STOCHASTIC PROGRAMS WITH COMPLETE RECOURSE: STABILITY AND AN APPLICATION TO POWER DISPATCH

Werner Römisch and Rüdiger Schultz

Humboldt-Universität Berlin Sektion Mathematik DDR-1086 Berlin

Abstract: For convex stochastic programs with (linear) complete recourse, we review and illustrate (quantitative) stability results for optimal solutions if the underlying probability distribution is subjected to perturbations. We show that the general results apply to a recourse model for economic dispatch of electric power with random demand.

1. Introduction

Consider the following stochastic program with (linear) complete recourse

(1.1) min
$$\{g(x) + Q_{\mu}(x) : x \in C\}$$

where

(1.2)
$$Q_{\mu}(x) = \int_{\mathbb{R}^{5}} \widetilde{Q}(h(x,z)) \, \mu(dz)$$

$$(1.3) \qquad \widetilde{Q}(t) = \min \left\{ q^{\mathsf{T}} y : \mathsf{W} y = t , y \geqslant 0 \right\}.$$

For the data in (1.1)-(1.3) we assume that g is a real-valued convex function on \mathbb{R}^m , C is a nonempty, closed, convex subset of \mathbb{R}^m , μ is a (Borel) probability measure on \mathbb{R}^s , $q \in \mathbb{R}^{\overline{m}}$, $W \in L(\mathbb{R}^{\overline{m}}, \mathbb{R}^{\Gamma})$ and h is a mapping from $\mathbb{R}^m \times \mathbb{R}^s$ to \mathbb{R}^r which is affine linear in x and globally Lipschitzian in z. Under the basic assumptions

- (A1) $\{Wy : y \in \mathbb{R}^{\overline{n}}, y \geqslant 0\} = \mathbb{R}^{\Gamma}$,
- (A2) there exists $u \in \mathbb{R}^{\Gamma}$ such that $W^{T}u \leq q$,
- (A3) $\int_{\mathbb{R}^{S}} ||z|| \, \mu(dz) < + \infty ,$

we have that $\widetilde{\mathbb{Q}}(h(x,z)) \in \mathbb{R}$ for all $x \in \mathbb{R}^m$, $z \in \mathbb{R}^s$ (this holds in view of (A1) and (A2)) and that \mathbb{Q}_{μ} is a real-valued convex function on \mathbb{R}^m (here, (A3) is used together with properties of $\widetilde{\mathbb{Q}}$ as an optimal-value function for linear programs with parameters in the right-hand side of the constraints), cf. e.g. [8].

The problem (1.1) arises as a deterministic equivalent to an improperly posed program

(1.4) min
$$\{g(x) : x \in C, h(x,z) = 0\}$$

where $z \in \mathbb{R}^{S}$ is a vector of random data.

The program (1.4) is improperly posed in the sense that no decision about feasibility of $x \in \mathbb{R}^m$ is possible before knowing the realization of z. In many practical situations, however, a decision on x has to be taken before knowing the realization of z, such that usually a deviation $h(x,z) \neq 0$ occurs. The basic idea in stochastic programming with recourse now consists in allowing for a compensation of such deviations by a second-stage decision, which is formalized in (1.3). The average costs of the compensation (cf. (1.2)) are added to the objective of (1.4) and we end up with (1.1). In this context, (A1) ensures that any deviation h(x,z) can be compensated, which leads to the terminus "complete recourse".

The present paper is concerned with the (quantitative) stability of optimal value and optimal solutions to (1.1) when the underlying probability measure μ is subjected to perturbations. Qualitative stability results for (more general) classes of stochastic programs with recourse were obtained in [9] and [13]. Such stability considerations are motivated by two essential difficulties one is generally confronted with when analyzing stochastic programs (among which recourse problems like (1.1) form only a specific but important class, cf. e.g. [8], [19] for further approaches):

Firstly, stochastic programs usually involve multidimensional integrals (like (1.2), for instance) which allow for numerical treatment only when approximating the underlying measure by simpler ones (see [5] and the references therein). Secondly, in practice information on the underlying measure is often available only in the form of statistical estimates, which leads to questions about the asymptotic behaviour of optimal values and optimal solutions (see [4], [17] and the references therein).

Based on metrizing the set $\mathfrak{P}(\mathbb{R}^{\mathbf{s}})$ of all (Borel) probability measures on $\mathbb{R}^{\mathbf{s}}$ in a suitable way we review in Section 2 for stochastic programs fitting into (1.1) local Lipschitz and Hölder properties for the optimal value and the set of optimal solutions, respectively. In Section 3 we discuss consequences of these results for a specific recourse model arising in optimal power dispatch.

2. Stability Properties

When aiming at a quantitative stability analysis for (1.1) that covers both approximations of μ in the situation of complete information and statistical estimates for μ when only incomplete information is available, it is recommendable to equip the set $\mathfrak{P}(\mathbb{R}^5)$ with a suitable probability metric. A metrization which, simultaneously,

induces convergence for a sufficiently broad class of measures and allows for a computation (or at least estimation) of the distance in important specific situations is given by L_p - Wasserstein distances W_p (p \geqslant 1) which are defined by

for all

$$\mu, \nu \in \mathfrak{M}_{p}(\mathbb{R}^{s}) := \left\{ \mu \in \mathfrak{P}(\mathbb{R}^{s}) : \int \|z\|^{p} \mu(dz) < \infty \right\},$$

where $D(\mu,\nu):=\left\{\gamma\in\mathfrak{P}(\mathbb{R}^{S}\times\mathbb{R}^{S}):\gamma\circ\pi_{1}^{-1}=\mu\;,\;\gamma\circ\pi_{2}^{-1}=\nu\right\}$ and π_{1} , π_{2} are the first and second projections, respectively. From the literature ([7],[12]) it is known that $(m_{p}(\mathbb{R}^{S}),w_{p})$ is a metric space and that convergence in w_{p} is characterized by: $w_{p}(\mu_{n},\mu)\longrightarrow 0$ for $\mu\in\mathfrak{M}_{p}(\mathbb{R}^{S})$ and $\mu_{n}\in\mathfrak{P}(\mathbb{R}^{S})$ if and only if the sequence (μ_{n}) converges weakly to μ (for the definition consult [1])

and
$$\lim_{n\to\infty} \int_{\mathbb{R}^s} \|z\|^p \, \mu_n(\mathrm{d}z) = \int_{\mathbb{R}^s} \|z\|^p \, \mu(\mathrm{d}z)$$
.

Now, let us denote by ψ the function from $(\mathbf{M}_1(\mathbb{R}^S), \mathbb{W}_1)$ to \mathbb{R} which assigns to $\nu \in \mathbf{M}_1(\mathbb{R}^S)$ the (global) optimal value of (1.1) with underlying measure ν . By ψ we denote the set-valued mapping from $(\mathbf{M}_1(\mathbb{R}^S), \mathbb{W}_1)$ to \mathbb{R}^m which assigns to $\nu \in \mathbf{M}_1(\mathbb{R}^S)$ the set of (global) minimizers of (1.1) with underlying measure ν .

Our first stability result asserts upper semicontinuity of ψ and a local Lipschitz property of ψ at μ (Recall that ψ is upper semicontinuous at μ if for each open set U containing $\psi(\mu)$ there exists $\delta_0>0$ such that $\psi(\nu)\subset U$ whenever $W_1(\mu,\nu)<\delta_0$.). It is a consequence of Theorem 2.4 and Remark 2.5 in [16].

Theorem 2.1:

Fix $\mu\in\mathfrak{P}(\mathbb{R}^3)$ and suppose (A1) - (A3). Let $\psi(\mu)$ nonempty and bounded. Then ψ is upper semicontinuous at μ and there exist constants L>0 and $\delta>0$ such that $\psi(\nu)$ $\not\models\emptyset$ and

The following example shows that, under the assumptions of the Theorem, ψ is in general not lower semicontinuous at μ (Recall that lower semicontinuity of ψ at μ means that for each open set U satisfying Un $\psi(\mu)$ \ddagger Ø there exists $\delta_o>0$ such that Un $\psi(\nu)$ \ddagger Ø whenever W₁(μ,ν) < δ_o .).

Example 2.2:

In (1.1), let m = s = r := 1, $\overline{m} := 2$, g(x) := -x, C := [0,1], $q := (1,1)^T$, W := (1,-1) and h(x,z) := xz. Then Q(t) = |t| and (A1), (A2) are satisfied. Let $\mu \in \mathcal{P}(\mathbb{R})$ be such that $\int |z| \mu(dz) = 1$ and $\mu_n \in \mathcal{P}(\mathbb{R})$

 $(n \in \mathbb{N})$ be chosen such that $M_n := \int\limits_{\mathbb{R}} |z| \mu_n(dz) > 1$, $M_n \longrightarrow 1$ and (μ_n) converges weakly to μ .

Hence, we have
$$W_1(\mu, \mu_n) \longrightarrow 0$$
 and $g(x) + Q_{\nu}(x) = -x + x + \sum_{n=1}^{\infty} |z| \nu(dz)$ $(x \in C, \nu \in M_1(\mathbb{R})).$

Then $\psi(\mu)=C$ and $\psi(\mu_n)=\left\{0\right\}$ for each $n\in\mathbb{N}$, thus implying that Ψ is not lower semicontinuous at μ.

However, for the case h(x,z) := z - Ax with a (non-stochastic) matrix $A \in L(\mathbb{R}^m, \mathbb{R}^s)$ Hausdorff-continuity of ψ at μ has even been quantified in [15], [16]. In this context, strong convexity properties of the function

$$Q_{\mu}^{*}(\xi) := \int_{\mathbb{R}^{S}} \widetilde{Q}(z - \xi) \, \mu(dz) \quad (\xi \in \mathbb{R}^{S})$$

with \widetilde{Q} as in (1.3) play an important role.

 $(Q_{ij}^*$ is strongly convex on a convex subset V of \mathbb{R}^S if there exists $k^{\mu} > 0$ such that for all $\xi, \xi \in V$ and $\lambda \in [0,1]$,

 $Q_{\mu}^{*}(\lambda\xi + (1-\lambda)\tilde{\xi}) \leq \lambda Q_{\mu}^{*}(\xi) + (1-\lambda)Q_{\mu}^{*}(\tilde{\xi}) - k\lambda(1-\lambda)\|\xi - \tilde{\xi}\|^{2}.)$ For this situation, the next result is proved in [16].

Theorem 2.3:

Let, in (1.1), g be convex quadratic and C a polyhedron.

Fix $\mu \in \mathcal{P}(\mathbb{R}^s)$ and suppose (A1) - (A3). Let further $\psi(\mu)$ nonempty, bounded and the function Q_{μ}^{\pm} strongly convex on a convex open set V containing $A(\Psi(\mu))$.

Then, there exist constants L>0 and δ >0 such that

$$d_{H}(\psi(\mu),\psi(\nu)) \leqslant L W_{1}(\mu,\nu)^{1/2}$$

whenever $y \in M_1(\mathbb{R}^s)$, $W_1(\mu, v) < \delta$. (Here d_H denotes the Hausdorff distance on subsets of \mathbf{R}^{m} .)

Before discussing consequences for a specific recourse model, let us add a few comments on the above results.

Theorem 2.1 can be extended to more general recourse problems (quadratic recourse, linear recourse with also q in (1.3) random), cf. Theorem 2.4 in [16]. Verification of the strong-convexity assumption in Theorem 2.3 is possible for specific instances of (1.3), namely if either $W \in L(\mathbb{R}^{s+1}, \mathbb{R}^s)$, $q \notin \text{im } W^T$ or $W \in L(\mathbb{R}^{2s}, \mathbb{R}^s)$,

W = (H,-H), q⁺ + q⁻ > 0 (where H \in L(\mathbb{R}^S , \mathbb{R}^S) is non-singular, q = (q⁺, q⁻)^T, q⁺, q⁻ \in \mathbb{R}^S and the strict inequality understood componentwise). The existence of a density for μ which is locally bounded below by a positive number and conditions on μ that ensure Q^*_{μ} to be continuously differentiable with locally Lipschitzian gradient then imply strong convexity of Q^*_{μ} , cf. Theorem 3.5, Corollary 3.6 in [15].

Theorem 2.3 does not hold for general convex g and C, and the exponent 1/2 on the right-hand side of the estimate is best possible, cf. Examples 4.5, 4.6 in [15], Remark 2.9 in [16].

Although calculation (or estimation) of distances between probability measures is in general a formidable task, explicit formulae for L_p -Wasserstein metrics are known in specific situations ([6], [7]). For probability measures on $\mathbb R$ the following holds (cf. e.g. [12]):

$$W_1(\mu,\nu) = \int_{-\infty}^{\infty} |F_{\mu}(t) - F_{\nu}(t)| dt$$

where F_{μ} , F_{ν} are the distribution functions for $\mu,\nu\in M_1(\mathbb{R})$. This formula is remarkable, since for measures $\mu,\nu\in M_p(\mathbb{R}^s)$ with independent one-dimensional marginal distributions $\mu_1,\,\nu_1$ (i=1,..,s) we have (cf. Remark 2.11 in [16])

$$W_{p}(\mu,\nu) \leq C_{o}(\sum_{i=1}^{s} W_{p}(\mu_{i},\nu_{i})^{p})^{1/p}$$

with a computable constant $C_o > 0$.

The recourse model which we discuss in the next section has the property that it only depends on the one-dimensional marginal distributions of the underlying probability measure, such that the above formulae apply.

3. Application to optimal power dispatch with uncertain demand

In this section, we consider an energy production system consisting of thermal power stations (tps), pumped (hydro) storage plants (psp) (serving as base- and peak-load plants, respectively) and an energy contract (ec) with connected systems. The problem of optimal power dispatch consists of allocating amounts of electric power to the generation units of the system (i.e. tps, psp and ec) such that the total generation costs are minimal while the actual power demand is met and certain operational constraints are satisfied.

The poculiarities of the model we shall discuss are the following:
(a) The model is designed for a daily operating cycle and assumes
that a unit commitment stage has been carried out before, (b) the
transmission losses are modeled by means of an adjusted portion of

the demand, (c) the cost functions of the thermal plants are taken to be strictly convex and quadratic. A special feature of our model is that we take into account the randomness of the electric power demand.

Let K and M denote the number of tps and psp, respectively, and N be the number of subintervals in the discretization of the planning period. Let $I_{r} \subset \{1,...,K\}$ denote the index set of available online tps within the time interval r $\in \{1,...,N\}$. The (unknown) outputs of the tps and psp at the time interval r are y_{1r} (l=1,...,K) and s_{1r} (generation mode of the psp $j \in \{1,...,M\}$), respectively. By w_{jr} we denote the input of the psp j during the pumping mode and by $e_{\mathbf{r}}$ the level of electric power which corresponds to the contract at time interval r.

Denoting $x:=(y,s,w,e)^T \in \mathbb{R}^m$ with m:=N(K+2M+1) our model for optimal power dispatch developed in [14] has the following shape

(3.1)
$$\min \{g(x) : x \in C, Ax = z \}.$$

In (3.1), g is a convex quadratic cost function defined on \mathbb{R}^m , ${\tt C} \subset {
m I\!R}^{\sf m}$ is a (nonempty) bounded convex polyhedron containing the restrictions for the power output, balances between generation and pumping in the psp, balances over the whole time horizon for the psp and according to the energy contract, fuel quotas in the tsp etc. The equation $Ax \approx z$ in (3.1) reads componentwise (i.e. at time interval r)

(3.2)
$$[Ax]_r := \sum_{i \in I_r} y_{ir} + \sum_{j=1}^{M} (s_{jr} - w_{jr}) + e_r = z_r$$

and means that the total generated output meets the demand $z=(z_1,\dots$ $(z_N)^T$ at each time interval.

We consider the demand z as a random vector and denote by $\mu \in \mathfrak{P}(\mathbb{R}^N)$ the probability distribution of z and by F, the distribution function of z_r (r=1,...,N). Distinct from the approach in [14], where the equilibrium between total generation and (random) demand has been modeled by a probabilistic constraint, we here consider a stochastic formulation of (3.1) as a recourse model. Its basic idea is to introduce a certain penalty cost for the deviation of the scheduled output from the actual demand for under- and over-dispatching, respectively. To be more precise, we define

$$\widetilde{\mathbb{Q}}(t) := \sum_{r=1}^{N} \widetilde{\mathbb{Q}}_{r}(t_{r}) := \sum_{r=1}^{N} \left\{ \begin{array}{l} q_{r}^{\dagger}t_{r} \;,\; t_{r} \geqslant 0 \\ -q_{r}^{\dagger}t_{r} \;,\; t_{r} < 0 \end{array} \right. \; (t \in \mathbb{R}^{N}),$$
 where q_{r}^{\dagger} and q_{r}^{\dagger} are the recourse costs for the under- and over-dis-

patching at time interval $r \in \{1,...,N\}$, respectively.

The power dispatch model then has the form

(3.3)
$$\min \left\{ g(x) + E\left[\widetilde{Q}(z - Ax)\right] : x \in C \right\}$$

where E[\cdot] denotes the expectation (i.e. the integral over \mathbf{R}^N with respect to the measure μ).

Similar power dispatch models are considered in [2] (Chapter 3.3), [3] and [20]. For more information on power dispatch, especially stochastic models, we refer to [18] and to several papers in [5].

Observing that $\widetilde{\mathbb{Q}}(t)=\inf\left\{q^Ty: \text{Wy}=t, y\geqslant 0\right\}$ holds with $q:=(q_1^+,\dots,q_r^+,q_1^-,\dots,q_r^-)^T$ and W:=(I,-I) (I denoting the identity matrix in $L(\mathbb{R}^N)$), (3.3) is a special instance of the general stochastic program with recourse (1.1). It is well-known that (A1),(A2) are satisfied if $q_r^++q_r^-\geqslant 0$ for each $r=1,\dots,N$ (cf.e.g. [8]). Now, we are in the position to apply the general stability results from Section 2 to the special recourse model (3.3). We still need the following 'distance' on $\mathfrak{P}(\mathbb{R}^N)$:

$$d(v_1, v_2) := \sum_{r=1}^{N} \int_{-\infty}^{\infty} |F_{1r}(t) - F_{2r}(t)| dt$$

where F_{1r} and F_{2r} are the one-dimensional marginal distribution functions of $v_1, v_2 \in \mathfrak{P}(\mathbb{R}^N)$.

Theorem 3.1:

Consider (3.3) with general assumptions as above, let $\mu \in \mathfrak{M}_1(\mathbb{R}^N)$ and $\mathfrak{q}_r^+ + \mathfrak{q}_r^- \geqslant 0$ for each r=1,...,N.

- (i) Then ψ is upper semicontinuous at μ (with respect to the distance d) and there exist constants L>0 and δ >0 such that $|\varphi(\mu) \varphi(\nu)| \le L d(\mu, \nu)$ whenever $d(\mu, \nu) < \delta$.
- (ii) Assume, additionally, that $q_r^+ + q_r^- > 0$ for each $r=1,\ldots,N$, μ has bounded marginal densities Θ_r ($r=1,\ldots,N$) and that there exists $c_0>0$ such that

exists $c_0>0$ such that $\Theta(t):=\frac{N}{\prod_{r=1}^N}\;\Theta_r(t_r)\geqslant c_0\quad\text{for all }t=(t_1,\ldots,t_N)^T$ in some open subset U of \mathbb{R}^N containing the set $A(\psi(\mu))$. Then there exist constants $L_1>0$ and $\delta_1>0$ such that

 $d_{\mu}(\psi(\mu), \psi(\nu)) \leqslant L_1 d(\mu, \nu)^{1/2}$ whenever $d(\mu, \nu) < \delta_1$.

Proof:

Part (i) is a consequence of Theorem 2.1. It remains to note that, since (3.3) only depends on the marginal distribution functions F_r (r=1,...,N), the final remark of Section 2 applies and $W_1(\mu,\nu)$ may be estimated by

$$C_0 \sum_{r=1}^{N} \int_{-\infty}^{\infty} |F_r(t) - F_{\gamma r}(t)| dt = C_0 d(\mu, \gamma)$$
,

where $C_0 > 0$ is a certain constant and F_{γ_r} (r=1,...,N) are the marginal distribution functions of $\gamma \in \mathfrak{P}(\mathbb{R}^N)$.

To prove (ii), we first remark that, according to the assumptions, Q_{μ}^{*} (for a definition see Section 2) is strongly convex on each bounded convex open set V with $A(\psi(\mu))\subset V\subset U$ (Theorem 3.5 in [15]). To see this, we note that again μ may be replaced by the measure $\bar{\mu}$ being the product of the marginal distributions of μ and that Θ is the density of $\bar{\mu}$. Now, Theorem 2.3 applies and the proof is complete.

This means that for our power dispatch model the optimal costs behave Lipschitz continuous, and, under suitable assumptions on the marginal densities of the random demand vector, the optimal sets enjoy a Hölder continuity property with respect to the computable distance d.

The following equivalent form of (3.3) via the introduction of a new variable $X \in \mathbb{R}^N$ (called 'tender') proves useful for numerical purposes:

(3.4)
$$\min \left\{ g(x) + Q(X) : x \in C, Ax = X \right\},$$

where
$$Q(X) := \sum_{r=1}^{N} E[\tilde{Q}_r(z_r - X_r)] (X \in \mathbb{R}^N).$$

(3.4) is a nonlinear convex separable program in which the number of variables occuring nonlinearly in the recourse part is N instead of $m \gg N$.

For an extensive discussion of numerical methods for the solution of (3.4) we refer to [10] (and also to their papers in [5]) and to the recent work in [2](Chapter 4) and in [11].

References:

- [1] P. Billingsley: Convergence of Probability Measures, Wiley, New York, 1968.
- [2] J. Böttcher: Stochastische lineare Programme mit Kompensation, Mathematical systems in economics, Vol. 115, Athenaum Verlag, Frankfurt am Main, 1989.
- [3] D.W. Bunn and S.N. Paschentis: Development of a stochastic model for the economic dispatch of electric power, European Journal of Operational Research 27(1986), 179-191.
- [4] J. Dupačová: Stability and sensitivity analysis for stochastic programming, Annals of Operations Research (to appear).
- [5] Y. Ermoliev and R.J-B. Wets (Eds.): Numerical Techniques for Stochastic Optimization, Springer-Verlag, Berlin, 1988.
- [6] M. Gelbrich: On a formula for the L²- Wasserstein metric between measures on Euclidean and Hilbert spaces, Mathematische Nachrichten (to appear).

- [7] C.R. Givens and R.M. Shortt: A class of Wasserstein metrics for probability distributions, Michigan Mathematical Journal 31 (1984), 231-240.
- [8] P. Kall: Stochastic Linear Programming, Springer-Verlag, Berlin,
- [9] P. Kall: On approximations and stability in stochastic programming, Parametric Optimization and Related Topics (J. Guddat, H.Th. Jongen, B. Kummer, F. Nožička, Eds.), Akademie-Verlag, Berlin, 1987, 387-407.
- [10] J.L. Nazareth and R.J-B. Wets: Algorithms for stochastic programs: The case of nonstochastic tenders, Mathematical Programming Study 28(1986), 1-28.
- [11] A. Prékopa: Dual method for the solution of a one-stage stochastic programming problem with random rhs obeying a discrete probability distribution, Zeitschrift für Operations Research (to appear).
- [12] S.T. Rachev: The Monge-Kantorovich mass transference problem and its stochastic applications, Theory of Probability and its Applications 29(1984), 647-676.
- [13] S.M. Robinson and R.J-B. Wets: Stability in two-stage stochastic programming, SIAM Journal on Control and Optimization 25 (1987), 1409-1416.
- [14] W. Römisch and R. Schultz: Distribution sensitivity for certain classes of chance constrained models - with application to power dispatch, manuscript, 1989 and submitted to Journal of Optimization Theory and Applications.
- [15] W. Römisch and R. Schultz: Stability of solutions for stochastic programs with complete recourse having C^{1,1} data, Manuskript, Institut für Operations Research der Universität Zürich, 1989.
- [16] W. Römisch and R. Schultz: Stability analysis for stochastic programs, Humboldt-Universität Berlin, Sektion Mathematik, Preprint Nr. 232 (1989).
- [17] A. Shapiro: Asymptotic analysis of stochastic programs, University of South Africa, Department of Mathematics, Research Report 90/89(12), 1989.
- [18] Hj. Wacker (Ed.): Applied Optimization Techniques in Energy Problems, Teubner, Stuttgart, 1985.
- [19] R.J-B. Wets: Stochastic programming, Handbooks in Operations Research and Management Science, Vol. 1, Optimization (G.L. Nemhauser, A.H.G. Rinnoy Kan, M.J. Todd, Eds.), North-Holland, 1989.
- [20] R. Wiebking: Stochastische Modelle zur optimalen Lastverteilung in einem Kraftwerksverbund, Zeitschrift für Operations Research 21(1977), B197-B217.