9th International Conference on Stochastic Programming

Berlin, August 25-31, 2001

Stochastic Integer Programming: A Tutorial

Rüdiger Schultz

(Gerhard-Mercator University Duisburg)

Stochastic integer programs are mixed-integer (linear) programs.

It always makes sense to check whether the model context allows for algorithmic shortcuts.

Without integers we may borrow from convex optimization, with integers there is no comfortable creditor so far.

Survey Article

Willem K. Klein Haneveld and Maarten H. van der Vlerk:

Stochastic integer programming: General models and algorithms, Annals of Operations Research 85 (1999), 39-57.

Issues to come:

1. Reasons for failure of traditional stochastic programming techniques

2. MILP techniques in stochastic integer programming:
   - Lagrangian Relaxation
   - Primal Methods - Test Sets (Outlook)
   - Dual Methods - Cutting Planes (Suvrajeet Sen)

3. Multistage Problems (Outlook)

Random Mixed-Integer Linear Program

$$\min\{cx + qy \ : \ Tx + Wy \geq h(\omega), \ Ax \leq b, \ x \in X, y \in Y\}$$

Information Constraints:

decision on x

$\downarrow$

observation of $h(\omega)$

$\downarrow$

decision on y

Basic Idea in Stochastic Programming with Recourse:

Find $x \in X$ such that direct costs cx plus expected costs from $h(\omega)$ and y are minimal.

$$\min_x\{cx \quad + \quad \min_y\{qy \ : \ Wy \geq h(\omega) - Tx, y \in Y\} \ : \ Ax \leq b, \ x \in X\}$$

$$\min_x\{cx \quad + \quad E_\omega(\min_y\{qy \ : \ Wy \geq h(\omega) - Tx, y \in Y\}) \ : \ Ax \leq b, \ x \in X\}$$

Stochastic Program with Mixed-Integer Recourse

$$\min\{cx + Q(x) \ : \ Ax \leq b, \ x \in X\}$$

where

$$Q(x) := \int_{\mathbb{R}^s} \Phi(z - Tx)\mu(dz)$$

and

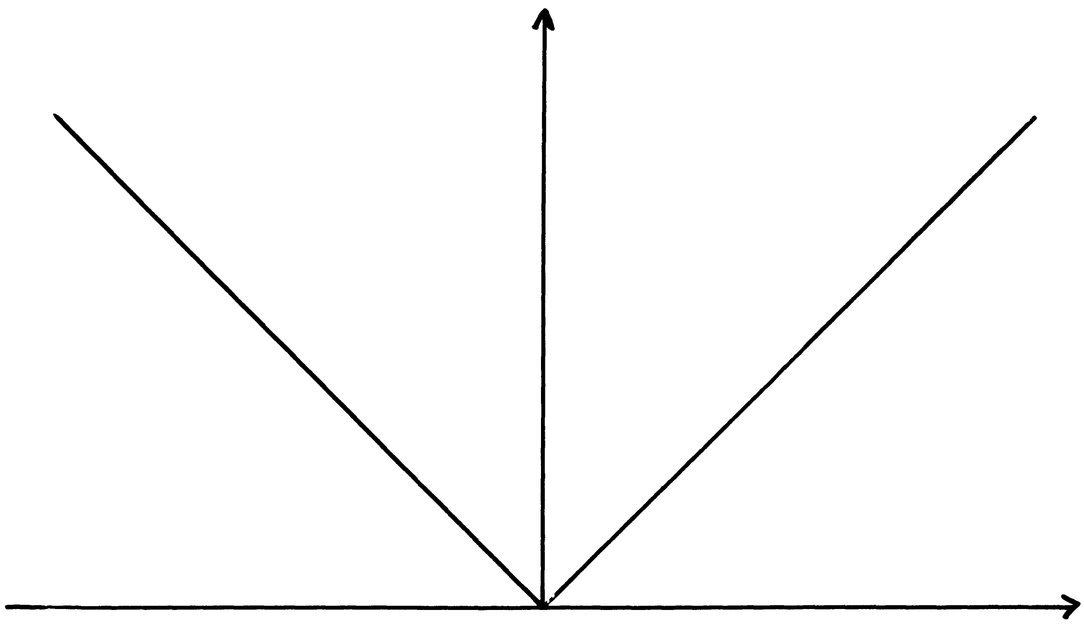
$$\Phi(t) := \min\{qy \ : \ Wy \geq t, \ y \in Y\}.$$

Topics of Interest:

- Structure of Q,
- Approximations and Stability,
- Algorithms !

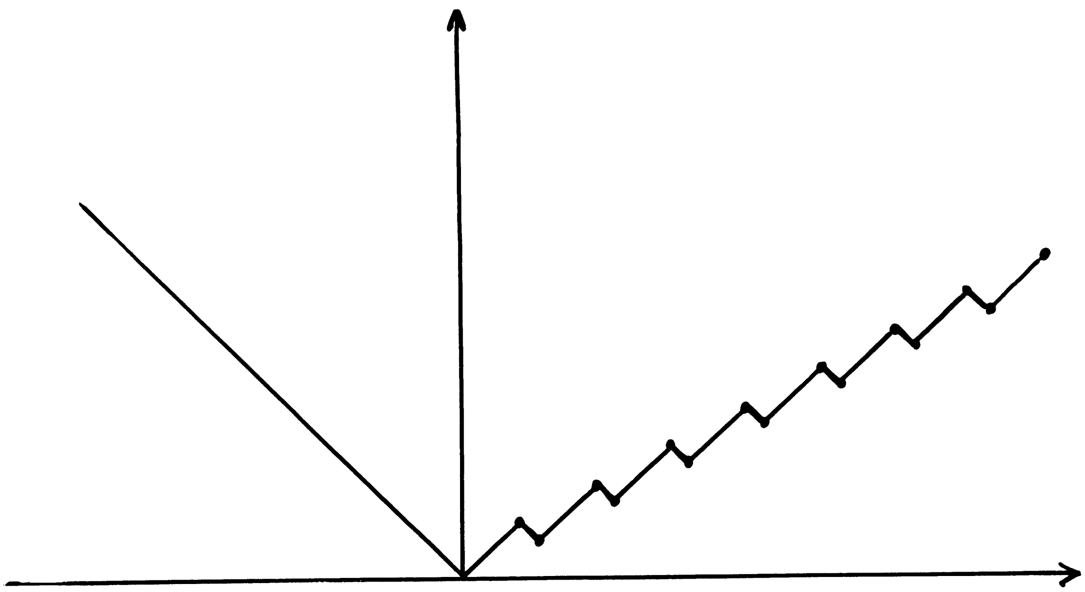
$$\Phi(t) = \min\{\, y^+ + y^- \;:\; y^+ - y^- = t\,,\; y^+ \in \mathbb{R}_+,\; y^- \in \mathbb{R}_+\}$$

$$= \max\{\, t \cdot u \;:\; -1 \leq u \leq 1\} = |t|$$

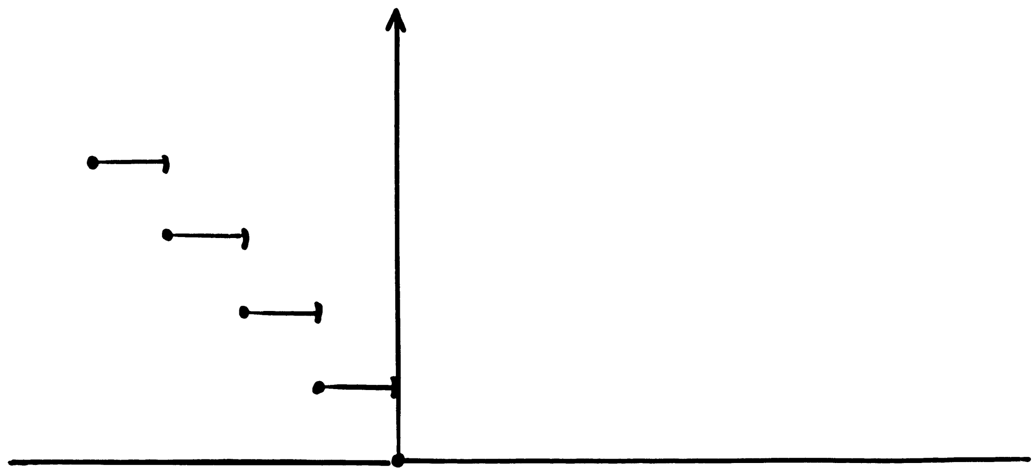


$$\Phi(t) = \min\{\frac{1}{2}v + y^+ + y^- : v + y^+ - y^- = t, \quad y^+ \in \mathbb{R}_+, y^- \in \mathbb{R}_+, v \in \mathbb{Z}_+\}$$

$$= \min\{\frac{1}{2}v + |t - v| : v \in \mathbb{Z}_+\}$$



$$\begin{aligned}\Phi(t) &= \min\{\, v^+ + v^- \;:\; y + v^+ - v^- = t, \\ &\qquad\qquad y \in \mathbb{R}_+,\ v^+ \in \mathbb{Z}_+,\ v^- \in \mathbb{Z}_+\} \\ &= \min\{\, v^+ + v^- \;:\; v^+ - v^- \le t\,,\ v^+ \in \mathbb{Z}_+, v^- \in \mathbb{Z}_+\} \\ &= \begin{cases} 0 & : \quad t \ge 0 \\ \lceil -t \rceil & : \quad t < 0 \end{cases}\end{aligned}$$



Properties of the Value Function

Proposition (Bank/Mandel, Blair/Jeroslow):

Assume that

$$W(\mathbb{Z}_+^{\bar{m}}) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s \text{ and } \{u \in \mathbb{R}^s \; : \; W^T u \leq q, \; W'^T u \leq q'\} \neq \emptyset.$$

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

$$|\Phi(t_1) - \Phi(t_2)| \leq \beta \|t_1 - t_2\| + \gamma$$

whenever $t_1, t_2 \in \mathbb{R}^s$.

Stochastic Program with (Mixed-)Integer Recourse

$$\min_x\{cx \quad + \quad \min_y\{qy \ : \ Wy \geq h(\omega) - Tx, y \in Y\} \ : \ Ax \leq b, \ x \in X\}$$

$$\min_x\{cx \quad + \quad E_\omega(\min_y\{qy \ : \ Wy \geq h(\omega) - Tx, y \in Y\}) \ : \ Ax \leq b, \ x \in X\}$$

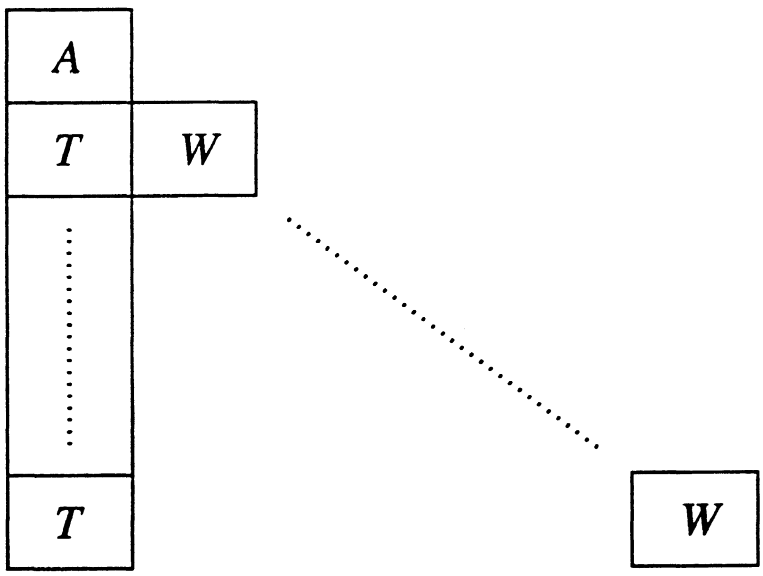
$$\min\{cx + Q(x) \ : \ Ax \leq b, \ x \in X\}$$

Suppose that $h(\omega)$ is discrete with realizations $h^1, \ldots, h^r$ and probabilities $p^1, \ldots, p^r$.

$\Rightarrow$ SP becomes large-scale mixed-integer linear program with block angular structure.

$$\begin{aligned}
\min\{cx + \textstyle\sum_{j=1}^r p^j q y^j \quad &: \ Ax \leq b, \qquad x \in X, \\
&\quad Tx + Wy^1 \qquad\qquad \geq \ h^1, \ \ y^1 \in Y, \\
&\qquad\qquad \vdots \\
&\quad Tx \qquad\qquad + Wy^r \ \geq \ h^r, \ \ y^r \in Y\}
\end{aligned}$$

Block Angular MILP



Dual Decomposition

(jointly with C.C. Carøe (Copenhagen))

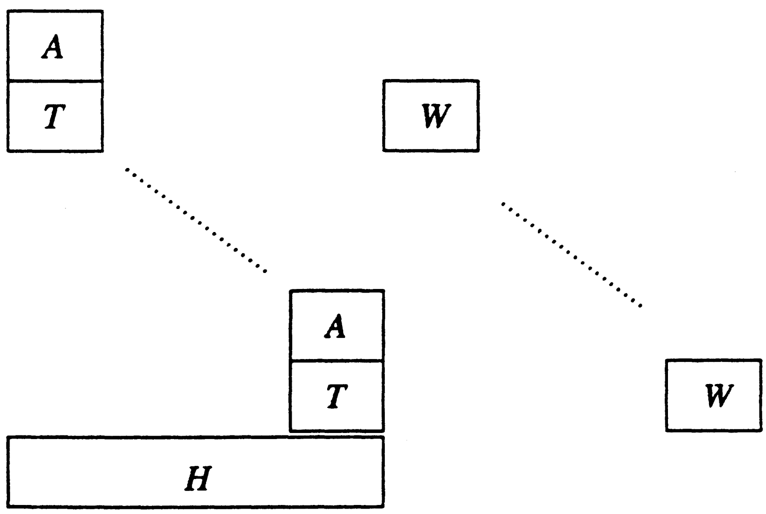
Introduce copies of x according to number of scenarios: $x^1, \ldots, x^r$

Add the constraints $x^1 = \ldots = x^r$ (non-anticipativity !)

Stochastic program transforms into (P):

$$\min\{\sum_{j=1}^{r} p^j(cx^j + qy^j) \;:\; \begin{array}{l} Ax^j \leq b,\ x^j \in X, \\ Tx^j + Wy^j \geq h^j,\ y^j \in Y, j = 1, \ldots, r, \\ x^1 = \ldots = x^r\} \end{array}$$

with the following block structure



where $x^1 = \ldots = x^r \iff \sum_{j=1}^{r} H^j x^j = 0,\ H = (H^1, \ldots, H^r)$

multi-stage $\Longrightarrow$ different H but same coupling structure !

Lagrangian Relaxation of Non-Anticipativity Constraints

Lagrangian:

$$L_j(x^j, y^j, \lambda) \;:=\; p^j(cx^j + qy^j) + \lambda(H^j x^j) \quad j = 1, \ldots, r$$

Lagrangian Relaxation (LR):

$$D(\lambda) = \min\{\sum_{j=1}^{r} L_j(x^j, y^j, \lambda) \quad : \quad Ax^j \leq b,\; x^j \in X,$$
$$Tx^j + Wy^j \geq h^j,\; y^j \in Y, j = 1, \ldots, r\}$$

!! Separability with respect to scenarios !!

!! Similarities among subproblems !!

Lagrangian Dual (D):

$$\max_{\lambda} D(\lambda)$$

!! Non-smooth convex minimization problem !!

!! Values and subgradients of D computable via (LR) !!

Theorem:

$$optval(P) \;\geq\; optval(D)$$

If, for some λ, an optimal solution to (LR) fulfils the non-anticipativity constraints then this solution is optimal for (P).

What do we end up with ?

solve Lagrangian dual

↓

guess a feasible solution

↓

observe a gap

Improvement – Branch-and-Bound

consider problem in form min$\{cx + Q(x) \,:\, Ax \leq b,\ x \in X\}$

Lagrangian dual provides lower bound

feasible solution provides upper bound

start branch-and-bound by subdividing $\{x \,:\, Ax \leq b,\ x \in X\}$

repeat decomposition procedure at each node of the tree

Advantages

- powerful codes available for mixed-integer subproblem solving (CPLEX) and for non-smooth minimization in the Lagrangian dual (NOA - K.C. Kiwiel)

- subproblems quite close to deterministic counterpart of the original (random) problem
(⇒ "deterministic experience" exploitable)

- simple structure of relaxed constraints ⇒ guessing feasible points "easy"

- Lagrangian-relaxation bound often tighter than LP bound

- optimality certificates of increasing quality, tradeoff with computation time

- potential for methods exploiting subproblem similarities
(test sets, Gröbner bases)

- extendability to the multi-stage situation

Problem sizes of deterministic equivalents

Form.	Scen.	Constr.	Var.	Int.	Mult.
Binary	4	47159	47327	7560	11424
	10	113639	109775	14616	28560
	16	180119	172223	21672	45696
Integer	4	32049	37257	5880	4704
	10	78369	89625	12936	11760
	16	124689	141993	19992	18816

Computational Results

- after 10 minutes of CPU-time,
- at a Digital Alpha Personal Workstation with 500MHz processor.

Generator Failure Instances

Form.	Scen.	NOA Steps	Best solution	Lower Bound	Gap	Without NOA
Binary	4	30	3.6417	3.6134	0.8%	3.2%
	10	10	3.6329	3.6050	0.8%	11.1%
	16	5	3.7869	3.6852	2.8%	9.7%
Integer	4	100	3.6249	3.6206	0.1%	3.2%
	10	40	3.6306	3.6251	0.2%	1.7%
	16	25	3.7208	3.7098	0.3%	2.2%

Inaccurate Load Forecast Instances

Form.	Scen.	NOA Steps	Best solution	Lower Bound	Gap	Without NOA
Binary	4	30	3.6598	3.6411	0.5%	1.4%
	10	10	3.6955	3.5781	3.3%	8.3%
	16	5	3.6225	3.5276	2.7%	10.1%
Integer	4	100	3.6579	3.6527	0.1%	4.1%
	10	40	3.6195	3.6080	0.3%	3.1%
	16	25	3.5698	3.5556	0.4%	2.5%

Preliminary computation

Table 2: Problem characteristics and sizes

Prob.	power exchange	Constr.	Variables	Integers	Binaries
A	–	18 641	11 089	1 008	1 512
others	10 scenarios	563 959	332 848	8 784	47 736

Table 3: Calculations

Prob.	Time	Best Solution	Lower Bound	Gap	Min. Saving
A	0:20:14	46 287 933	46 287 933	0.00 %	0.00 %
B	3:43:32	45 950 044	45 419 304	1.16 %	0.72 %
C	3:43:41	46 221 885	45 644 079	1.25 %	0.14 %
D	3:38:31	46 220 219	45 649 030	1.24 %	0.15 %
E	3:39:56	45 990 322	45 585 417	0.90 %	0.62 %
F	22:12:36	46 206 060	45 651 110	1.22 %	0.17 %
G	23:50:43	46 097 080	45 601 185	1.08 %	0.41 %
H	3:28:55	46 109 997	45 633 638	1.03 %	0.38 %

Universal Test Sets

Consider the family of optimization problems

$$(IP)_{c,b}: \qquad \min\{c^\intercal z : Az = b, z \in \mathbb{Z}^n_{\geq 0}\}.$$

$\mathcal{T} \subseteq \mathbb{Z}^n$ is called a universal test set for $(IP)_{c,b}$ if for any $c \in \mathbb{R}^n$, any $b \in \mathbb{R}^d$, and any non-optimal feasible point z_0 of $(IP)_{c,b}$, there exists a $t \in \mathcal{T}$ such that

- $z_0 - t$ is feasible for $(IP)_{c,b}$ and
- $c^\intercal(z_0 - t) < c^\intercal z_0$.

Augmentation Algorithm

Input: a finite test set $\mathcal{T}$, a feasible point z_0 of $(IP)_{c,b}$

Output: an optimum $z_{\min}$ to $(IP)_{c,b}$

while there is $t \in \mathcal{T}$ with $c^\intercal t > 0$ such that $z_0 - t$ is feasible do

$z_0 := z_0 - t$

return z_0

Hilbert Bases

Let C be a rational cone. A finite set $H = \{h_1, \ldots, h_t\} \subseteq C \cap \mathbb{Z}^n$ is called a Hilbert basis of C if every $z \in C \cap \mathbb{Z}^n$ has a representation of the form

$$z = \sum_{i=1}^{t} \lambda_i h_i,$$

with non-negative integral multipliers $\lambda_1, \ldots, \lambda_t$.

Every pointed rational cone possesses a unique Hilbert basis that is minimal with respect to inclusion.

IP Graver test sets

Let $\mathbb{O}_j$ be the j^{th} orthant of $\mathbb{Z}^n$ and $H_j(A)$ the unique minimal Hilbert basis of $\ker(A) \cap \mathbb{O}_j$. Then we define $\mathcal{G}_{IP}(A) := \bigcup H_j(A)$ to be the IP Graver test set (or IP Graver basis) of A.

$\mathcal{G}_{IP}(A)$ is the set of all vectors in $\ker(A) \setminus \{0\}$ minimal with respect to the relation $\sqsubseteq$ defined by

$$u \sqsubseteq v \Leftrightarrow u^{(j)}v^{(j)} \geq 0 \text{ and } |u^{(j)}| \leq |v^{(j)}| \text{ for all } j.$$

Computation of IP Graver test sets

Input: $F = \bigcup_{f \in F(A)} \{f, -f\}$, where $F(A)$ generates $\ker(A)$ over $\mathbb{Z}$

Output: a set G which contains the IP Graver test set $\mathcal{G}_{IP}(A)$

Buchberger-type algorithm

$G := F$
$C := \bigcup_{f,g \in G} \{f + g\}$
while $C \neq \emptyset$ do
  $s :=$ an element in C
  $C := C - \{s\}$
  $f :=$ normalForm(s, G)
  if $f \neq 0$ then
    $C := C \cup \bigcup_{g \in G} \{f + g\}$
    $G := G \cup \{f\}$
return G.

IP normalForm Algorithm

Input: a vector s, a set G of vectors

Output: a normal form of s with respect to G

while there is some $g \in G$ such that $g \sqsubseteq s$ do

$s := s - g$

return s

The Problem

$$\min\{c^\intercal z : A_N z = b, z \in \mathbb{Z}_{\geq 0}^{n_t + N \cdot n_w}\}$$

$$A_N := \begin{pmatrix} A & 0 & 0 & \cdots & 0 \\ T & W & 0 & \cdots & 0 \\ T & 0 & W & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T & 0 & 0 & \cdots & W \end{pmatrix}$$

A_N is a rational matrix, b, c are real vectors.

N indicates number of T's and W's used.

Building Blocks

$$z = (u, v_1, \ldots, v_N) \in \ker(A_N) \Leftrightarrow (u, v_1), \ldots, (u, v_N) \in \ker(A_1)$$

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = A_N z = \begin{pmatrix} Au \\ Tu + Wv_1 \\ \vdots \\ Tu + Wv_N \end{pmatrix}$$

Call $u, v_1, \ldots, v_N$ the building blocks of z.

Notation

$\mathcal{G}_N$ = Graver test set associated with A_N

$\mathcal{H}_N$ = set of building blocks of elements of $\mathcal{G}_N$

$\mathcal{H}_\infty = \bigcup_{N=1}^{\infty} \mathcal{H}_N$

Arrange building blocks in $\mathcal{H}_N$ into pairs (u, V_u) where

$$V_u := \{v \in \mathcal{H}_N : (u,v) \in \ker(A_1)\}.$$

Questions

- Is $\mathcal{H}_\infty$ always (ever) finite?
- How can we reconstruct an improving vector from $\mathcal{H}_\infty$?
- How can we compute $\mathcal{H}_\infty$?

Main Finiteness Result

Theorem 1 *(Maclagan, 1999)*

Let $\{I_1, I_2, \ldots\}$ *be a sequence of monomial ideals in* $k[x_1, \ldots, x_n]$ *such that* $I_j \not\subset I_i$ *whenever* $i < j$. *Then this sequence is finite.*

Theorem 2 $\mathcal{H}_\infty$ *is finite for 2-stage stochastic LPs and IPs.*

Computation of $\mathcal{H}_\infty$ for IP

Input: $F = \mathcal{H}_1 \cup \{0\}$ in (u, V_u)-notation

Output: a set G which contains $\mathcal{H}_\infty$

$G := F$

$C := \bigcup_{f,g \in G} \{f+g\}$

while $C \neq \emptyset$ do

- $s :=$ an element in C
- $C := C - \{s\}$
- $f :=$ normalForm(s, G)
- if $f \neq 0$ then
  - $C := C \cup \bigcup_{g \in G \cup \{f\}} \{f+g\}$
  - $G := G \cup \{f\}$

return G.

$\mathcal{H}_\infty$ normalForm Algorithm for IP

We define $(u', V_{u'}) \sqsubseteq (u, V_u)$ if and only if

1. $u' \sqsubseteq u$, and
2. for all $v_i \in V_u$ there exists $v_i' \in V_{u'}$ such that $v_i' \sqsubseteq v_i$.

(u, V_u) reduces to $(u, V_u) - (u', V_{u'}) := (u - u', \{v_i - v_i' : v_i' \in V_{u'}\})$.

$$(u, V_u) + (u', V_{u'}) := (u + u', \{v + v' : v \in V_u, v' \in V_{u'}\})$$

Reconstructing Improving Vectors

Theorem 2 *Let* $(u', V_{u'})$ *satisfy*

1. $u' \leq u$ *and*
2. $v'_i \leq v_i$, $v'_i \in V_{u'}$, *for* $i = 1, \ldots, N$.

For every $i = 1, \ldots, N$ *choose* $v'_i \in V_{u'}$ *such that* $0 \leq v'_i \leq v_i$ *and* $c_i^\intercal v'_i$ *is maximal.*

If $z' = (u', v'_1, \ldots, v'_N)$ *does not satisfy* $c^\intercal z' > 0$, *then no improving vector can be constructed from the pair* $(u', V_{u'})$.

Computational Example

$$\min\{35x_1 + 40x_2 + \frac{1}{N}\sum_{i=1}^{N} 16y_1^{(i)} + 19y_2^{(i)} + 47y_3^{(i)} + 54y_4^{(i)} :$$

$$x_1 + y_1^{(i)} + y_3^{(i)} \geq \delta_1^{(i)}, \quad 2y_1^{(i)} + y_2^{(i)} \leq \gamma_1^{(i)},$$
$$x_2 + y_2^{(i)} + y_4^{(i)} \geq \delta_2^{(i)}, \quad y_1^{(i)} + 2y_2^{(i)} \leq \gamma_2^{(i)},$$

$$x_1, x_2, y_1^{(i)}, y_2^{(i)}, y_3^{(i)}, y_4^{(i)} \in \mathbb{Z}_+\}$$

(δ_1, δ_2) and (γ_1, γ_2) form a uniform grid on the squares $[300..500] \times [300..500]$ and $[0..2000] \times [0..2000]$

Timings for Computational Example

(CPU seconds on a SUN Enterprise 450, 300 MHz Ultra-SPARC)

Time to compute $\mathcal{H}_\infty$: 18 seconds

(δ, γ)–grids: $(5 \times 5, 3 \times 3)$, $(5 \times 5, 21 \times 21)$, $(9 \times 9, 21 \times 21)$

scenarios	variables	optimum	Aug(H_∞)	CPLEX	dualsip
225	902	(100, 150)	1.52	0.63	> 1800
11025	44102	(100, 100)	67.37	696.10	–
35721	142886	(108, 96)	190.63	> 1 day	–

About Information Constraints

Framework

- finite horizon sequential decision process under uncertainty,
- decision made at stage t based only on information available up to t ($1 \leq t \leq T$),
- information given by a discrete-time stochastic process $\{\xi_t\}_{t=1}^T$:

$$\xi \; : \; (\Omega, \mathcal{A}, I\!\!P) \longrightarrow \times_{t=1}^T I\!\!R^{m_t},$$

- decisions do not influence future information (!)

Non-Anticipativity:

stochastic decision $x_t \in \mathbb{R}^{m_t}$ at stage t

depends only on $\xi^t := (\xi_1, \ldots, \xi_t)$,

i.e., the information available up to time t

Equivalent Formulation:

$$\mathcal{A}_t := \sigma((\xi^t)^{-1}(\mathcal{B}^{m_t}))$$

(smallest $\sigma-$algebra $\mathcal{A}_t \subseteq \mathcal{A}$ containing all the pre-images $(\xi^t)^{-1}(B)$ of Borel sets $B \in \mathcal{B}^{m_t}$ in $\mathbb{R}^{m_t}$)

Then

$$\underline{\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \ldots \subseteq \mathcal{A}_T}$$

where we assume

$$\mathcal{A}_T = \mathcal{A} \text{ and } \mathcal{A}_1 = \{\emptyset, \Omega\}$$

(i.e., ξ_1 and x_1 are deterministic).

$x := (x_1, \ldots, x_T)$ non-anticipative

iff

$\underline{x_t \text{ measurable with respect to } \mathcal{A}_t \text{ for all } t = 1, \ldots, T}$

iff

$$x_t = E[x_t | \mathcal{A}_t]$$

The Discrete Case

- Ω finite, $\Omega = \{\omega_j\}_{j=1}^{r}$,
- $\mathcal{A} = 2^{\Omega}$ power set,
- $\mathbb{P}(\{\omega_j\}) = \pi_j\,,\ j = 1, \ldots, r$,
- $\xi_t^j := \xi_t(\omega_j)$ value of the data scenario j at stage t,

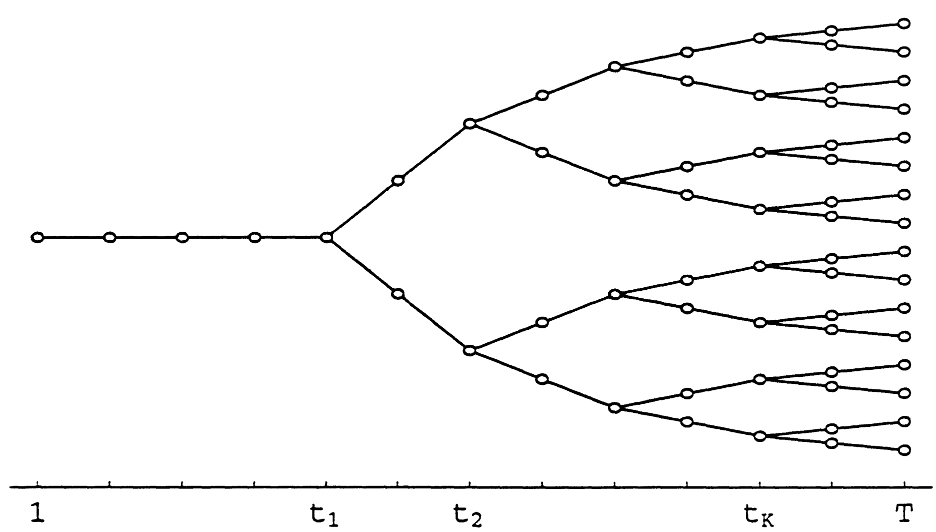
  x_t^j the value of the decision scenario j at stage t

$\forall t = 1, \ldots, T \quad \exists \mathcal{E}_t \subseteq \mathcal{A}_t$:

$\mathcal{E}_t$ partition of Ω and $\sigma(\mathcal{E}_t) = \mathcal{A}_t$,

$card\,\mathcal{E}_t$ = number of realizations of ξ and x at time t

$\Rightarrow$ relations between elements of $\mathcal{E}_t$ and $\mathcal{E}_{t+1}$ representable by a tree



Analytical Expression of Non-Anticipativity

$$E[x_t|\mathcal{A}_t] = \sum_{C\in\mathcal{E}_t} \frac{1}{P(C)} \int_C x_t(\omega) \mathbb{P}(d\omega)\chi_C$$

$$= \sum_{C\in\mathcal{E}_t} (\sum_{\substack{j=1 \\ \omega_j\in C}}^{r} \pi_j)^{-1} (\sum_{\substack{j=1 \\ \omega_j\in C}}^{r} \pi_j x_t^j)\chi_C$$

Non-Anticipativity is equivalent to

$$x_t^\sigma = \sum_{\substack{C\in\mathcal{E}_t \\ \omega_\sigma\in C}} (\sum_{\substack{j=1 \\ \omega_j\in C}}^{r} \pi_j)^{-1} \sum_{\substack{j=1 \\ \omega_j\in C}}^{r} \pi_j x_t^j\,, \quad \sigma = 1,\ldots,r,\, t = 1,\ldots,T$$

$t = 1 \Rightarrow \mathcal{E}_1 = \{\Omega\}$

$$x_1^\sigma = \sum_{j=1}^{r} \pi_j x_1^j\,, \quad \sigma = 1,\ldots,r$$

$(\Leftrightarrow x_1^1 = \ldots = x_1^r)$

Multi-Stage Stochastic Programs, The Link with Dynamic Programming

Random Minimization Problem

Constraints:

$$x \in \times_{t=1}^{T} L_\infty(\Omega, \mathcal{A}, I\!P; I\!R^{m_t}) \text{ fulfils}$$

(1) feasibility of the t-th stage decision x_t

$$x_t \in X_t\,,\; B_t(\xi_t)x_t \geq d_t(\xi_t)\,,\; I\!P - a.s.\,,\; t = 1, \ldots, T$$

(2) relations between decisions at different stages

$$\sum_{\tau=1}^{t} A_{t\tau}(\xi_t)x_\tau \geq g_t(\xi_t)\,,\; I\!P - a.s.\,,\; t = 2, \ldots, T$$

(3) non-anticipativity

x_t measurable with respect to $\mathcal{A}_t$, $t = 1, \ldots, T$

where

- A, B, c, d, g affine,
- $conv\,(X_t)$ compact, polyhedral $\rightarrow$ integer requirements possible

Random objective

$$\sum_{t=1}^{T} c_t(\xi_t)x_t$$

gives rise to different criteria:

Minimize

1. the expectation of (ω–pointwise) minimal costs !

2. the probability that (ω–pointwise) minimal costs do not exceed a preselected threshold !

3. the variance of (ω–pointwise) minimal costs !

subject to:

(functional) non-anticipativity of decisions.

Minimizing the Expectation:

$$\min\Big\{ \int_\Omega \min_{x(\omega)} \{ \sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega) \ : \ (1),\ (2)\} \, I\!P(d\omega) \ : \ x \text{ fulfilling } (3)\Big\}$$

or equivalently

$$\min\Big\{ \int_\Omega \sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega) \, I\!P(d\omega) \ : \ x \text{ fulfilling } (1),\ (2),\ (3)\Big\}$$

Remarks:

- the "classical" approach in stochastic programming,
- "nice problem" in the absence of integrality in x.

Minimizing the Probability:

$$\min\Big\{ \mathbb{P}\Big(\Big\{\omega \in \Omega \; : \; \min_{x(\omega)}\Big\{ \sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega) \; : \; (1),\ (2)\Big\} > \varphi_o\Big\}\Big)$$

$$x \text{ fulfilling } (3)\Big\}$$

or equivalently

$$\min\Big\{ \int_{\Omega} \theta(\omega)\, \mathbb{P}(d\omega) \quad : \quad \sum_{t=1}^{T} c_t(\xi_t(\omega))x_t(\omega) \; - \; \varphi_o \; \leq \; M \cdot \theta(\omega),$$

$$\theta(\omega) \in \{0,1\} \; \mathbb{P}\text{–a.s.}, \;\; x \text{ fulfilling } (1),\ (2),\ (3)\Big\}$$

($M > 0$ sufficiently big constant)

Remarks:

- "inherently integral",
- proposed by Bereanu (1981) as "minimum risk criterion".

The Link with Dynamic Programming

Unified Stochastic Program (USP):

$$\min\{\mathbb{E}[\varphi(x_1,\ldots,x_T,\omega)] \quad : \quad x_t \in L_\infty(\Omega,\mathcal{A},\mathbb{P};\mathbb{R}^{m_t}),$$
$$x_t \text{ is measurable w.r.t. } \mathcal{A}_t\,,\ t=1,\ldots,T\}$$

φ is extended-real-valued and $+\infty$ in case of infeasibility,
further properties of φ:

- joint measurability,
- integrable minorant,
- compact lower level sets.

Questions:

- Solvability of (USP) ?
- Computation of Solutions to (USP) ?

Backward Recursion:

$$\begin{aligned} \psi_{T+1} &:= \varphi , \\ \varphi_t(y_1,\ldots,y_t,\omega) &:= \mathbb{E}^r[\psi_{t+1}(y_1,\ldots,y_t,\cdot)|\mathcal{A}_t](\omega) , \\ \psi_t(y_1,\ldots,y_{t-1},\omega) &:= \inf_y \varphi_t(y_1,\ldots,y_{t-1},y,\omega) , \end{aligned}$$

for all $t = T,\ldots,1$, $\omega \in \Omega$, $y_\tau \in X_\tau$, $\tau = 1,\ldots,T$.

Dynamic Programming:

- At stage t, the infimum w.r.t. y_t is taken.
- Its (conditional) expectation forms the objective at stage $t-1$.

Theorem (Dynkin/Evstigneev, Rockafellar/Wets)

(i) $\{\bar{x}_t\}_{t=1}^T$ is a solution to (USP) iff

$$\varphi_t(\bar{x}^t(\omega),\omega) = \psi_t(\bar{x}^{t-1}(\omega),\omega) , \ \mathbb{P}-\text{a.s.} , \ t = 1,\ldots,T.$$

(ii) In particular, there exists a solution $\bar{x}_1$ to the first-stage problem

$$\min\{\varphi_1(x_1) = \mathbb{E}[\psi_2(x_1,\omega)] \ : \ x_1 \in X_1 , \ B_1x_1 \geq d_1\}.$$

Key Tools: theorems on measurable selections of multifunctions.

! Remark: structure of $\varphi_1(.)$ in (ii) almost undetected. !

The Two-Stage Case

Random Mixed-Integer Linear Program

$$\min\{c^T x + q^T y + q'^T y' \ : \ Tx + Wy + W'y' = h(\omega),$$
$$x \in X, \ \ y \in \mathbb{Z}_+^{\bar{m}}, \ \ y' \in \mathbb{R}_+^{m'}\}$$

first-stage decision x, second-stage decision (y, y'),

W, W' rational matrices,

$h(\omega) \in \mathbb{R}^s$ random vector on $(\Omega, \mathcal{A}, \mathbb{P})$,

$X \subseteq \mathbb{R}^m$ nonempty closed polyhedron.

Value function of MILP:

$$\Phi(t) := \min\{q^T y + q'^T y' \;:\; Wy + W'y' = t,\; y \in Z_+^{\bar{m}},\; y' \in R_+^{m'}\}.$$

Minimizing expected costs:

$$\min\left\{Q_{\mathbb{E}}(x) \;:=\; \int_\Omega (c^T x + \Phi(h(\omega) - Tx))\, \mathbb{P}(d\omega) \;:\; x \in X\right\}$$

Minimizing probability of excessive costs:

$$\min\left\{Q_{\mathbb{P}}(x) \;:=\; \mathbb{P}(\{\omega \in \Omega \;:\; c^T x + \Phi(h(\omega) - Tx) > \varphi_o\}) \;:\; x \in X\right\}$$

Structure of $Q_{\mathbb{E}}$

Proposition (non-integer):

Assume $\bar{m} = 0$, $W'(\mathbb{R}_{+}^{m'}) = \mathbb{R}^{s}$, $\{u \in \mathbb{R}^{s} \;:\; W'^{T}u \leq q'\} \neq \emptyset$, and $\int_{\mathbb{R}^{s}} \|h\| \, \mu(dh) < \infty$.

Then $Q_{\mathbb{E}} : \mathbb{R}^{m} \longrightarrow \mathbb{R}$ is a real-valued convex function.

Proposition (mixed-integer):

Assume that

$$W(\mathbb{Z}_{+}^{\bar{m}}) + W'(R_{+}^{m'}) = \mathbb{R}^{s}, \;\; \{u \in \mathbb{R}^{s} \;:\; W^{T}u \leq q, \; W'^{T}u \leq q'\} \neq \emptyset,$$

and $\int_{\mathbb{R}^{s}} \|h\| \, \mu(dh) < \infty$.

Then it holds

(i) $Q_{\mathbb{E}} : \mathbb{R}^{m} \longrightarrow \mathbb{R}$ is a real-valued lower semicontinuous function,

(ii) if μ has a density, then $Q_{\mathbb{E}}$ is continuous on $\mathbb{R}^{m}$.

Proof techniques:

Fatou's Lemma, Lebesgue's Dominated Convergence Theorem

Structure of $Q_{I\!P}$

Denote

$$M_e(x) := \{h \in I\!R^s : c^T x + \Phi(h - Tx) = \varphi_o\},$$
$$M_d(x) := \{h \in I\!R^s : \Phi \text{ is discontinuous at } h - Tx\}.$$

Proposition:

Assume that

$$W(Z\!\!Z_+^{\bar{m}}) + W'(R_+^{m'}) = I\!R^s \text{ and } \{u \in I\!R^s : W^T u \leq q, \ {W'}^T u \leq q'\} \neq \emptyset.$$

Then it holds

(i) $Q_{I\!P} : I\!R^m \longrightarrow I\!R$ is a real-valued lower semicontinuous function,

(ii) if $\mu(M_e(x) \cup M_d(x)) = 0$, then $Q_{I\!P} : I\!R^m \longrightarrow I\!R$ is continuous at x and $Q_{I\!P} : I\!R^m \times \mathcal{P}(I\!R^s) \longrightarrow I\!R$ is continuous at (x, μ).

Proof techniques:

(semi) continuity of probability measure on sequences of sets, weak convergence of image measures

Discrete Probability Measure – Equivalent MILPs

Let μ be discrete with realizations h_j and probabilities $\pi_j, j = 1, \ldots, r$.

Moreover

$$\Phi(t) := \min\{q^T y \ : \ Wy \geq t, \ y \in Y := \mathbb{Z}_+^{\bar{m}} \times R_+^{m'}\}.$$

Then there are equivalent

$$\min\left\{Q_{\mathbb{E}}(x) \;:=\; \int_{\Omega}(c^T x + \Phi(h(\omega) - Tx))\,\mathbb{P}(d\omega) \;:\; x \in X\right\}$$

and

$$\min_{x,y}\{\sum_{j=1}^{r}\pi_j(c^T x + q^T y_j) \quad : \quad W y_j \;\geq\; h_j - Tx,$$
$$x \in X,\;\; y_j \in Y,\;\; j = 1, \ldots, r\}$$

as well as

$$\min\left\{Q_{\mathbb{P}}(x) \;:=\; \mathbb{P}(\{\omega \in \Omega \;:\; c^T x + \Phi(h(\omega) - Tx) > \varphi_o\}) \;:\; x \in X\right\}$$

and

$$\min_{x,y,\theta}\{\sum_{j=1}^{r}\pi_j\theta_j \quad : \quad W y_j \;\geq\; h_j - Tx, \quad q^T y_j + c^T x - \varphi_o \;\leq\; M_1\theta_j.$$
$$x \in X,\;\; y_j \in Y,\;\; \theta_j \in \{0,1\},\;\; j = 1, \ldots, r\}$$

($M_1 > 0$ sufficiently large)

Both readily extendible to multistage situation !

Decomposition Algorithms

$$\min\{\Lambda(\mathrm{x}) : \quad \mathrm{x}_t \in X_t\,,\ B_t(\xi_t)\mathrm{x}_t \geq d_t(\xi_t)\,,\ I\!\!P - a.s.\,,\ \forall t, \quad (1)$$

$$\sum_{\tau=1}^{t} A_{t\tau}(\xi_t)\mathrm{x}_\tau \geq g_t(\xi_t)\,,\ I\!\!P - a.s.\,,\ \forall t\,, \quad (2)$$

$$\mathrm{x} \in \times_{t=1}^{T} L_\infty(\Omega, \mathcal{A}, I\!\!P; I\!\!R^{m_t})\,,\ \mathrm{x} \in \mathcal{N}_{na}\} \quad (3)$$

Lagrangian Decomposition (Relaxation):

- of (1): single-unit subproblems, (→ see W. Römisch)
- of (2): single-node subproblems, (→ see W. Römisch)
- of (3): single-scenario subproblems.

Outlook:

<u>Structure</u>

- properties of objective functions
- qualitative and quantitative stability under perturbations

<u>Algorithms</u>

- approximation and decomposition techniques
- implementation and numerical experience

	Structure		Algorithms	
	Exp.	Prob.	Exp.	Prob.
linear 2-st.	+	*	+	*
linear m-st.	+	?	+	?
mixed-integer 2-st.	+	*	(+)	*
mixed-integer m-st.	?	?	*	?

+ well understood

* work in progress

? open

Acknowledgements:

C. Carøe (Copenhagen),

R. Hemmecke (Duisburg, Davis),

A. Märkert (Duisburg),

M. Riis (Aarhus),

W. Römisch (Berlin),

S. Tiedemann (Duisburg),

M. Westphalen (Duisburg).

`http://www.uni-duisburg.de/FB11/Math-Net/Preprints.html`