



Konrad-Zuse-Zentrum für Informationstechnik Berlin
Takustraße 7, D-14195 Berlin

Rüdiger Schultz

**A Note on Preprocessing via Fourier - Motzkin Elimination in
Two-Stage Stochastic Programming**

A Note on Preprocessing via Fourier - Motzkin Elimination in Two-Stage Stochastic Programming

Rüdiger Schultz

Abstract

Preprocessing in two-stage stochastic programming is considered from the viewpoint of Fourier-Motzkin elimination. Although of exponential complexity in general, Fourier-Motzkin elimination is shown to provide valuable insights into specific topics such as solving integer recourse stochastic programs or verifying stability conditions. Test runs with the computer code PORTA [2] are reported.

Key Words: Stochastic programming, preprocessing, Fourier-Motzkin elimination

1991 Mathematics Subject Classification: 90 C 15, 90 C 11, 90 C 31

1 Introduction

Two-stage stochastic programs arise as deterministic equivalents to random optimization problems that are characterized by a two-stage scheme of alternate decision and observation. First, a here-and-now decision has to be taken without knowing the outcomes of random problem data. After realization of the random data, a second-stage (recourse) decision is possible, which is the solution of a subordinate optimization problem depending on the first-stage solution and the outcome of the random data. In this paper, we consider problems where the second-stage is a linear program with possibly integer requirements on the variables. In two-stage stochastic programming one encounters an interplay of algebraic and probabilistic difficulties. Preprocessing in stochastic programming is directed to analyzing the underlying algebraic structures. This may be helpful for supporting solution procedures or improving structural understanding of the problem.

In this note, we consider stochastic programs of the following form

$$\min\{c^T x + Q(x) : x \in C\} \quad (1)$$

where

$$Q(x) = \int_{\mathbb{R}^s} \Phi(z - Ax)\mu(dz) \quad (2)$$

and

$$\Phi(t) = \min\{q^T y : Wy \geq t, y \in Y\} \quad \text{with } Y = \mathbb{R}_+^m \quad \text{or} \quad Y = \mathbb{Z}_+^m. \quad (3)$$

Here, $C \subset \mathbb{R}^n$ is a non-empty polyhedron and c, q, W are vectors and matrices of proper dimensions. The above mentioned scheme of alternate decision and observation is reflected by the first-stage variables x , the random vector z with underlying probability measure μ and the second-stage variables y . The two-stage stochastic program aims at finding a first-stage decision x such that the sum of direct costs $c^T x$ and expected recourse costs $Q(x)$ is minimal. For further reading on basics in two-stage stochastic programming we refer to [5, 6]. Let us assume that $Y = \mathbb{R}_+^m$ and impose the following assumptions

$$\text{the matrix } W \text{ has full rank,} \quad (4)$$

$$M_D = \{u \in \mathbb{R}_+^s : W^T u \leq q\} \neq \emptyset, \quad (5)$$

$$\int_{\mathbb{R}^s} \|z\| \mu(dz) < +\infty. \quad (6)$$

These assumptions imply that M_D has vertices, and according to the decomposition theorem for polyhedra it admits a representation

$$M_D = \text{conv}(V_0) + \text{cone}(V_1) \quad (7)$$

where V_0, V_1 are finite sets of vectors in \mathbb{R}^s and conv and cone denote the convex and conical hulls, respectively. Preprocessing, as discussed in this note, concerns the algorithmic transformation of the representation in (5) into the one of (7). To see that explicit knowledge of V_0, V_1

can be beneficial, recall the following identities

$$\begin{aligned}
\text{cone}(V_I) &= \{u \in \mathbb{R}_+^s : W^T u \leq 0\} \\
&\quad (\text{since } \text{cone}(V_I) \text{ is the recession cone of } M_D) \\
&= \{u \in \mathbb{R}^s : u^T w_i \leq 0, i = 1, \dots, m, \quad u^T (-e_j) \leq 0, j = 1, \dots, s\} \\
&\quad (w_i \text{ are the columns of } W \text{ and } e_j \text{ the canonical unit vectors}) \\
&= \{u \in \mathbb{R}^s : u^T (\sum_i \lambda_i w_i + \sum_j \lambda_j (-e_j)) \leq 0 \quad \forall \lambda_i \geq 0, \forall \lambda_j \geq 0\} \\
&= \{u \in \mathbb{R}^s : u^T w \leq 0 \quad \forall w \in \text{pos}(W, -I)\} \\
&= (\text{pos}(W, -I))^*
\end{aligned}$$

where pos denotes the positive span and $*$ indicates the polar cone. Therefore, the elements of V_I are the coefficients in an inequality description of $\text{pos}(W, -I)$. In computations, the probability measure μ is in general assumed to be discrete, i. e., with mass points z_1, \dots, z_L and probabilities p_1, \dots, p_L . Then, $Q(x)$ is well defined if $z_l - Ax \in \text{pos}(W, -I)$ for all $l = 1, \dots, L$. In the literature, the latter are called induced constraints. These are implicit conditions that are made explicit if V_I is known. Moreover, for $t \in \text{pos}(W, -I)$ it holds that

$$\Phi(t) = \max_{v_i \in V_0} v_i^T t. \quad (8)$$

such that computation of Φ becomes easy if V_0 is known, and the domain of definition of Φ becomes explicit if V_I is known.

Nevertheless, the above considerations have only limited impact on solution procedures for state-of-the-art linear recourse problems (i.e., if $Y = \mathbb{R}_+^n$), since far too many elements arise in V_0 and V_I . Methods like L-shaped, regularized or stochastic decomposition [14, 8, 4], for instance, generate only those elements of V_0 and V_I that are relevant for the solution process. On the other hand, the above transformation may be useful to support solution procedures for smaller problems of more complicated nature (e.g., if $Y = \mathbb{Z}_+^n$) or for answering questions in the theory of stochastic programming (e.g., verifying stability conditions). The emphasis of our paper is on these two specific issues. A more general view on preprocessing including its impact on modeling is adopted in [15, 16] and Chapter 5 in [5].

In Section 2, we recall the role of Fourier-Motzkin elimination when transforming an inequality description of a polyhedron into the representation as Minkowski sum of a convex and a conical hull. In Section 3, we show how preprocessing via Fourier-Motzkin elimination enters into an algorithm for integer recourse stochastic programs. Section 4 deals with the verification of stability conditions for stochastic programs with linear recourse. Here Fourier-Motzkin elimination can be beneficial in generating information about the polyhedral complex of lineality regions of the second-stage value function Φ . Finally, we have a conclusions section.

2 Theoretical Background

In this section we recall the essence of Fourier-Motzkin elimination and put it into the context of computing extreme points and extreme rays of the polyhedron M_D (cf. (5)). Enumeration of extreme points and extreme rays of a polyhedron is a well studied problem in the literature (see, e.g., [18] and the references therein). In connection with stochastic programming, an excellent account is given in [15]. Our intention here is to show the relation to Fourier-Motzkin elimination, which is interesting from the practical computations point of view, since there exists a freely available implementation ([2]) of the transformation procedure described below. This supplements the implementations reported in [15].

For a polyhedron $\Pi = \{x \in \mathbb{R}^n : Ax \leq b\}$, Fourier - Motzkin elimination provides an algorithmic way for projection along the k -th coordinate ($1 \leq k \leq n$), i.e., for eliminating the variable x_k from the system $Ax \leq b$. Let $a_i, b_i (i = 1, \dots, l)$ be the rows of A and components of b , respectively, and let a_{ik} denote the k -th component of a_i . We introduce subsets $I_>, I_<, I_ =$ of $\{1, \dots, l\}$ such that

$$\begin{aligned} a_{ik} &> 0 && \text{for all } i \in I_>, \\ a_{ik} &< 0 && \text{for all } i \in I_<, \\ a_{ik} &= 0 && \text{for all } i \in I_ =. \end{aligned}$$

Simple manipulations then provide for any $x \in \Pi$

$$x_k \leq \frac{1}{a_{ik}}(a_{ik}x_k - a_i^T x + b_i) \quad \text{for all } i \in I_>$$

and

$$x_k \geq \frac{1}{-a_{ik}}(-a_{ik}x_k + a_i^T x - b_i) \quad \text{for all } i \in I_<.$$

The right-hand sides in these inequalities are independent of x_k such that we obtain for the projection $\Pi_{(k)}$ of Π along the k -th coordinate

$$\Pi_{(k)} = \left\{ \begin{array}{l} x_{(k)} \in \mathbb{R}^{(n-1)} : \\ \max_{i \in I_>} \left\{ \frac{1}{-a_{ik}}(-a_{ik}x_k + a_i^T x - b_i) \right\} \leq \min_{i \in I_<} \left\{ \frac{1}{a_{ik}}(a_{ik}x_k - a_i^T x + b_i) \right\}, \\ a_i^T x \leq b_i \quad \text{for all } i \in I_ = \end{array} \right\}$$

The first inequality in the above description of $\Pi_{(k)}$ can be equivalently expressed by $|I_>| \cdot |I_<|$ many linear inequalities, where $|\cdot|$ denotes cardinality. When projecting further down to smaller dimensions, the above scheme has to be iterated, which in general produces inequality systems of enormous size. This prevents algorithmic use of the method for large-scale problems. Strategies going back to Tschernikow ([13], see also [3]) allow to produce a description of $\Pi_{(k)}$ without redundant inequalities such that in practical computations intermediate inequality systems can be kept as small as possible.

Proposition 2.1 *Let $M_D = \{u \in \mathbb{R}_+^s : W^T u \leq q\}$ be such that $0 \in M_D$, then a representation $M_D = \text{conv}(V_0) + \text{cone}(V_1)$ can be computed by using the above elimination procedure.*

Proof:

We consider the polar polyhedron M_D^* of M_D , which is given as follows

$$M_D^* = \{\xi \in \mathbb{R}^s : \xi^T u \leq 1 \quad \forall u \in M_D\}.$$

Then it holds (cf. [9], Theorem 9.1) that M_D^* is a polyhedron again and that $M_D^{**} = M_D$. Let M_D be written as

$$M_D = \{u \in \mathbb{R}_+^s : (W, -I)^T u \leq \begin{pmatrix} q \\ 0 \end{pmatrix}\}$$

where W is scaled in such a way that $q_i \in \{0, 1\}$ for all components q_i of q , ($i = 1, \dots, m$). The latter is possible since $0 \in M_D$ implies $q \geq 0$. Another standard result on polarity (cf. again [9], Theorem 9.1) now yields that

$$M_D^* = \text{conv}\left(\{0\} \cup \bigcup_{q_i=1} \{w_i\}\right) + \text{cone}\left(\bigcup_{q_i=0} \{w_i\} \cup \bigcup_{j=1}^m \{-e_j\}\right)$$

where w_i and e_j are as in Section 1. Therefore

$$M_D^* = \left\{ u \in \mathbb{R}^s : \begin{array}{l} \exists \lambda \in \mathbb{R}_+^{|q_i=1|} \exists \mu^1 \in \mathbb{R}_+^{|q_i=0|} \exists \mu^2 \in \mathbb{R}_+^m \\ u = \sum_i \lambda_i w_i + \sum_k \mu_k^1 w_k + \sum_j \mu_j^2 (-e_j), \\ \sum_i \lambda_i = 1 \end{array} \right\}.$$

Hence

$$M_D^* = \mathcal{M}_{D(\lambda, \mu^1, \mu^2)}^*$$

where

$$\mathcal{M}_D^* = \left\{ (u, \lambda, \mu^1, \mu^2) : \begin{array}{l} u = \sum_i \lambda_i w_i + \sum_k \mu_k^1 w_k + \sum_j \mu_j^2 (-e_j), \\ \sum_i \lambda_i = 1, \lambda \geq 0, \mu^1 \geq 0, \mu^2 \geq 0 \end{array} \right\}$$

and $\mathcal{M}_{D(\lambda, \mu^1, \mu^2)}^*$ denotes the projection of \mathcal{M}_D^* along (λ, μ^1, μ^2) . Eliminating the variables λ, μ^1, μ^2 from the above system by the Fourier-Motzkin procedure yields an inequality description

$$M_D^* = \{u \in \mathbb{R}^s : H_0 u \leq h, H_1 u = 0\}$$

which has no redundant rows if we apply the above mentioned Tschernikow rules. Since also $0 \in M_D^*$, we may assume that, after proper scaling, $h_i \in \{0, 1\}$ for all components h_i of h . Now, $M_D = M_D^{**}$, and again Theorem 9.1 in [9] implies

$$M_D = (M_D^*)^* = \text{conv}\left(\{0\} \cup \bigcup_{h_i=1} \{h_{0i}\}\right) + \text{cone}\left(\bigcup_{h_i=0} \{h_{0i}\} \cup \bigcup_i \{h_{1i}\} \cup \bigcup_i \{-h_{1i}\}\right)$$

where h_{0i}, h_{1i} are the rows of H_0, H_1 . This is the desired representation and our proof is complete. ■

Recall that, due to our basic assumptions (4), (5), the polyhedron M_D has vertices. Therefore H_1 in the above proof has to be the zero matrix.

The main algorithmic step in the above proof, the elimination of the variables λ, μ^1, μ^2 from the system in (9), is implemented in the code PORTA [2]. As input, the user has to supply an inequality description of the relevant polyhedron, in our situation $(W, -I)^T u \leq \binom{q}{0}, u \geq 0$. The output of PORTA contains the list of vertices and extreme rays, in our situation the row vectors of H_0 .

In [15] the authors report on numerical experience with the algorithm *support* ([17]) that, although different in appearance, follows similar principles as the Fourier-Motzkin procedure. Our intention with the above proposition is to point out the relation to Fourier-Motzkin elimination and to give an impression on the key procedure implemented in PORTA.

3 Lower Bounds for Integer Recourse Problems

The present section deals with two-stage stochastic programs where the second stage problem is an integer linear program. The basic model is again given by (1) - (3) but now $Y = \mathbb{Z}_+^p$. We also assume that all entries in W are rational numbers, and (3) reads

$$\Phi(t) = \min\{q^T y : Wy \geq t, y \in \mathbb{Z}_+^m\}. \quad (9)$$

This value function Φ is in general non-convex and discontinuous, in fact, lower semicontinuous, and these properties of Φ , obviously, are transferred to Q . Therefore, algorithms for linear recourse problems with continuous variables, essentially resting on convexity of Q , break down for this class of problems. If the underlying probability measure μ is discrete, which we assume also in the present section, then (1) - (3) may be equivalently rewritten as a large-scale mixed-integer linear program with dual block angular structure. It is tempting to tackle this problem by decomposition. This, however, leads to master problems whose objectives are essentially governed by Φ , and, again, we are facing lower semicontinuous objectives ([1]).

In [12], an algorithm for the above integer recourse stochastic program is proposed that combines enumeration of Q with an efficient procedure for computing its function values. The latter employs Gröbner bases methods from computational algebra: Using Buchberger's algorithm, a Gröbner basis of a polynomial ideal related to the integer program in (9) is computed. This basis only depends on the objective and the coefficient matrix of the integer program. For the various right-hand sides, solution of the integer programs then is accomplished by a scheme of generalized division of multivariate polynomials. The latter is much faster than solving anew the integer program with conventional methods each time another right-hand side arises. Computing the Gröbner basis, however, is the bottleneck of the method such that second-stage problems with moderate size can be handled only.

As to the enumeration of Q , bounds restricting the search are most welcome. Here, preprocessing

does an important job which we will explain now. Consider the continuous relaxation

$$\Phi_R(t) = \min\{q^T y : Wy \geq t, y \in \mathbb{R}_+^m\}$$

and the corresponding relaxed expected recourse function

$$Q_R(x) = \int_{\mathbb{R}^s} \Phi_R(z - Ax)\mu(dz).$$

Of course, $Q_R(x) \leq Q(x)$, and therefore any optimal solution to the integer recourse stochastic program belongs to the level set

$$\{x \in C : c^T x + Q_R(x) \leq c^T \bar{x} + Q(\bar{x})\} \quad (10)$$

where $\bar{x} \in C$ is an arbitrary feasible point. The enumeration part of the algorithm in [12] rests on searching the above level sets. Each time a feasible point \bar{x} with improved objective function value is found the level set (10) can be shrunk.

In view of the representation (8), the function Q_R is convex piecewise-linear if the measure μ is discrete. Therefore, all the level sets in (10) are non-empty polyhedra. If we assume that $\text{pos}(W, -I) = \mathbb{R}^s$ then $\Phi_R(t) = \max_{v_i \in V_0} v_i^T t$ where V_0 is the vertex set of M_D . This leads to the following lower bound for Q_R

$$\begin{aligned} Q_R(x) &= \sum_{l=1}^L p_l \Phi_R(z_l - Ax) \\ &\geq \Phi_R\left(\sum_{l=1}^L p_l z_l - Ax\right) \\ &= \max_{v_i \in V_0} v_i^T \left(\sum_{l=1}^L p_l z_l - Ax\right) =: Q_{RL}(x). \end{aligned}$$

Here, the second line is a consequence of Jensen's inequality. If we replace Q_R in (10) by Q_{RL} then again any optimal solution to the integer recourse stochastic program belongs to the respective level set. The advantage over the previous situation is that Q_{RL} is explicitly known via the vertices of M_D . If the latter are obtained by the procedure described in the previous section, then the enumeration part of the algorithm in [12] becomes algorithmically feasible. The disadvantage that Fourier-Motzkin elimination breaks down at large-scale instances is not of great significance in this case, since the above mentioned bottleneck in Gröbner basis computation restricts application of the algorithm in [12] to second-stage problems of moderate size.

As an example let us consider two-stage investment problems with multi-knapsack second stage. In contrast to (1) these are maximization problems which can be formalized as follows.

$$\max\{c^T x + Q(x) : x \in C\} \quad (11)$$

where

$$Q(x) = \int_{\mathbb{R}^s} \Phi(z - Ax)\mu(dz) \quad (12)$$

and

$$\Phi(t) = \max\{q^T y : Wy \leq t, y \in \{0, 1\}^m\} \quad (13)$$

with non-negative components in c, q and A, W .

The first-stage investment decision x in (11) - (13) is selected from some feasible set C and yields an immediate revenue $c^T x$. Further revenue is gained from projects for which investment is done in the second stage after having observed the random vector $z \in \mathbb{R}^s$ leading to the budget $z - Ax$. Spending money in the first stage decreases possibilities in the second stage. However, negative entries in x may be permitted leading to the possibility to contract loans in the first stage to enlarge possibilities in the second stage. The objective in (11) - (13) is to find a first-stage investment decision x such that the sum of direct revenue from the first stage and expected revenue from the second stage is maximal.

Computational experience with solving (11) - (13) is reported in [12]. Here, we concentrate on the preprocessing part, i.e., on finding a representation (7) for the dual polyhedron

$$M_D = \{u \in \mathbb{R}_+^{s+m} : (W^T, I)u \geq q\}. \quad (14)$$

Although the assumption that $pos(W, I) = \mathbb{R}^s$ is not met here, such that Φ is not defined on the whole of \mathbb{R}^s , vertices of M_D , nevertheless, can be used in the above way to bound enumeration. As mentioned in Section 2, the bottleneck of Fourier-Motzkin elimination is that, in general, the size of the iteration system of linear inequalities grows quadratically per elimination of one variable. Moreover, Fourier-Motzkin elimination is an all-or-nothing procedure. The complete list of vertices and extreme rays is generated only in the very last step. If the method breaks down since the iteration system of linear inequalities becomes too big, then no partial list of vertices and extreme rays is available.

Our experience with PORTA on a SPARCstation 20 Model 61 with 160 MB of main memory indicates that, within seconds, vertex sets with up to several hundreds of elements can be enumerated. If there are several hundred thousands of vertices, then there is a pretty high chance that PORTA breaks down due to excessive size of the iteration system. Vertex sets with up to one hundred thousand elements have a good chance to be enumerable, although this might cost several hours of CPU time. To illustrate these statements, we tested PORTA on some instances of the polyhedron (14). For a matrix W with 2 rows and 29 columns the complete list of 241 vertices was found after 2 seconds, for a 2×50 matrix W we ended up with 545 vertices after 14 seconds and for a 4×50 matrix W it took 5550 seconds to find all the 37887 vertices.

4 Verification of Stability Conditions

The purpose of this section is to demonstrate how the preprocessing tool from Section 2 can be used for the verification of assumptions in theoretical considerations on stochastic programs.

In recent years, studies on the stability of the stochastic program (1) - (3) with respect to perturbations of the underlying measure μ have attracted some interest. This is mainly motivated

by the incomplete information on μ that one often faces in applications and by numerical problems that arise in computations of the integral in (2) if μ is multivariate continuous. A crucial issue in stability analysis is that sufficient stability conditions are verifiable from the data in the unperturbed problem, i.e., in our situation, from (1) - (3) with some fixed measure μ . In the following, we illustrate at a result on the stability of optimal solution sets how preprocessing can be employed to extend verifiability of stability conditions.

Let $Y = \mathbb{R}_+^m$ and consider problem (1) - (3) as a parametric program in μ

$$P(\mu) \quad \min\{c^T x + Q_\mu(x) : x \in C\} \quad (15)$$

where

$$Q_\mu(x) = \int_{\mathbb{R}^s} \Phi(z - Ax)\mu(dz) \quad (16)$$

and

$$\Phi(t) = \min\{q^T y : Wy = t, y \in \mathbb{R}_+^m\}. \quad (17)$$

(The only reason for writing the second-stage linear program in equality form is consistency with the settings in [7], [10].)

The following proposition was established in [7]. It provides a Lipschitz estimate for the Hausdorff distance of optimal solution sets $\psi(\mu), \psi(\nu)$ to stochastic programs $P(\mu)$ and $P(\nu)$, respectively. The estimate is in terms of some distance of probability measures $d(\mu, \nu; U)$ that we will not explain here. For details we refer to [7] where it is also shown that $d(\mu, \nu; U)$ can be majorized by the uniform distance of distribution functions of probability measures closely related to μ and ν . The function \tilde{Q}_μ arising in the statement is defined by $\tilde{Q}_\mu(\chi) = \int_{\mathbb{R}^s} \Phi(z - \chi)\mu(dz)$.

Proposition 4.1 *Let pos $W = \mathbb{R}^s, M_D \neq \emptyset$ and $\int_{\mathbb{R}^s} \|z\|\mu(dz) < +\infty$. Suppose further that $\psi(\mu)$ is non-empty and bounded. Assume that there exists a convex open subset V of \mathbb{R}^s such that $A(\psi(\mu)) \subset V$ and the function \tilde{Q}_μ is strongly convex on V . Let $U = \text{cl } U_o$, where U_o is an open, convex, bounded set such that $\psi(\mu) \subset U_o$ and $A(U) \subset V$.*

Then there exist constants $L > 0, \delta > 0$ such that

$$d_H(\psi(\mu), \psi(\nu)) \leq L \cdot d(\mu, \nu; U)$$

for all probability measures ν such that $\int_{\mathbb{R}^s} \|z\|\nu(dz) < +\infty$ and $d(\mu, \nu; U) < \delta$.

As to verification of assumptions, the critical part of the above statement is the strong convexity of \tilde{Q}_μ , which means that there exists some $\kappa > 0$ such that for all $\chi, \chi' \in V$ and all $\lambda \in [0, 1]$

$$\tilde{Q}_\mu(\lambda\chi + (1 - \lambda)\chi') \leq \lambda\tilde{Q}_\mu(\chi) + (1 - \lambda)\tilde{Q}_\mu(\chi') - \kappa\lambda(1 - \lambda)\|\chi - \chi'\|^2.$$

Strong convexity of \tilde{Q}_μ can be verified via the following result from [10].

Proposition 4.2 *Let pos $W = \mathbb{R}^s$, the interior of M_D be non-empty and $\int_{\mathbb{R}^s} \|z\|\mu(dz) < +\infty$. Suppose further that there exist a convex open set $V \subset \mathbb{R}^s$, constants $r > 0, \rho > 0$ as well as a density θ of μ such that $\theta(t') \geq r$ for all $t' \in V_\rho := \{t' \in \mathbb{R}^s : \text{dist}(t', V) \leq \rho\}$. Then \tilde{Q}_μ is strongly convex on V .*

The density assumption in the above proposition restricts application of the result to stochastic programs (15) - (17) where all components of z are random, i.e., not constant almost surely. This is quite restrictive in applications, and [11] analyses models where only a part of $z \in \mathbb{R}^s$, say $z_1 \in \mathbb{R}^{s_1}$, is random, and the remaining part $z_2 \in \mathbb{R}^{s_2}$ ($s_1 + s_2 = s$) is fixed. Then Q_μ becomes

$$Q_\mu(x) = \int_{\mathbb{R}^{s_1}} \Phi^o(z_1 - A_1x, z_2 - A_2x) \mu(dz_1)$$

with

$$\Phi^o(t_1, t_2) = \min\{q^T y : W_1y = t_1, W_2y = t_2, y \in \mathbb{R}_+^m\}. \quad (18)$$

Consider

$$\tilde{Q}_\mu^o(\chi_1, \chi_2) = \int_{\mathbb{R}^{s_1}} \Phi^o(z_1 - \chi_1, z_2 - \chi_2) \mu(dz_1)$$

which, for fixed $\bar{\chi}_2$, is studied in [11] as a function of the first argument χ_1

$$\tilde{Q}_\mu(\chi_1) = \int_{\mathbb{R}^{s_1}} \Phi^o(z_1 - \chi_1, z_2 - \bar{\chi}_2) \mu(dz_1). \quad (19)$$

The central result in [11] is a counterpart to Proposition 4.2 for the above function \tilde{Q}_μ . This leads to a generalization of the Lipschitz estimate in Proposition 4.1 for stochastic programs with partially random right-hand side (see [11] for details). The counterpart to Proposition 4.2 contains a density assumption for z_1 that is more involved than the one for z in Proposition 4.2. Its verification can be essentially supported by preprocessing. Let us explain this in more detail. The more complicated assumption, called generalized density assumption in the following, reads:

There exist a convex open set $V \subset \mathbb{R}^{s_1}$, constants $r > 0, \rho > 0$, points $e_{ij}^ \in \mathcal{F}_{ij}^*$ and a density θ of μ such that $\theta(t') \geq r$ for all $t' \in \cup_{(i,j)}\{e_{ij}^* + V_\rho\}$.*

The sets \mathcal{F}_{ij}^* here denote the unbounded facets arising in the polyhedral complex of lineality regions of the piecewise linear convex function $\Phi(t_1) = \Phi^o(t_1, z_2 - \bar{\chi}_2)$ with Φ^o as in (18). In Proposition 4.2, this complex is especially simple: Lineality regions of Φ coincide with the outer normal cones to the dual polyhedron M_D , which is compact in this case such that the polyhedral complex is a fan of cones with common vertex zero. Therefore, the points e_{ij}^* can all be selected as zero and the condition $t' \in \cup_{(i,j)}\{e_{ij}^* + V_\rho\}$ turns into $t' \in V_\rho$.

To see how the complex looks like in the more general situation we consider the dual polyhedron $M_D^o = \{(u_1, u_2) \in \mathbb{R}^{s_1+s_2} : W_1^T u_1 + W_2^T u_2 \leq q\}$. By duality, we obtain (with $\bar{t}_2 = z_2 - \bar{\chi}_2$)

$$\Phi(t_1) = \max\{t_1^T u_1 + \bar{t}_2^T u_2 : (u_1, u_2) \in M_D^o\}. \quad (20)$$

Suppose that M_D^o has vertices, and let $\tilde{d}_1, \dots, \tilde{d}_N$ be the vertices of M_D^o that arise as optimal ones in (20) when t_1 varies in \mathbb{R}^{s_1} . Denoting by $\tilde{d}_{i1}, \tilde{d}_{i2}$ ($i = 1, \dots, N$) the projections of \tilde{d}_i on \mathbb{R}^{s_1} and \mathbb{R}^{s_2} , respectively, we obtain

$$\Phi(t_1) = t_1^T \tilde{d}_{i1} + \bar{t}_2^T \tilde{d}_{i2}$$

for all $t_1 \in \mathbb{R}^{s_1}$ such that (t_1, \bar{t}_2) belongs to the outer normal cone \mathcal{K}_i to M_D^o at the vertex \tilde{d}_i . The lineality regions \mathcal{K}_i^* of Φ then are given by

$$\mathcal{K}_i^* = \pi_I(\mathcal{K}_i \cap \{\mathbb{R}^{s_1} \times \{\bar{t}_2\}\}), \quad i \in \{1, \dots, N\} \quad (21)$$

where π_I denotes the projection from \mathbb{R}^s to \mathbb{R}^{s_1} . They form a polyhedral complex that arises by intersecting the fan of outer normal cones to M_D^o with the affine subspace $\mathbb{R}^{s_1} \times \{\bar{t}_2\}$. The sets \mathcal{F}_{ij}^* are the unbounded facets ($(s_1 - 1)$ -dimensional faces) in this complex.

Some insight into the polyhedral complex in (21) can be gained by preprocessing. Let us illustrate this at the following example treated in more detail in [11].

Example: In (18), (19) we put

$$\begin{aligned} q^T &= (21, 21, 21, 21, 7, 7, 3, 3, 1, 0) \in \mathbb{R}^{10}, \\ W_1 &= \begin{pmatrix} -3 & -3 & -3 & 3 & -1 & 1 & 0 & 0 & 0 & 0 \\ -5 & 1 & 2 & -2 & -2 & 2 & -1 & 1 & 0 & 0 \end{pmatrix} \in L(\mathbb{R}^{10}, \mathbb{R}^2), \\ W_2 &= \begin{pmatrix} 12 & 12 & 9 & 9 & 3 & 3 & 0 & 0 & 1 & -1 \end{pmatrix} \in L(\mathbb{R}^{10}, \mathbb{R}^1), \\ z_2 &= 1, \quad \bar{\chi}_2 = 0. \end{aligned}$$

Using PORTA [2], we computed the vertices of M_D^o together with a vertex-inequality incidence table displaying the binding inequalities for each vertex. Gradients of the binding (linear) inequalities then generate the respective outer normal cones:

$$\begin{aligned} \tilde{d}_1 &= (-7, 0, 0), \\ \mathcal{K}_1 &= \text{cone} \{(0, 0, -1), (-1, -2, 3), (-3, 2, 9), (-3, 1, 12), (-3, -5, 12)\}, \\ \tilde{d}_2 &= (-1, -3, 0), \\ \mathcal{K}_2 &= \text{cone} \{(0, 0, -1), (0, -1, 0), (-1, -2, 3)\}, \\ \tilde{d}_3 &= (5, -3, 0), \\ \mathcal{K}_3 &= \text{cone} \{(0, 0, -1), (0, -1, 0), (3, -2, 9)\}, \\ \tilde{d}_4 &= (7, 0, 0), \\ \mathcal{K}_4 &= \text{cone} \{(0, 0, -1), (1, 2, 3), (3, -2, 9)\}, \\ \tilde{d}_5 &= (1, 3, 0), \\ \mathcal{K}_5 &= \text{cone} \{(0, 0, -1), (1, 2, 3), (0, 1, 0)\}, \\ \tilde{d}_6 &= (-5, 3, 0), \\ \mathcal{K}_6 &= \text{cone} \{(0, 0, -1), (0, 1, 0), (-3, 2, 9)\}, \end{aligned}$$

$$\begin{aligned}
\tilde{d}_7 &= (-3, 0, 1), \\
\mathcal{K}_7 &= \text{cone} \{(0, 0, 1), (-3, 1, 12), (-3, -5, 12)\}, \\
\tilde{d}_8 &= (2, -3, 1), \\
\mathcal{K}_8 &= \text{cone} \{(0, 0, 1), (0, -1, 0), (-1, -2, 3), (3, -2, 9), (-3, -5, 12)\}, \\
\tilde{d}_9 &= (4, 0, 1), \\
\mathcal{K}_9 &= \text{cone} \{(0, 0, 1), (1, 2, 3), (3, -2, 9)\}, \\
\tilde{d}_{10} &= (-2, 3, 1), \\
\mathcal{K}_{10} &= \text{cone} \{(0, 0, 1), (1, 2, 3), (0, 1, 0), (-3, 2, 9), (-3, 1, 12)\}.
\end{aligned}$$

One possibility to fulfill the generalized density assumption is to select \tilde{e}_j as *vertices* of unbounded facets \mathcal{F}_{ij}^* . Here, the above list of generators can be helpful since it allows to compute *all* vertices in the complex (21). Indeed, each intersection of a positive multiple of some generator with the affine subspace $\mathbb{R}^2 \times \{1\}$ yields a vertex and each vertex has to arise as such an intersection. Hence, there is a one-to-one correspondence between the vertices and the generators with positive third component. In this way, we obtain a list of 7 vertices of which 3 do not belong to unbounded facets. For more complicated examples, the latter extraction may be non-trivial. In such cases the generalized density assumption can be fulfilled by using all vertices of (21) instead of the points e_{ij}^* .

Another possibility to fulfill the generalized density assumption is to claim that $\theta(t) \geq r$ for all $t \in \mathcal{B} + V_\rho$ where \mathcal{B} is a bounded set containing all the vertices from (21). To this end, we compute an upper bound for the norm of these vertices. Again preprocessing is helpful. The outer normal cone \mathcal{K}_i to M_D^o at \tilde{d}_i can be written as

$$\mathcal{K}_i = \{u \in \mathbb{R}^s : u = W(i)v, v \geq 0\}$$

where $W(i) \in L(\mathbb{R}^{m_i}, \mathbb{R}^s)$ is given by the generators of \mathcal{K}_i computed above. Then it holds that

$$\begin{aligned}
\mathcal{K}_i^* &= \{u_1 \in \mathbb{R}^{s_1} : u_1 = W(i)_1 v, \bar{t}_2 = W(i)_2 v, v \geq 0\} \\
&= W(i)_1(\{v \in \mathbb{R}^{m_i} : W(i)_2 v = \bar{t}_2, v \geq 0\}).
\end{aligned}$$

hence, the vertices of \mathcal{K}_i^* are among the $W(i)_1$ -images of the vertices $v(i)_j$ ($j = 1, \dots, J_i$) of $\{v \in \mathbb{R}^{m_i} : W(i)_2 v = \bar{t}_2, v \geq 0\}$. The latter can be computed using PORTA [2]. For the desired upper bound all possible basis submatrices of all $W(i)_2$ have to be extracted which, together with \bar{t}_2 and submatrices of the $W(i)_1$, yields representations for the vertices of \mathcal{K}_i^* . These are bounded above by the usual estimates using matrix norms (see [11] for details).

It is evident that the procedures discussed in this section are not suitable for large-scale problems. However, they can serve well for improving our theoretical understanding for problems with moderate size.

5 Conclusions

Fourier-Motzkin elimination provides an elegant way to look at preprocessing in stochastic programming. Moreover there is a well tested code, PORTA [2], that is based on Fourier-Motzkin elimination, such that, in addition to the codes reported in [15], [16], another convenient computer tool for preprocessing in stochastic programming is available. It is well known that Fourier-Motzkin elimination is of exponential complexity such that application of the method has to be restricted to problems of moderate size. In the present paper we discussed two applications with natural size limitation. For stochastic programs with integer recourse, preprocessing is helpful for restricting the search in enumeration algorithms. In the stability analysis of stochastic programs, preprocessing tools can be used to widen the class of problems for which sufficient stability conditions can be verified.

Acknowledgement:

I wish to thank Thomas Christof (Universität Heidelberg) and Andreas Löbel (Konrad-Zuse-Zentrum für Informationstechnik Berlin) as well as Stein W. Wallace (Norwegian Institute of Technology Trondheim) and Roger J-B Wets (University of California, Davis) for beneficial discussions.

References

- [1] Carøe, C.C., Tind, J.: L-shaped decomposition of two-stage stochastic programs with integer recourse, Technical Report, Institute of Mathematics, University of Copenhagen, 1995.
- [2] Christof, T.: PORTA – A Polyhedron Representation Transformation Algorithm, version 1.2.1., 1994, written by T. Christof (Univ. Heidelberg), revised by A. Löbel and M. Stoer, available from the ZIB electronic library eLib by anonymous ftp from `elib.zib.de`, in the directory `/pub/mathprog/polyth`.
- [3] Duffin, R.J.: On Fourier's analysis of linear inequality systems, Math. Programming Study 1, American Elsevier Publishing Company, New York, 1974.
- [4] Higle, J.L., Sen, S.: Stochastic decomposition: an algorithm for two-stage linear programs with recourse, *Mathematics of Operations Research* 16 (1991) 650–669.
- [5] Kall, P., Wallace, S.W.: *Stochastic Programming*, Wiley, Chichester, 1994.
- [6] Prékopa, A.: *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, 1995.

- [7] Römisch, W., Schultz, R.: Lipschitz stability for stochastic programs with complete recourse, *SIAM Journal on Optimization* 6 (1996) 531–547.
- [8] Ruszczyński, A.: A regularized decomposition method for minimizing a sum of polyhedral functions, *Mathematical Programming* 35 (1986) 309–333.
- [9] Schrijver, A.: *Theory of Linear and Integer Programming*, Wiley, Chichester, 1986.
- [10] Schultz, R.: Strong convexity in stochastic programs with complete recourse, *Journal of Computational and Applied Mathematics* 56 (1994) 3–22.
- [11] Schultz, R.: Strong convexity in stochastic programs with complete recourse II: partially random right-hand side, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Preprint SC 95-21, 1995.
- [12] Schultz, R., Stougie, L., van der Vlerk, M.H.: Solving stochastic integer programs with complete recourse: a framework using Gröbner bases, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Preprint SC 95-23, 1995, revised 1996.
- [13] Tschernikow, S.N.: *Lineare Ungleichungen*, Deutscher Verlag der Wissenschaften, Berlin, 1971.
- [14] van Slyke, R., Wets, R.J-B: L-shaped linear programs with applications to optimal control and stochastic programming, *SIAM Journal of Applied Mathematics* 17 (1969), 638–663.
- [15] Wallace, S.W., Wets, R.J-B: Preprocessing in stochastic programming: the case of linear programs, *ORSA Journal on Computing* 4 (1992) 45–59.
- [16] Wallace, S.W., Wets, R.J-B: Preprocessing in stochastic programming: the case of capacitated networks, *ORSA Journal on Computing* 7 (1995) 44–62.
- [17] Wets, R.J-B: Elementary, constructive proofs of the Theorem of Farkas, Minkowski and Weyl, in: J. Gabszewicz, J.F. Richard, L.A. Wolsey (eds.) "Economic Decision Making: Games, Econometrics and Optimization: Contributions in Honour of Jacques Drèze, North Holland, Elsevier Science, Amsterdam, 427–432.
- [18] Ziegler, G.M.: *Lectures on Polytopes*, Springer - Verlag, New York, 1995.