p_i f(\boldsymbol{x}, \omega^i) \ : \ \sum_{i=1}^{I} p_i g_k(\omega^i) = y_k \, \forall k, \ \sum_{i=1}^{I} p_i = 1, \ p_i \geq 0 \, \forall i\}. \quad (13)$$

This situation may occur for problems with an ad hoc given finite support of ω or may result from identification of discrete extremal distributions as in (11) or (12), may be obtained by sampling or relying on past data and on experts' suggestions. Additional polyhedral constraints on probabilities p_i, cf. [4, 34], can be included.

Further simplifications are possible when $f(\boldsymbol{x}, \omega)$ is convex separable with respect to individual components of ω, or Ω is a rectangle and the components of ω are independent, or when Ω is a simplex. Then we may even get explicit formulas for $U(\boldsymbol{x}, \boldsymbol{y})$, cf. [9, 10], or obtain extremal probability distributions which do not depend on the choice of the convex random objective function $f(\bullet, \omega)$.

Example 2.2 Assume that $f(\boldsymbol{x}, \boldsymbol{z}) = \sum_{j=1}^{m} f_j(\boldsymbol{x}, z_j)$ where for a fixed $\boldsymbol{x}$, $f_j \, \forall j$ are convex functions of z_j and that $\mathcal{P}_{\boldsymbol{y}}$ is defined by the following conditions: The marginal distributions of ω_j are supported on given nondegenerate compact intervals $[a_j, b_j]$ with Ω their Cartesian product, $E_P \omega_j = y_j$, with given values $y_j \in (a_j, b_j) \, \forall j$.

Then

$$U(\boldsymbol{x}, \boldsymbol{y}) = \max_{P \in \mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x}, \omega) = \sum_{j=1}^{m} \lambda_j f_j(\boldsymbol{x}, a_j) + \sum_{j=1}^{m} (1 - \lambda_j) f_j(\boldsymbol{x}, b_j) \quad (14)$$

with $\lambda_j = (b_j - y_j)/(b_j - a_j) \, \forall j$.

Moreover, for arbitrary values $y_j \in (a_j, b_j), j = 1, \ldots, m$, the minimax solution is an *efficient solution of a multicriteria problem* with objective functions $f(\boldsymbol{x}, \boldsymbol{z})$, $\boldsymbol{z} \in \mathcal{Z}^*$ where $\mathcal{Z}^*$ is the set of vertices of Ω, cf. [11]. *Specifying the expectation $\boldsymbol{y}$ we get one of these efficient solutions.*

Example 2.3 Consider Example 2.2 with $m = n = 1$, and $f(x, z) = (x - z)^+$. Such terms (with $\boldsymbol{z}$ replaced by ω) appear as the random part of objective functions of various popular stochastic programs, for example of the simple recourse problem whose well known classroom example is the newsboy problem, or of the risk measure CVaR — Conditional Value at Risk often used in financial mathematics; see e.g. [25]. Let us modify conditions identifying the class of probability distributions and consider $\mathcal{P} = \mathcal{P}_{\mu, V}$ defined as follows:

There is a known upper bound V for the range of variation of ω and a prescribed expectation $E_P \omega = \mu, \forall P \in \mathcal{P}_{\mu, V}$.

Then we evidently have $\omega \in [\mu - V, \mu + V] \, \forall P \in \mathcal{P}_{\mu, V}$ with probability 1 and

$$
\begin{aligned}
U(x, \mu, V) := \max_{P \in \mathcal{P}_{\mu, V}} \ E_P(x - \omega)^+ &= 0 \quad \text{for } x < \mu - V \\
&= x - \mu \quad \text{for } x > \mu + V \\
&= \tfrac{1}{4V}(V + x - \mu)^2 \quad \text{for } \mu - V \leq x \leq \mu + V.
\end{aligned}
\quad (15)
$$

The last formula in (15) follows by application of (14): For an arbitrary $x \in [\mu - V, \mu + V]$ there exists $P \in \mathcal{P}_{\mu, V}$ such that x belongs to its support. We shall prove first that $\tilde{U}(x) := \frac{1}{4V}(V + x - \mu)^2 \geq E_P(x - \omega)^+$ for $P \in \mathcal{P}_{\mu, V}$ and $\mu - V \leq x \leq \mu + V$.

The support of an arbitrary $P \in \mathcal{P}_{\mu,V}$ can be enlarged to $[a, b]$ so that $\mu \in [a, b]$ and $b - a = V$. If $x \in [a, a + V]$ then according to (14)

$$E_P(x - \omega)^+ \leq (x - a)\frac{V + a - \mu}{V},$$

$\max_a\{(x - a)\frac{V+a-\mu}{V} : a \in [\mu - V, \mu]\}$ is attained for $a^* = \frac{1}{2}(x + \mu - V)$ and the maximal value equals $\tilde{U}(x)$. If $\mu - V \leq x < a$, $E_P(x - \omega)^+ = 0 \leq \tilde{U}(x)$. Finally, by a simple algebra we get for $b = a + V < x \leq \mu + V$ the sought inequality $E_P(x - \omega)^+ = x - \mu \leq \tilde{U}(x)$. Hence,

$$\tilde{U}(x) = \frac{1}{4V}(V + x - \mu)^2 \geq \sup_{P \in \mathcal{P}_{\mu,V}} E_P(x - \omega)^+ \text{ for } x \in [\mu - V, \mu + V].$$

The supremum is attained for the extremal discrete distribution $P^* \in \mathcal{P}_{\mu,V}$ which is supported by a^* and $a^* + V$ with probabilities $p^* = \frac{V - \mu + x}{2V}$ and $1 - p^*$.

Without compactness of Ω, existence of the optimal solution in the inner optimization problem (10) depends much on the properties of the functions $f(\boldsymbol{x}, \bullet)$ and $g_k \forall k$. An example is $\Omega = \mathbb{R}^m$, $f(\boldsymbol{x}, \bullet)$ positive conical (i.e. $f(\boldsymbol{x}, \boldsymbol{z}) > 0 \forall \boldsymbol{z} \neq 0$ and epi$f(\boldsymbol{x}, \bullet)$ a convex polyhedral cone in $\mathbb{R}^{m+1}$ pointed at the origin) with moment constraints on expectations and on the second order moments; see [17, 25]. A special instance (cf. [9, 10]) is obtained by a direct solution of (10):

Example 2.4 Assume again that $f(x, z) = (x - z)^+$ and define $\mathcal{P} := \mathcal{P}_{\mu,\sigma^2}$ by moment conditions $E_P\omega = \mu$, $\text{var}_P\omega = \sigma^2$ for all $P \in \mathcal{P}_{\mu,\sigma^2}$.

Then

$$\max_{P \in \mathcal{P}_{\mu,\sigma^2}} E_P(x - \omega)^+ = \frac{1}{2}\left(x - \mu + \sqrt{\sigma^2 + (x - \mu)^2}\right) := U(x, \mu, \sigma^2). \qquad (16)$$

For a generalization of Example 2.4 to piecewise linear convex functions $f(\boldsymbol{x}, \bullet)$ and/or g_k see [17].

Convexity properties of $f(\boldsymbol{x}, \bullet)$ play an essential role. For example, in the two-stage stochastic linear program

$$f(\boldsymbol{x}, \omega) = \boldsymbol{c}^\top \boldsymbol{x} + \min_u\{\boldsymbol{q}(\omega)^\top \boldsymbol{u} : \boldsymbol{W}(\omega)\boldsymbol{u} = \boldsymbol{h}(\omega) - \boldsymbol{T}(\omega)\boldsymbol{x}, \boldsymbol{u} \geq 0\}.$$

Convexity of $f(\boldsymbol{x}, \bullet)$ is achieved by a restriction to a *fixed* recourse matrix $\boldsymbol{W}$, *fixed* coefficients $\boldsymbol{q}$ in the second-stage objective function and to right-hand sides $\boldsymbol{h}$ and technological matrix $\boldsymbol{T}$ *linear* in ω. There are also parallel results for saddle functions $f(\boldsymbol{x}, \bullet)$ that allow inclusion of random coefficients $\boldsymbol{q}$, e.g. [14].

2.2.   *Stability with respect to input information*

The prescribed values of moments used in definition of $\mathcal{P}_{\boldsymbol{y}}$ play the role of *input information* which influences the resulting minimax bounds and minimax solutions. However, this input information is not always completely known, it can be based on sample or past information, on expert opinion, etc. In the sequel, we shall deal with stability of minimax bounds and minimax decisions under rather simplifying assumptions postponing possible generalizations.

Assumption 2.5 Assume that

- $\mathcal{X} \subset \mathbb{R}^n$ is a nonempty convex compact set,
- $\Omega \subset \mathbb{R}^m$ is a nonempty compact set,
- $g_1, \ldots, g_K$ are given continuous functions on Ω,
- $f : \mathcal{X} \times \Omega \to \mathbb{R}$ is a continuous function on Ω for any fixed $\boldsymbol{x} \in \mathcal{X}$ and it is a lower semicontinuous convex function of $\boldsymbol{x}$ for every $\omega \in \Omega$,
- the interior of the moment space $\mathcal{Y} := \mathrm{conv}\,\{\boldsymbol{g}(\Omega)\}$ is nonempty and $\boldsymbol{y} \in \mathrm{int}\mathcal{Y}$.

Remark 1 : Observe that lower semicontinuity and convexity of $f(\bullet, \omega)$ implies lower semicontinuity and convexity of the expected value functions $E_P f(\bullet, \omega)$ (see [35]), as well as of the upper bounds $U(\bullet, \boldsymbol{y}) = \sup_{P \in \mathcal{P}_{\boldsymbol{y}}} E_P f(\bullet, \omega)$.

In this section, we shall present selected applications of parametric programming to stability and sensitivity of moment bounds (7)–(8) and of minimax decisions with respect to the prescribed values of moments and/or to the choice of set Ω.

2.2.1. *Prescribed moment values*

Being concave with respect to $\boldsymbol{y} \in \mathcal{Y}$ the optimal value function $U(\boldsymbol{x}, \bullet)$ of the semiinfinite program (10) has *directional derivatives* on $\mathrm{int}\mathcal{Y}$ in all directions, see e.g. [28]. For the special problems (12) and (13), stability analysis with respect to $\boldsymbol{y}$ reduces to the standard stability analysis for linear programs with respect to right-hand sides and the optimal value function $U(\boldsymbol{x}, \bullet)$ is concave, piecewise linear on $\mathrm{int}\mathcal{Y}$.

Concerning the optimal value $U(\boldsymbol{y}) = \min_{\boldsymbol{x} \in \mathcal{X}} U(\boldsymbol{x}, \boldsymbol{y})$ and the minimax solutions, one can apply results on stability for nonlinear parametric programs as in [11] to obtain:

Theorem 2.6 : *Under Assumption 2.5, $U(\boldsymbol{y}) := \min_{\boldsymbol{x} \in \mathcal{X}} U(\boldsymbol{x}, \boldsymbol{y})$ is concave on $\mathcal{Y}$ and the point-to-set mapping $\boldsymbol{y} \mapsto \mathcal{X}^*(\boldsymbol{y})$ is upper semicontinuous on $\mathcal{Y}$.*

This implies again that $U(\boldsymbol{y})$ has *directional derivatives* on $\mathrm{int}\mathcal{Y}$ in all directions. Gradients of $U(\boldsymbol{y})$ exist almost everywhere on $\mathrm{int}\mathcal{Y}$, nevertheless, *differentiability* of $U(\boldsymbol{y})$ holds true only under additional smoothness assumptions, e.g. second order differentiability of $U(\boldsymbol{x}, \boldsymbol{y})$, second order sufficient conditions, and suitable regularity conditions, see e.g. [16]. Under such conditions, there is a unique minimax solution, say $\boldsymbol{x}(\boldsymbol{y})$, and $\nabla U(\boldsymbol{y}) = \nabla_{\boldsymbol{y}} U(\boldsymbol{x}(\boldsymbol{y}), \boldsymbol{y})$. However, the assumption of second order differentiability of $U(\boldsymbol{x}, \boldsymbol{y})$ is not always realistic. (For an example when it is fulfilled see [7] and Example 2.4.) Therefore, in *postoptimality analysis* of the moment bounds one can rely mainly on the results on directional differentiability, cf. [12].

2.2.2. *Choice of the set Ω*

The direct analysis of explicit formulas such as (14) shows that due to a change of Ω the upper bound function $U(\boldsymbol{x}, \boldsymbol{y})$ may change significantly; see also Theorem 3.1.1 of [1]. The relaxation of the assumption of a known set Ω as done in Example 2.3 leads to a rather different upper moment bound as well.

The situation is relatively simple in the special case of probability distributions supported on a given finite set of scenarios but with not precisely known probabilities of their occurrence, cf. (13), or for the class $\mathcal{P}_{\boldsymbol{y}}$ determined by prescribed expectations $\boldsymbol{y}$ with a convex polyhedral set Ω and convex $f(\boldsymbol{x}, \bullet)$, cf. (12) in Section 2.1.1. For each fixed $\boldsymbol{x} \in \mathcal{X}$ the worst case probabilities can be then obtained as optimal solutions of a linear program with a compact set of feasible solutions. Changes of scenarios or of the extremal points of Ω influence the objective function and the matrix of coefficients of the linear program. Nevertheless, these linear programs are *stable* (small changes of the data cause only small changes of

the optimal solutions, cf. [27]) under the nonemptiness and boundedness condition on the sets of optimal solutions of the corresponding dual programs

$$\inf_{\boldsymbol{d}\in\mathcal{D}} d_0 + \sum_{k=1}^{K} d_k y_k \tag{17}$$

with $\mathcal{D}$ given by

$$\mathcal{D} = \{\boldsymbol{d} \in \mathrm{I\!R}^{K+1} \,:\, d_0 + \sum_{k=1}^{K} d_k g_k(\omega^i) \geq f(\boldsymbol{x}, \omega^i), i = 1, \ldots, I\}$$

or

$$\mathcal{D} = \{\boldsymbol{d} \in \mathrm{I\!R}^{K+1} \,:\, d_0 + \sum_{k=1}^{K} d_k z_k^{(h)} \geq f(\boldsymbol{x}, \boldsymbol{z}^{(h)}), h = 1, \ldots, H\}. \tag{18}$$

For $\boldsymbol{y} \in \mathrm{int}\mathcal{Y}$ this condition is fulfilled, cf. [20]. The optimal value function U is then a continuous function of all coefficients on a neighborhood of the initial data $\boldsymbol{x}, \boldsymbol{y}$ and $\boldsymbol{z}^{(h)}, 1 \leq h \leq H$, or $\boldsymbol{g}(\omega^i) \forall i$. A unique and nondegenerate optimal solution of the primal LP (12) or (13) is a special well known example. The size of the neighborhood is limited e.g. by the condition that the perturbed vector $\boldsymbol{y}$ remains an interior point of the convex hull of the perturbed moment space.

Another possibility is to analyze the dual linear programs (17)–(18) allowing some uncertainty in selection of ω^i or $\boldsymbol{z}^{(h)}$. Inspired by [15] consider problem (12) with extremal points $\boldsymbol{z}^h$ which belong to an ellipsoid around $\boldsymbol{z}^{(h)}$, say

$$\boldsymbol{z}^h = \boldsymbol{z}^{(h)} + \boldsymbol{E}_h \delta^h, \quad \|\delta^h\|_2 \leq \varrho, \tag{19}$$

and ask for the best solution of program (17)–(18) which is feasible for all choices of $\boldsymbol{z}$ obtained by the special structure of perturbations (19). In the simplest case of $\boldsymbol{E}_h = \boldsymbol{I}$ the h-th constraint of (18) is fulfilled if

$$d_0 + \boldsymbol{d}^\top \boldsymbol{z}^{(h)} + \boldsymbol{d}^\top \delta^h - f(\boldsymbol{x}, \boldsymbol{z}^{(h)} + \delta^h) \geq 0 \quad \forall \|\delta^h\|_2 \leq \varrho. \tag{20}$$

The Lipschitz property of $f(\boldsymbol{x}, \bullet)$ on the neighborhood (19) implies that there is a constant l such that

$$|f(\boldsymbol{x}, \boldsymbol{z}^h) - f(\boldsymbol{x}, \boldsymbol{z}^{(h)})| \leq l\|\delta^h\|_2 \leq l\varrho.$$

By an adaptation of results in section 5.3 of [15], to satisfy constraint (20) it is sufficient that

$$d_0 + \boldsymbol{d}^\top \boldsymbol{z}^{(h)} - f(\boldsymbol{x}, \boldsymbol{z}^{(h)}) - \varrho\sqrt{\|\boldsymbol{d}\|_2^2 + l^2} \geq 0. \tag{21}$$

Again, when the optimal solution of the unperturbed linear program (17)–(18) is unique and nondegenerate, then there exists $\varrho_{\max} > 0$ such that for all problems with perturbed constraints (21) with $0 < \varrho < \varrho_{\max}$ the optimal solutions are unique and nondegenerate, too. A similar analysis applies to (13) under suitable assumptions about the mapping $\boldsymbol{g}$.

Even for classes $\mathcal{P}$ which do not explicitly assume a fixed known carrier set Ω, such as (4) or (5), various assumptions about Ω are exploited in output analysis, e.g. the set Ω is supposed to be compact convex. As a special case, existence of a ball of radius R that contains the support of the unknown probability distribution is frequently assumed where the magnitude of R may follow from "an educated and conservative guess"; cf. [6, 23].

Convergence properties have been studied for finite supports which are consecutively improved to approximate the uncountable support; cf. [26] and the next example.

Example 2.7 Assume that $\mathcal{P}_{\boldsymbol{y}}$ is the class of probability distributions on $\Omega \subset \mathbb{R}^m$, which fulfill the moment conditions (3) and Assumption 2.5 is satisfied.

Let $\{\Omega^\nu\}_{\nu \geq 1}$ be a sequence of finite sets in $\mathbb{R}^m$ such that $\Omega^\nu \subseteq \Omega^{\nu+1} \subseteq \Omega$ and $\bigcup_\nu \Omega^\nu$ is dense in Ω. Choose ν_0 such that $\boldsymbol{y} \in \mathrm{int}\,\mathrm{conv}\{\boldsymbol{g}(\Omega^{\nu_0})\}$. For $\nu \geq \nu_0$ consider classes $\mathcal{P}_{\boldsymbol{y}}^\nu$ of probability distributions supported on Ω^ν for which moment conditions (3) are fulfilled. The following statement can be viewed as a special case of Proposition 2.1 and Example 2.1 of [26]:

If for every $P \in \mathcal{P}_{\boldsymbol{y}}$ there is a sequence $\{P^\nu\}_{\nu \geq \nu_0}$, $P^\nu \in \mathcal{P}_{\boldsymbol{y}}^\nu$ which for $\nu \to \infty$ converges weakly to P, then for $\nu \to \infty$

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_{\boldsymbol{y}}^\nu} E_P f(\boldsymbol{x}, \omega) \to \min_{\boldsymbol{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_{\boldsymbol{y}}} E_P f(\boldsymbol{x}, \omega).$$

Indeed, for all $P \in \mathcal{P}_{\boldsymbol{y}}$, lower semicontinuity and convexity of the expected value $E_P f(\boldsymbol{x}, \omega)$ follows from the assumed lower semicontinuity and convexity of $f(\bullet, \omega)$. Joint continuity of $E_P f(\boldsymbol{x}, \omega)$ with respect to $\boldsymbol{x}$ and P follows from its convexity with respect to $\boldsymbol{x}$ and continuity with respect to P (apply Theorem 10.7 of [28] under Assumption 2.5 and Remark 1).

Moreover, also upper semicontinuity of sets of minimax solutions with respect to the considered convergence of classes $\mathcal{P}_{\boldsymbol{y}}^\nu$ to $\mathcal{P}_{\boldsymbol{y}}$ can be proved.

2.2.3. *Additional input information*

If the class of probability distributions is defined not only by the set Ω and the moment conditions (3) but also by other conditions such as unimodality then it is often possible to remove these conditions by a suitable transformation of probability distributions and functions and to reduce the problem to the basic moment problem. With reference to [24, 31] for more general situations, we shall delineate here only the approach for univariate unimodal probability distributions which was motivated by [21] and detailed in [9–11].

Let $\mathcal{P}_y^M$ be the class of unimodal probability distributions on a compact interval $\Omega \subset \mathbb{R}$ with the given mode M such that the moment conditions (3) are fulfilled. All distributions in $\mathcal{P}_y^M$ are mixtures of uniform distributions over intervals (u, M) and (M, u'), $u, u' \in \Omega$ complemented by the degenerated distribution concentrated at M.

For a continuous function $h : \Omega \to \mathbb{R}$ consider its transform h^* defined as follows:

$$h^*(z) = \frac{1}{z - M} \int_M^z h(u)du \text{ for } z \in \Omega, \, z \neq M \text{ and } h^*(z) = h(z) \text{ for } z = M. \quad (22)$$

According to [21] and in view of Assumption 2.5, the moment bound for unimodal

distributions can be obtained as follows:

$$\max_{P \in \mathcal{P}_y^M} E_P f(\boldsymbol{x}, \omega) = \max_P \{E_P f^*(\boldsymbol{x}, \omega) \, : \, E_P g_k^*(\omega) = y_j, \, k = 1, \ldots, K\}. \qquad (23)$$

Moreover, the transform h^* of a convex function h is convex. The next simple example details a possible application.

Example 2.8 As in Example 2.2 with $m = 1$ let $f(\boldsymbol{x}, z)$ be a convex function of z. Let $\mathcal{P}_y^M$ be the class of unimodal probability distributions with the given mode M, supported on a compact interval $[a, b]$ and such that for all $P \in \mathcal{P}_y^M$, $E_P \omega = y$. For $g(u) = u$ the transform (22) gives $g^*(z) = \frac{1}{2}(z + M)$ and equation (23) then reads

$$\max_{P \in \mathcal{P}_y^M} E_P f(\boldsymbol{x}, \omega) =$$

$$\max_P \{E_P f^*(\boldsymbol{x}, \omega) : E_P g^*(\omega) = y, P(\omega \in [a, b]) = 1\} := \tilde{U}(\boldsymbol{x}, y, M). \qquad (24)$$

Define $\mu = 2y - M$; then (24) is nothing else but the usual moment problem with the class $\mathcal{P}_\mu = \{P \, : \, E_P \omega = \mu, \, P(\omega \in [a, b]) = 1\}$; moreover, if the expectation y and the mode M coincide, then $\mu = y$. The transformed objective $f^*(\boldsymbol{x}, z)$ is convex in z. This means that the maximal expectation $E_P f^*(\boldsymbol{x}, \omega)$ over the class $\mathcal{P}_\mu$ is

$$U(\boldsymbol{x}, \mu) = \lambda f^*(\boldsymbol{x}, a) + (1 - \lambda) f^*(\boldsymbol{x}, b) = \tilde{U}(\boldsymbol{x}, y, M)$$

with $\lambda = \frac{b - 2y + M}{b - a} = \frac{b - \mu}{b - a}$. Substituting for $f^*(\boldsymbol{x}, z)$ according to (22) we get

$$\tilde{U}(\boldsymbol{x}, y, M) = \frac{b - 2y + M}{(b - a)(M - a)} \int_a^M f(\boldsymbol{x}, u) du + \frac{2y - M - a}{(b - a)(b - M)} \int_M^b f(\boldsymbol{x}, u) du.$$

It is easy to recognize two densities of uniform distributions on $[a, M]$ and $[M, b]$ weighted by λ and $(1 - \lambda)$, respectively.

If the mode is not known, additional maximization with respect to $M \in [a, b]$ is possible. As a result, the worst case probability distribution is uniform on $[a, b]$ if $y = \frac{1}{2}(a + b)$ or is a mixture of the uniform distribution over $[a, b]$ and the degenerated one concentrated at a or at b if $y > \frac{1}{2}(a + b)$ or $y < \frac{1}{2}(a + b)$, respectively.

3.   Stability with respect to estimated moment values

Assume now that sample information was used to estimate the moment values which identify the class $\mathcal{P}_{\boldsymbol{y}}$. Assume that these parameters were consistently estimated using e.g. a sequence of i.i.d. observations of ω. Let $\boldsymbol{y}^\nu$ be based on the first ν observations. Using continuity of function $U(\boldsymbol{x}, \bullet)$ on int$\mathcal{Y}$ and theorems about convergence of transformed random variables, cf. [30], we get for (strongly) consistent estimates $\boldsymbol{y}^\nu$ of the true parameter $\boldsymbol{y}$ the pointwise convergence

$$u^\nu(\boldsymbol{x}) := U(\boldsymbol{x}, \boldsymbol{y}^\nu) \to U(\boldsymbol{x}, \boldsymbol{y}) \quad \text{a.s.} \qquad (25)$$

valid at an arbitrary element $\boldsymbol{x} \in \mathcal{X}$. The same conclusion, based on Theorem 2.6, holds true for minimax bounds:

Theorem 3.1: *Under Assumption 2.5 and for strongly consistent estimates $\boldsymbol{y}^\nu$ of $\boldsymbol{y}$,*

$$\min_{x \in \mathcal{X}} U(\boldsymbol{x}, \boldsymbol{y}^\nu) \to \min_{x \in \mathcal{X}} U(\boldsymbol{x}, \boldsymbol{y}) \quad \text{a.s.}$$

In general the pointwise convergence (25) does not imply consistency of minimax solutions. To get the consistency result one may employ epi-convergence of the approximate objectives $u^\nu(\boldsymbol{x}) := U(\boldsymbol{x}, \boldsymbol{y}^\nu)$; for epi-convergence consult e.g. Chapter 7 of [29].

Definition 3.2: A sequence of functions $\{u^\nu : \mathrm{I\!R}^n \to \bar{\mathrm{I\!R}}, \nu = 1, \dots \}$ is said to epi-converge to $u : \mathrm{I\!R}^n \to \bar{\mathrm{I\!R}}$ if for all $\boldsymbol{x} \in \mathrm{I\!R}^n$ the two following properties hold:

$$\liminf_{\nu \to \infty} u^\nu(\boldsymbol{x}^\nu) \geq u(\boldsymbol{x}) \text{ for all } \{\boldsymbol{x}^\nu\} \to \boldsymbol{x} \tag{26}$$

and for some $\{\boldsymbol{x}^\nu\}$ converging to $\boldsymbol{x}$

$$\limsup_{\nu \to \infty} u^\nu(\boldsymbol{x}^\nu) \leq u(\boldsymbol{x}). \tag{27}$$

Whereas the pointwise convergence implies condition (27), additional assumptions are needed to get validity of condition (26). Fortunately, *pointwise convergence of lower semicontinuous convex functions u^ν, u with $\mathrm{int}(\mathrm{dom}\, u) \neq \emptyset$ implies epi-convergence;* see e.g. Theorem 7.17 of [29]. In such case, $\limsup\{\arg\min u^\nu\} \subset \arg\min u$.

Theorem 3.3: *Under Assumption 2.5 and for strongly consistent estimates $\boldsymbol{y}^\nu$ of $\boldsymbol{y}$, the approximate objectives $u^\nu(\boldsymbol{x})$ epi-converge almost surely to $U(\boldsymbol{x}, \boldsymbol{y})$ as $\nu \to \infty$. This implies that with probability 1 all cluster points of sequence of minimizers $\boldsymbol{x}^\nu$ of $u^\nu(\boldsymbol{x})$ on $\mathcal{X}$ are minimizers of $U(\boldsymbol{x}, \boldsymbol{y})$ on $\mathcal{X}$.*

Proof: Let us examine epi-convergence of the sequence of $u^\nu(\boldsymbol{x}) := U(\boldsymbol{x}, \boldsymbol{y}^\nu), \nu = 1, \dots$, to $u(\boldsymbol{x}) := U(\boldsymbol{x}, \boldsymbol{y})$. According to Remark 1, $u^\nu(\boldsymbol{x}), u(\boldsymbol{x})$ are lower semicontinuous convex functions on $\mathcal{X}$ and for each $\boldsymbol{x} \in \mathcal{X}$, $u^\nu(\boldsymbol{x}) \to u(\boldsymbol{x})$ a.s. according to (25). It implies that there is a countable set $\mathcal{C}$ dense in $\mathcal{X}$ and a probability 1 set $\mathcal{Y}_0$ in the sample space such that $u^\nu(\boldsymbol{x}) \to u(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{C}$ and for all sample paths $\{\boldsymbol{y}^\nu\}_1^\infty \in \mathcal{Y}_0$. Hence, see Theorem 10.8 of [28], with probability 1, $u^\nu(\boldsymbol{x}) \to u(\boldsymbol{x})$ uniformly on any compact set $\mathcal{K} \subset \mathrm{int}\mathcal{X}$ which is equivalent to epi-convergence of $\{u^\nu\}$ to u; cf. Theorem 7.17 of [29]. The sought consistency of the corresponding minimax solutions $\boldsymbol{x}^\nu$ follows by an application of Theorem 7.33 of [29].   □

A similar consistency result can be obtained also for the special convex problem treated in Section 2.2.2 with perturbed both $\boldsymbol{y}$ and Ω.

Example 3.4 Assume that parameters a, b, μ identifying the class of one-dimensional probability distributions on the interval $[a, b]$ with mean value μ are known to belong to the interior of a compact set in $\mathrm{I\!R}^3$. Assume further that their values can be obtained by an estimation procedure based on a sample path of i.i.d. observations of ω from the true probability distribution P. Their consistent estimates based on a sample size ν are the minimal/maximal sample values and the arithmetic mean, i.e. $\omega_{\nu:1}, \omega_{\nu:\nu}$ and $\mu^\nu = \frac{1}{\nu} \sum_{i=1}^\nu \omega_i$. In this case, we know the explicit form of all approximate objective functions

$$u^\nu(\boldsymbol{x}) := \lambda^\nu f(\boldsymbol{x}, \omega_{\nu:1}) + (1 - \lambda^\nu) f(\boldsymbol{x}, \omega_{\nu:\nu})$$

with $\lambda^\nu = (\omega_{\nu:\nu} - \mu^\nu)/(\omega_{\nu:\nu} - \omega_{\nu:1})$; see Example 2.2 for $m = 1$. This is a continuous function of parameters provided that $\omega_{\nu:1} < \omega_{\nu:\nu}$. For $f(\bullet, \omega)$ convex, $u^\nu(\boldsymbol{x})$ are convex in $\boldsymbol{x}$. For compact set $\mathcal{X}$, existence of the true minimax solution $\boldsymbol{x}$ follows from lower semicontinuity of $f(\bullet, a)$ and $f(\bullet, b)$. Hence, the consistency statements of Theorems 3.1 and 3.3 hold true. An extension to the corresponding "data-driven" version of Example 2.2 is obvious.

Additional asymptotic results can be proved when the function $U(\boldsymbol{x}, \bullet)$ is differentiable. This is the case of the "data-driven" version of Example 2.4.

Example 3.5 Assume that σ is known and μ is replaced by arithmetic mean $\mu^\nu = \frac{1}{\nu} \sum_i \omega^i$ of i.i.d. realizations of ω, which is an asymptotically normal estimate.

Using differentiability of the function $U(x, \mu, \sigma^2)$ in (16) with respect to x, μ and theorems about distributions of transformed sequences of random variables, here of $U(x, \mu^\nu, \sigma^2)$, we get an asymptotically normal distribution of function values of the approximate minimax objectives. Second order differentiability of function $U(x, \mu, \sigma^2)$ is used to obtain the rate of convergence $O(\nu^{-1/2})$ in probability, based on the Berry-Esséen inequality; see [7].

Additional assumptions are needed to prove asymptotic normality of the approximate minimax bounds and minimax decisions.

4.   Extensions

Up to now we have assumed that the set of feasible solutions does not depend on the probability distribution P. Let us remove this assumption and consider the stochastic program

$$\text{minimize } F(\boldsymbol{x}, P) := E_P f(\boldsymbol{x}, \omega) \text{ on the set } \mathcal{X} \cap \mathcal{X}(P) \tag{28}$$

where $\mathcal{X}$ does not depend on P and $\mathcal{X}(P) = \{\boldsymbol{x} \in \mathcal{X} : G_i(\boldsymbol{x}, P) \leq 0, i = 1, \ldots, k\}$. Stochastic programs with probabilistic constraints are a special type of (28). Theoretically, it is enough to deal just with one constraint $G(\boldsymbol{x}, P) \leq 0$.

When the probability distribution P of ω in (28) is only known to belong to a specified class $\mathcal{P}$ of probability distributions, [23, 36] suggest to solve the "robustified" version of (28):

$$\min_{\boldsymbol{x} \in \mathcal{X}} \max\{F(\boldsymbol{x}, P) : P \in \mathcal{P}\} \tag{29}$$

subject to $G(\boldsymbol{x}, P) \leq 0 \,\forall P \in \mathcal{P}$ or equivalently, subject to

$$\max_{P \in \mathcal{P}} G(\boldsymbol{x}, P) \leq 0. \tag{30}$$

Assume that $G(\boldsymbol{x}, P)$ is convex in $\boldsymbol{x}$ on $\mathcal{X}$ and *linear (in the sense that it is both convex and concave) in P on $\mathcal{P}$*. Then for convex, compact classes $\mathcal{P}$ and for a fixed $\boldsymbol{x}$, the maxima in (29), (30) are attained at extremal points of $\mathcal{P}$. Hence for the class $\mathcal{P}_{\boldsymbol{y}}$ identified by moment conditions (3) it is possible to pass in (29) and in (30) to discrete distributions $P \in \mathcal{P}_{\boldsymbol{y}}$ provided that the correspondingly extended basic assumptions are fulfilled. This convenient property carries over also to $G(\boldsymbol{x}, P)$ in (30) and/or $F(\boldsymbol{x}, P)$ in (29) *convex in P*.

Stochastic programs (28) whose constraints depend on the probability distribution have been applied frequently in portfolio optimization under *risk constraints.*

Whereas expected disutility functions or CVaR$(\boldsymbol{x}, P)$ are linear in P, other popular portfolio characteristics are not even convex in P: the variance is concave in P, the mean absolute deviation is neither convex nor concave in P. This means that extensions of the minimax approach to risk functionals nonlinear in P are carried through only under special circumstances delineated in the next example.

Example 4.1 Denote by ω the random vector of unit returns of assets included to the portfolio and let $f(\boldsymbol{x}, \omega) = -\omega^\top \boldsymbol{x}$ quantify the random loss of the investment $\boldsymbol{x}$. The probability distribution P of ω is known to belong to a class $\mathcal{P}_\mu$ of distributions for which, inter alia, the expectation $E_P \omega = \mu$ is fixed (independent of P). Then for a fixed $\boldsymbol{x}$, $\mathrm{var}_P f(\boldsymbol{x}, \omega) = E_P(\omega^\top \boldsymbol{x})^2 - (\mu^\top \boldsymbol{x})^2$ is linear in $P \in \mathcal{P}_\mu$ and the mean absolute deviation $\mathrm{MAD}_P f(\boldsymbol{x}, \omega) = E_P |\omega^\top \boldsymbol{x} - \mu^\top \boldsymbol{x}|$ is linear in $P \in \mathcal{P}_\mu$ as well.

5. Conclusions

The presented approach to the stability analysis of minimax stochastic programs with respect to input information was elaborated for the class $\mathcal{P}$ defined by generalized moment conditions (3) and a given set Ω. It is suitable also for other "parametric" classes $\mathcal{P}$ whereas stability for "nonparametric" classes, e.g. (5), would require different techniques. We did not aim at the most general statements and results on stability and sensitivity of minimax bounds and minimax decisions with respect to the model input. Specifically, various convexity assumptions were exploited: convexity and compactness of the class $\mathcal{P}_{\boldsymbol{y}}$, convexity of the random objective function $f(\boldsymbol{x}, \omega)$ with respect to the decision variable $\boldsymbol{x}$ on a compact convex set of feasible decisions, convexity of functionals $F(\boldsymbol{x}, P)$, $G(\boldsymbol{x}, P)$ with respect to the probability distribution P.

Convexity of the random objective with respect to $\boldsymbol{x}$ can be replaced by a saddle property and under suitable conditions, also unbounded sets $\mathcal{X}$ can be treated. An open question is under what general assumptions the presented approach can be applied to minimax problems with functionals nonconvex in P.

Acknowledgments

References

[1] G. A. Anastassiou, *Moments in Probability and Approximation Theory*. Longman Sci. & Tech, Harlow, UK, 1993.
[2] A. Ben-Tal and E. Hochman, *More bounds on the expectation of a convex function of a random variable*, J. Appl. Prob. 9 (1972), pp. 803–812.
[3] J. R. Birge and R. J.-B. Wets, *Computing bounds for stochastic programming problems by means of a generalized moment problem*, Math. Oper. Res. 12 (1987), pp. 149–162.
[4] W. Bühler, *Characterization of the extreme points of a class of special polyhedra*, Zeitschrift für Operations Research 19 (1975), pp. 131–137.
[5] G. C. Calafiore, *Ambiguous risk measures and optimal robust portfolios*, SIAM J. Optim. 18 (2007), pp. 853–877.
[6] E. Delage and Y. Ye, *Distributionally robust optimization under moment uncertainty with application to data-driven problems*, Operations Research 58 (2010), pp. 596–612.
[7] J. Dupačová, *Stability in stochastic programming with recourse – Estimated parameters*, Math. Program. 28 (1984), pp. 72–83.
[8] J. Dupačová, *Minimax stochastic programs with nonconvex nonseparable penalty functions*, in *Progress in Operations Research*, A. Prékopa, ed., J. Bolyai Math. Soc./North Holland, Budapest, 1976, pp. 303-316.

[9]   J. Dupačová, *Minimax approach to stochastic linear programming and the moment problem* (in Czech), EMO, 13 (1977), pp. 279–307; extended abstract ZAMM, 58 (1977), pp. T466–T467.

[10] J. Dupačová, *On minimax decision rule in stochastic linear programming*, in *Studies on Mathematical Programming*, A. Prékopa, ed., Akadémiai Kiadó, Budapest, 1980, pp. 47–60.

[11] J. Dupačová, *The minimax approach to stochastic programming and an illustrative application*, Stochastics 20 (1987), pp. 73–88.

[12] J. Dupačová, *Moment bounds for stochastic programs in particular for recourse problems*, in: *Distributions with given Marginals and Moment Problems*, V. Beneš and J. Štěpán, eds., Kluwer Academic Publ., Dordrecht, 1997, pp. 199–204.

[13] J. Dupačová, *Stochastic programming: minimax approach*, in *Encyclopedia of Optimization, Vol. V.*, Ch. A. Floudas and P. M. Pardalos, eds., Kluwer Academic Publ., Dordrecht 2001, pp.327–330.

[14] N. C. P. Edirisinghe and W. T. Ziemba, *Bounds for two-stage stochastic programs with fixed recourse*, Math. Oper. Res. 19 (1994), pp. 292–313.

[15] L. El Ghaoui, F. Oustry, and H. Lebret, *Robust solutions to uncertain semidefinite programs*, SIAM J. Optim. 9 (1998), pp.33–52.

[16] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

[17] P. Kall, *Stochastic programming with recourse: Upper bounds and moment problems - a review*, in *Advances in Mathematical Optimization*, J. Guddat, et al., eds., Akademie-Verlag, Berlin, 1988, pp. 86–103.

[18] V. Kaňková, *On an ε-solution of minimax problem in stochastic programming*, in: *Distributions with given Marginals and Moment Problems*, V. Beneš and J. Štěpán, eds., Kluwer Academic Publ., Dordrecht, 1997, pp. 211–216.

[19] A. F. Karr, *Extreme points of certain sets of probability measures with applications*, Math. Oper. Res. 8 (1983), pp. 74–85.

[20] J. M. B. Kemperman, *The general moment problem, a geometric approach*, Ann. Math. Statist. 39 (1968), pp. 93–122.

[21] J. M. B. Kemperman, *Moment problem with convexity conditions I.*, in: *Optimizing Methods in Statistics*, J. S. Rustagi, ed., Academic Press, 1971, pp. 115–178.

[22] A. Madansky, *Inequalities for stochastic programming problems*, Manag. Sci. 6 (1960), pp. 197–204.

[23] G. Pflug and D. Wozabal, *Ambiguity in portfolio selection*, Quantitative Finance 7 (2007), pp. 435–442.

[24] I. Popescu, *A semidefinite programming approach to optimal moment bounds for convex classes of distributions*, Math. Oper. Res. 30 (2005), pp. 632–657.

[25] A. Prékopa, *Stochastic Programming*, J. Bolyai Math. Soc./Kluwer Academic Publ., Budapest, 1995.

[26] M. Riis and K. A. Andersen, *Applying the minimax criterion in stochastic recourse programs*, European J. of Oper. Res. 165 (2005), pp. 569–584.

[27] S. M. Robinson, *A characterization of stability in linear programming*, Operations Research 25 (1977), pp. 435–447.

[28] R. T. Rockafellar, *Convex Analysis*, 2nd Ed., Princeton University Press, 1972.

[29] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

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

[31] A. Shapiro, *Worst-case distribution analysis of stochastic programs*, Math. Program. B 107 (2006), pp. 91–96.

[32] A. Shapiro and S. Ahmed, *On a class of minimax stochastic programs*, SIAM J. Optim. 14 (2004), pp. 1237–1249.

[33] A. Shapiro and A. Kleywegt, *Minimax analysis of stochastic programs*, GOMS 17 (2002), pp. 532–542.

[34] A. Thiele, *Robust stochastic programming with uncertain probabilities*, IMA J. of Manag. Math. 19 (2008), pp. 289–321.

[35] D. W. Walkup and R. J.-B. Wets, *Stochastic programs with recourse II: On the continuity of the objective*, SIAM J. Appl.Math. 17 (1969), pp. 98–103.

[36] D. Wozabal, *A framework for optimization under ambiguity*, TR2008-08, Dept of Statistics and Decision Support Systems, University of Vienna, August 2008; submitted.