Workshop on Integer Programming and Continuous Optimization

Chemnitz, November 7-9, 2004

SIP from NLP Perspective

Rüdiger Schultz

(Universität Duisburg-Essen)

SIP Objects of Desire

Random mixed-integer optimization problem:

$$\min\left\{c^\top x + q^\top y \ : \ Tx + Wy = z(\omega), \ \ x \in X, \ y \in Y\right\}$$

$(X, Y$ - mixed-integer, polyhedral)

with information constraints (non-anticipativity):

$$\text{decide } x \ \mapsto \ \text{observe } z(\omega) \ \mapsto \ \text{decide } y = y(x, z(\omega))$$

$$\begin{aligned}
&= \quad \min_x \left\{c^\top x + \min_y\{q^\top y \ : \ Wy = z(\omega) - Tx, \ y \in Y\} \ : \ x \in X\right\} \\
&= \quad \min\left\{c^\top x + \Phi(z(\omega) - Tx) \ : \ x \in X\right\} \\
&= \quad \min\left\{f(x, z(\omega)) \ : \ x \in X\right\}
\end{aligned}$$

here

$$\Phi(t) := \min\{q^\top y \ : \ Wy = t, \ y \in Y\}$$

mixed-integer value function

How to find “best” element in family

$$\{f(x, z(\omega)) \, : \, x \in X\}$$

of random variables ?

Answer: Mean-Risk Model

$$\min \left\{ \mathbb{E}_z f(x,z) + \rho \cdot \mathcal{R}_z f(x,z) \, : \, x \in X \right\} \qquad (\rho \geq 0)$$

This creates a “whole zoo” of NLPs.

Specification of Risk - Deviation Based

Variance:

$$\mathcal{R}_V f(x,z) := \mathbb{E}\left(f(x,z) - \mathbb{E} f(x,z)\right)^2$$

Central Deviation:

$$\mathcal{R}_{CD} f(x,z) := \mathbb{E}\left|f(x,z) - \mathbb{E} f(x,z)\right|$$

Semideviation:

$$\mathcal{R}_{SD} f(x,z) := \mathbb{E} \max\{f(x,z) - \mathbb{E} f(x,z), 0\}$$

Expected Excess of Target $\eta \in \mathbb{R}$:

$$\mathcal{R}_{EE} f(x,z) := \mathbb{E} \max\{f(x,z) - \eta, 0\}$$

Specification of Risk - Quantile Based

Excess Probability of Target $\eta \in I\!R$:

$$\mathcal{R}_{EP} f(x,z) \;:=\; I\!P\left(f(x,z) > \eta\right)$$

α-Value-at-Risk (αVaR):

$$\mathcal{R}_{VaR} f(x,z) \;:=\; \min\left\{\eta \;:\; I\!P\left(f(x,z) \leq \eta\right) \;\geq\; \alpha\right\} \quad (=: \eta_\alpha(x))$$

(smallest of $(1-\alpha)100\%$ worst outcomes)

α-Conditional-Value-at-Risk (αCVaR):

$$\mathcal{R}_{CVaR} f(x,z) \;:=\; I\!E\left(f(x,z) \,|\, f(x,z) \geq \eta_\alpha(x)\right)$$

(expectation of $(1-\alpha)100\%$ worst outcomes)
(definition needs modification for discretely distributed $f(x,z)$)

Integers in y:

$$\Phi(t) = \min\{q^\top y + q'^\top y' \,:\, Wy + W'y' = t,\, y \in \mathbb{Z}_+^{\bar{m}},\, y' \in \mathbb{R}_+^{m'}\}$$

Basic assumptions:

(A1) complete recourse:
$$W(\mathbb{Z}_+^{\bar{m}}) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s,$$

(A2) sufficiently expensive recourse:
$$\{u \in \mathbb{R}^s \,:\, W^T u \le q,\, W'^T u \le q'\} \neq \emptyset,$$

(A3) finite first moment:
$$\mathbb{E}_\mu \|z\| \;:=\; \int_{\mathbb{R}^s} \|z\| \mu(dz) \;<\; +\infty.$$

Proposition [Blair/Jeroslow 1977, Bank/Mandel 1988]:

Assume (A1) , (A2). Then it holds

(i) Φ is real-valued and lower semicontinuous on $\mathbb{R}^s$,

(ii) there exists a countable partition $\mathbb{R}^s = \cup_{i=1}^\infty \mathcal{T}_i$ such that the restrictions of Φ to $\mathcal{T}_i$ are piecewise linear and Lipschitz continuous with a uniform constant $L > 0$ not depending on i,

(iii) each of the sets $\mathcal{T}_i$ has a representation
$$\mathcal{T}_i = \{t_i + \mathcal{K}\} \setminus \cup_{j=1}^N \{t_{ij} + \mathcal{K}\}$$
where $\mathcal{K}$ denotes the polyhedral cone $W'(\mathbb{R}_+^{m'})$ and t_i, t_{ij} are suitable points from $\mathbb{R}^s$, moreover, N does not depend on i,

(iv) there exist positive constants β, γ such that
$$|\Phi(t_1) - \Phi(t_2)| \le \beta \|t_1 - t_2\| + \gamma$$
whenever $t_1, t_2 \in \mathbb{R}^s$.

Analytical Properties - Convexity

Departure point: Without integers (!!), $f(.,z)$ is convex.

Mean-risk models preserving convexity:

central deviation (for $0 \leq \rho \leq 1/2$):

$$\mathbb{E}f(x,z) + \rho \cdot \mathbb{E}|f(x,z) - \mathbb{E}f(x,z)|$$

$$= (1-2\rho) \cdot \mathbb{E}f(x,z) + 2\rho \cdot \mathbb{E}\max\{f(x,z), \mathbb{E}f(x,z)\}$$

semideviation (for $0 \leq \rho \leq 1$):

$$\mathbb{E}f(x,z) + \rho \cdot \mathbb{E}\max\{f(x,z) - \mathbb{E}f(x,z), 0\}$$

$$= (1-\rho)\mathbb{E}f(x,z) + \rho\mathbb{E}\max\{f(x,z), \mathbb{E}f(x,z)\}$$

expected excess (for $\eta \in \mathbb{R}$ and $\rho \geq 0$):

$$\mathbb{E}f(x,z) + \rho \cdot \mathbb{E}\max\{f(x,z) - \eta, 0\}$$

conditional value-at-risk:

$$\mathcal{R}_{CVaR}f(x,z) = \min\left\{\eta + \frac{1}{1-\alpha}\mathbb{E}\max\{f(x,z) - \eta, 0\} : \eta \in \mathbb{R}\right\}$$

Minimizing a jointly convex function with respect to one variable gives a function that is convex in the other variable.

Analytical Properties - Lipschitz Continuity

Deterring Result:

Proposition:

Suppose that

- q, q', W, W' all have rational entries,
- (A1)-(A3) hold,
- μ has a density,
- for any nonsingular linear transformation $B \in L(\mathbb{R}^s, \mathbb{R}^s)$ all one-dimensional marginal distributions of $\mu \circ B$ have bounded densities which, outside some bounded interval, are monotonically decreasing with growing absolute value of the argument.

Then $\mathbb{E} f(x, z)$ is Lipschitz continuous on bounded sets.

Remark:

Last assumption indispensable. Counterexamples exist.

Analytical Properties - Lower Semicontinuity

Typical Result:

Proposition:

Assume (A1)-(A3). Then $I\!\!E f(.,z) + \rho \cdot \mathcal{R}_{SD} f(.,z)$, with $0 \leq \rho \leq 1$, is lower semicontinuous on $I\!\!R^m$.

Remark:

Result invalid for $\mathcal{R}_V$ (variance), leading to ill-posed mean-risk problems (infimum finite, but not attained).

Analytical Properties - Continuity

Typical Result:

Proposition:

Assume (A1)-(A3) and that $\mu(E(x)) = 0$ where

$$E(x) = \{z \in \mathbb{R}^s : \Phi \text{ is discontinuous at } z - Tx\}.$$

Then $\mathbb{E}f(.,z) + \rho \cdot \mathcal{R}_{SD}f(.,z)$, with $0 \leq \rho \leq 1$, is continuous at x.

Remark:

Discontinuities of Φ contained in countable union of hyperplanes.
Result thus valid if μ has a density.

Analytical Properties - Joint Continuity and Stability

Parametric Optimization Problem:

$$P(\mu) \qquad \min\left\{ \mathbb{E}^{\mu} f(x,z) + \rho \cdot \mathcal{R}^{\mu} f(x,z) \ : \ x \in X \right\}$$

Denote:

$$Q(x,\mu) \ := \ \mathbb{E}^{\mu} f(x,z) + \rho \cdot \mathcal{R}^{\mu} f(x,z)$$

Parameter Space:

$\mathcal{P}(\mathbb{R}^s)$ - set of all Borel probability measures on $\mathbb{R}^s$, equipped with weak convergence of probability measures.

Strengthened (uniform) integrability:

$$\Delta_{p,K}(\mathbb{R}^s) \ := \ \{\nu \in \mathcal{P}(\mathbb{R}^s) \ : \ \int_{\mathbb{R}^s} \|z\|^p \, \nu(dz) \ \leq \ K\}$$

where $p > 1$ and $K > 0$ are fixed constants.

Proposition:

Assume (A1), (A2). Let $\mu \in \Delta_{p,K}(\mathbb{R}^s)$ for some $p > 1$ and $K > 0$, and $\mu(E(x)) = 0$.
Then, with $\mathcal{R} := \mathcal{R}_{SD}$, the function $Q : \mathbb{R}^m \times \Delta_{p,K}(\mathbb{R}^s) \longrightarrow \mathbb{R}$ is continuous at (x,μ).

Remark:

This induces (Berge) stability of the parametric program $P(\mu)$ and, among others, justifies approximation of μ by simpler measures, e.g., discrete ones.

Algorithms

Non-convex global optimization problem:

$$\min \Big\{ Q(x) := \mathbb{E}_z f(x,z) + \rho \cdot \mathcal{R}_z f(x,z) \;\; : \;\; x \in X \Big\}$$

Assume that μ is discrete and finite !

<u>Branch-and-Bound:</u>

Upper Bounding:

- just function evaluation, although somehow "guided" by lower bounds,
- no descent part, yet.

Lower Bounding:

- "expanded" problem formulation, with explicit y-variables,
- yields large-scale, block-structured MILP,
- depending on risk measure, block structure is decomposable or not,
- decomposable case: Lagrangean relaxation of nonanticipativity leads to single-scenario subproblems,
- non-decomposable case: identify decomposable bounds better than just $\mathbb{E}_z f(x,z)$.

Equivalent MILPs - Expectation Problem

μ discrete with realizations z_j and probabilities $\pi_j, j = 1, \ldots, J$

$$\min \left\{ Q_{\mathbb{E}}(x, \mu) \ : \ x \in X \right\}$$

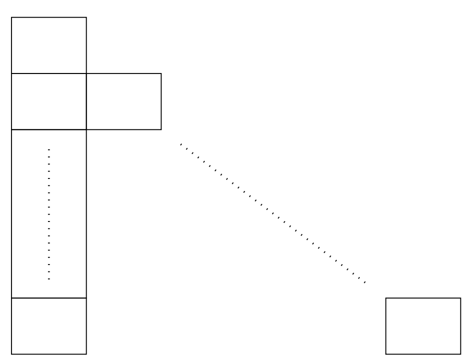
$$= \min \left\{ \mathbb{E}_z[c^\top x + \Phi(z - Tx)] \ : \ x \in X \right\}$$

$$= \min \left\{ c^\top x + \mathbb{E}_z[\Phi(z - Tx)] \ : \ x \in X \right\}$$

$$= \min_x \left\{ c^\top x + \mathbb{E}_z[\min_y \{q^\top y \ : \ Wy = z - Tx, \ y \in Y\}] \ : \ x \in X \right\}$$

$$= \min_x \left\{ c^\top x + \mathbb{E}_z[\min_y \{q^\top y \ : \ Tx + Wy = z, \ y \in Y\}] \ : \ x \in X \right\}$$

$$= \min_{x, y_j} \{ c^T x \quad + \quad \sum_{j=1}^{J} \pi_j q^T y_j \ : \quad Tx + Wy_j \ = \ z_j, \quad x \in X, \ y_j \in Y, \ j = 1, \ldots, J \}$$



Equivalent MILPs - Expected Excess

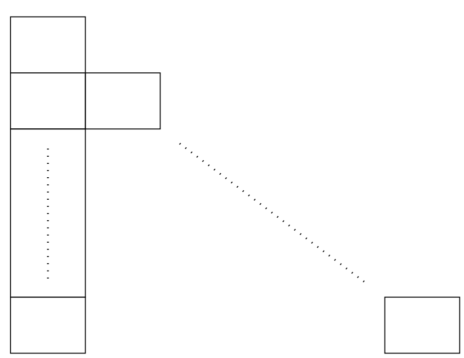
$$Q_{\mathbb{E}}(x,\mu) + \rho \cdot Q_{\mathcal{D}^\eta}(x,\mu)$$

$$= \mathbb{E}_z f(x,z) + \rho \mathbb{E}_z \max\{f(x,z) - \eta, 0\}$$

$$= \mathbb{E}_z[c^\top x + \Phi(z - Tx)] \; + \; \rho \mathbb{E}_z \max\{c^\top x + \Phi(z - Tx) - \eta, 0\}$$

Equivalent minimization problem:

$$\min \Big\{ c^\top x \quad + \quad \sum_{j=1}^{J} \pi_j q^\top y_j + \rho \cdot \sum_{j=1}^{J} \pi_j v_j \; : \\ Tx + W y_j \; = \; z_j, \\ c^\top x + q^\top y_j - \eta \; \leq \; v_j, \\ x \in X, \; y_j \in Y, \; v_j \in \mathbb{R}_+, \; j = 1, \ldots, J \Big\}$$



Equivalent MILPs - Semideviation

$$
\begin{aligned}
&Q_{\mathbb{E}}(x,\mu) + \rho Q_{\mathcal{D}^+}(x,\mu) \\
&\quad = \mathbb{E}_z f(x,z) + \rho \mathbb{E}_z \max\{f(x,z) - \mathbb{E}_z f(x,z), 0\} \\
&\quad = \mathbb{E}_z f(x,z) + \rho\Big(\mathbb{E}_z \max\{f(x,z), \mathbb{E}_z f(x,z)\} - \mathbb{E}_z f(x,z)\Big) \\
&\quad = (1-\rho)\mathbb{E}_z f(x,z) \ + \ \rho \mathbb{E}_z \max\{f(x,z), \mathbb{E}_z f(x,z)\} \\
&\quad = (1-\rho)\mathbb{E}_z[c^\top x + \Phi(z - Tx)] \\
&\qquad\qquad + \rho \mathbb{E}_z \max\Big\{c^\top x + \Phi(z-Tx),\ \mathbb{E}_z[c^\top x + \Phi(z - Tx)]\Big\}
\end{aligned}
$$

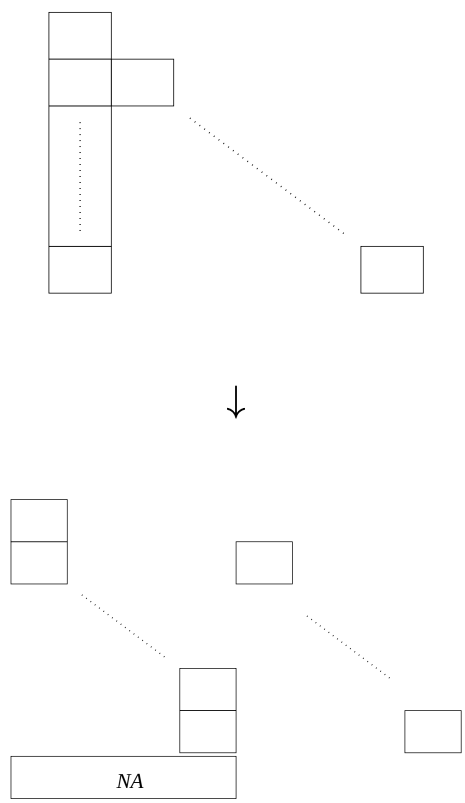
Equivalent minimization problem:

$$
\begin{aligned}
\min\Big\{(1-\rho)c^\top x \ + \ (1-\rho)\sum_{j=1}^{J} \pi_j q^\top y_j + \rho \cdot \sum_{j=1}^{J} \pi_j v_j \ :& \\
Tx + Wy_j \ &= \ z_j, \\
c^\top x + q^\top y_j \ &\leq \ v_j, \\
\boxed{c^\top x + \sum_{i=1}^{J} \pi_i q^\top y_i \ \leq \ v_j,}& \\
x \in X,\ y_j \in Y,\ v_j \in \mathbb{R},\ j = 1,\ldots,J\Big\}&
\end{aligned}
$$

Lower Bounding I:

Relaxation of Nonanticipativity for Expected-Excess Model

Problem reformulation with explicit nonanticipativity ($x_1 = \ldots = x_J$)



Lagrangian function and Lagrangian dual:

$$L(x,y,v,\lambda) \ := \ \sum_{j=1}^{J} L_j(x_j,y_j,v_j,\lambda)$$

with

$$L_j(x_j,y_j,v_j,\lambda) := \pi_j(c^\top x_j + q^\top y_j + \rho v_j) \ + \ \lambda^\top H_j x_j, \ \ j = 1,\ldots,J,$$

and

$$\max\{\sum_{j=1}^{J} D_j(\lambda) \ : \ \lambda \in I\!R^l\}$$

with

$$\begin{aligned} D_j(\lambda) = \min\{L_j(x_j,y_j,v_j,\lambda) \quad : \quad & Tx_j + Wy_j \ = \ z_j, \\ & c^\top x_j + q^\top y_j - \eta \ \leq \ v_j, \\ & x_j \in X, \, y_j \in Y, \, v_j \in I\!R_+\}. \end{aligned}$$

Advantages:

- $D_j(\lambda)$ given by scenario-specific MILP $\mapsto$ **decomposition !**
- powerful algorithms and codes for solving Lagrangian dual and scenario-specific MILPs (ILOG-CPLEX, CONIC BUNDLE)

Lower Bounding II:

Separable Minorants for Semideviation Model

Problem reformulation with explicit NA possible, but constraints

$$\boxed{c^\top x_j + \sum_{i=1}^{J} \pi_i q^\top y_i \;\leq\; v_j,\; j = 1, \ldots, J}$$

prevent separability after relaxation of NA.

Question: Separable lower bounds for objectives ?

Answers:

- trivial bound $Q_{\mathbb{E}}(x, \mu)$,
- improvement by next lemma:

Lemma:

Fix $x \in X$, let $\eta \leq Q_{\mathbb{E}}(x, \mu)$ and $0 \leq \alpha \leq 1$. Then

$$\begin{aligned} Q_{\mathbb{E}}(x, \mu) &\leq \boxed{(1 - \rho) Q_{\mathbb{E}}(x, \mu) + \rho Q_{\mathcal{D}^\eta}(x, \mu) + \rho \eta} \\ &\leq Q_{\mathbb{E}}(x, \mu) + \rho Q_{\mathcal{D}^+}(x, \mu). \end{aligned}$$

Remarks:

- Wait-and-see solution $\mathbb{E}\Phi_{WS}(z)$ with

$$\Phi_{WS}(z) \; := \; \min\left\{c^\top x + q^\top y \; : \; Tx + Wy = z, \; x \in X, \; y \in Y\right\}$$

  provides feasible choice for η in the above lemma.

- Lower bound is strictly tighter than $Q_{\mathbb{E}}(x,\mu)$ if

$$\mu\left\{z \in \mathbb{R}^s \; : \; \mathbb{E}\Phi_{WS}(z) > f(x,z)\right\} \; > \; 0.$$

Computational impact of improved bound:

Semideviation extension of real-life expectation model from chemical engineering.

- first stage: m=24 variables, all integer or binary, together with 3 constraints,
- second stage: $\bar{m}$=108 integer or binary and m'=224 continuous variables, together with 311 constraints,
- J=10 scenarios,
- 4 hours of cpu time, Sun V880 with 880 MHz processor and 4 GB of main memory,

- gaps in %,
- CPLEX: direct application of ILOG-CPLEX 8.1 to full equivalent MILP,
- B&B/EXP: our branch-and-bound algorithm with lower bounds by $Q_{\mathbb{E}}$,
- B&B/ENH: our branch-and-bound algorithm with lower bounds enhanced by lemma.

Instance	CPLEX	B&B/ENH	B&B/EXP
1	86.40	3.01	5.05
2	94.30	16.16	47.41
3	57.80	4.02	6.94
4	10.99	4.31	4.43
5	89.26	7.49	20.86
6	8.73	4.46	7.54
7	6.06	3.62	7.41
8	5.31	5.34	8.64
9	5.34	1.18	5.45
10	97.03	3.87	6.79