Rutgers University
Center for Operations Research

DIMACS
Center for Discrete Mathematics & Theoretical Computer Science
Founded as a National Science Foundation Science and Technology Center

SOLUTION OF A STOCHASTIC NETWORK DESIGN PROBLEM WITH PROBABILISTIC CONSTRAINT AND DISCRETE RANDOM VARIABLES

ANDRÁS PRÉKOPA
MERVE UNUVAR

RUTCOR-DIMACS-CCICADA WORKSHOP ON STOCHASTIC NETWORKS
OCTOBER 12-13, 2011
RUTGERS UNIVERSITY, PISCATAWAY, NJ

Summary

1. Definition of a (single commodity) network according to Gale, 1957.
2. Examples:
   - Interconnected Power Systems
   - Flood Control Reservoir System Design
   - Parking Lots, Transportation, Location Problems.
3. The Gale–Hoffman Feasibility Theorem.
4. The Prékopa–Boros and the Wallace–Wets Theorems.
5. Elimination of the Redundant Gale–Hoffman Inequalities.
6. A Theorem on *p*-efficient Points
7. The Capacity Design Problem.
8. Solution by the Dentcheva–Prékopa–Ruszczyński Algorithm.
9. Solution by the Prékopa–Vizvári–Badics Algorithm.
10. Numerical Example.

A *network* $G = (N, A)$ is a finite collection of nodes N and a subset A of $N \times N$, which is the collection of arcs. We assume that if $(i,k) \in A$, then also $(k,i) \in A$. The *arc capacity* function is a real-valued function $y(i,k)$, $(i,k) \in A$ on the set of arcs. A *flow* is a real-valued function $f(i,k)$, $(i,k) \in A$ which satisfies the conditions

$$\begin{aligned} & f(i,k) + f(k,i) = 0 \\ & f(i,k) \le y(i,k) \ \text{ for } \ (i,k) \in A. \end{aligned} \tag{1}$$

The definition of y and f can be extended to the entire set $N \times N$, so we write $f(i,k) = y(i,k) = 0$ for $(i,k) \in N \times N$, and $(i,k) \notin A$. We will use the notation

$$\begin{aligned} y(B,C) &= \sum_{i \in B,\, k \in C} y(i,k) \\ f(B,C) &= \sum_{i \in B,\, k \in C} f(i,k), \end{aligned}$$

where B and C are subsets of N.

A *demand function* $d(i)$, $i \in N$ is a real-valued function on the set of nodes. If $B \subseteq N$, then we assign a demand value $d(B)$, to B which is defined by

$$d(B) = \sum_{i \in B} d(i).$$

A demand function (briefly: *demand*) is said to be *feasible* if there exists a flow f such that

$$f(N, i) \geq d(i) \text{ for every } i \in N. \qquad (2)$$

Relations (1) and (2) contain the variables $f(i,k)$, $y(i,k)$ and $d(i)$. It is an important problem to find the projection of the convex polyhedron defined by (1) and (2) onto the space of the variables $y(i,k)$ and $d(i)$, i.e., to give the necessary and sufficient condition in terms of these variables for the existence of a flow satisfying (1) and (2). This problem was solved by Gale (1957) and Hoffman (1960) and the result is contained in the following theorem.

Theorem (Gale and Hoffman). *The demand function* $d(i),\ i \in N$ *is feasible if and only if, for every set* $S \subseteq N$, *we have the inequality*

$$d(S) \le y(\bar{S}, S).$$

For a short proof the reader is referred to Gale (1957).

In power system engineering one node of the network represents one area. To each node i a deterministic generating capacity x_i is assigned, which is diminished by a random deficiency ζ_i, so that the available generating capacity is $x_i - \zeta_i$. Moreover, there exists a random local demand η_i, corresponding to node i, which is to be satisfied first by the use of the generating capacity $x_i - \zeta_i$.

Let $\xi_i = \eta_i + \zeta_i$, $i \in N$. The function

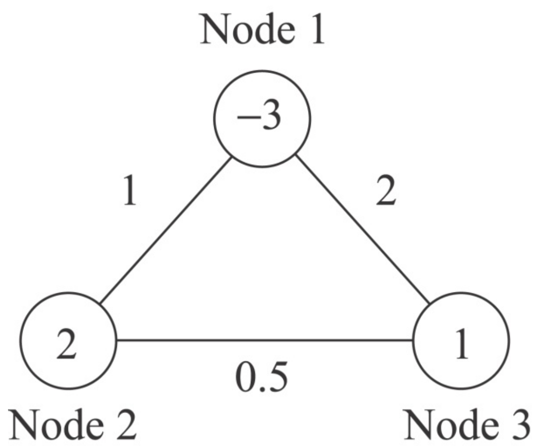
$$d(i) = \xi_i - x_i, \quad i \in N$$

is a demand function corresponding to the network (network demand). If $\xi_i - x_i > 0$, then at node i we need an amount of power $\xi_i - x_i$; and if $\xi_i - x_i < 0$, then at node i there is a surplus generating capacity of $x_i - \xi_i$, which we call the supply. If

$$\sum_{i \in N} x_i \geq \sum_{i \in N} \xi_i$$

then the total available power generating capacity is enough to meet the total demand. However, the transmission system may not be able to allow the individual areas to assist each other to the extent that is necessary. The above stated theorem by Gale and Hoffman provides us with a necessary and sufficient condition for this, i.e., for the existence of a feasible flow.

Simple example: $|N|=3$



$$d(1) = -3, \quad d(2) = 2, \quad d(3) = 1.$$

There is enough supply but if $y(1,2) = 1$, $y(1,3) = 2$, $y(2,3) = 0.5$ then there is no feasible flow

$$d(2) \not\leq y(\{1,3\}, \{2\}).$$

Elimination of the Redundant Inequalities

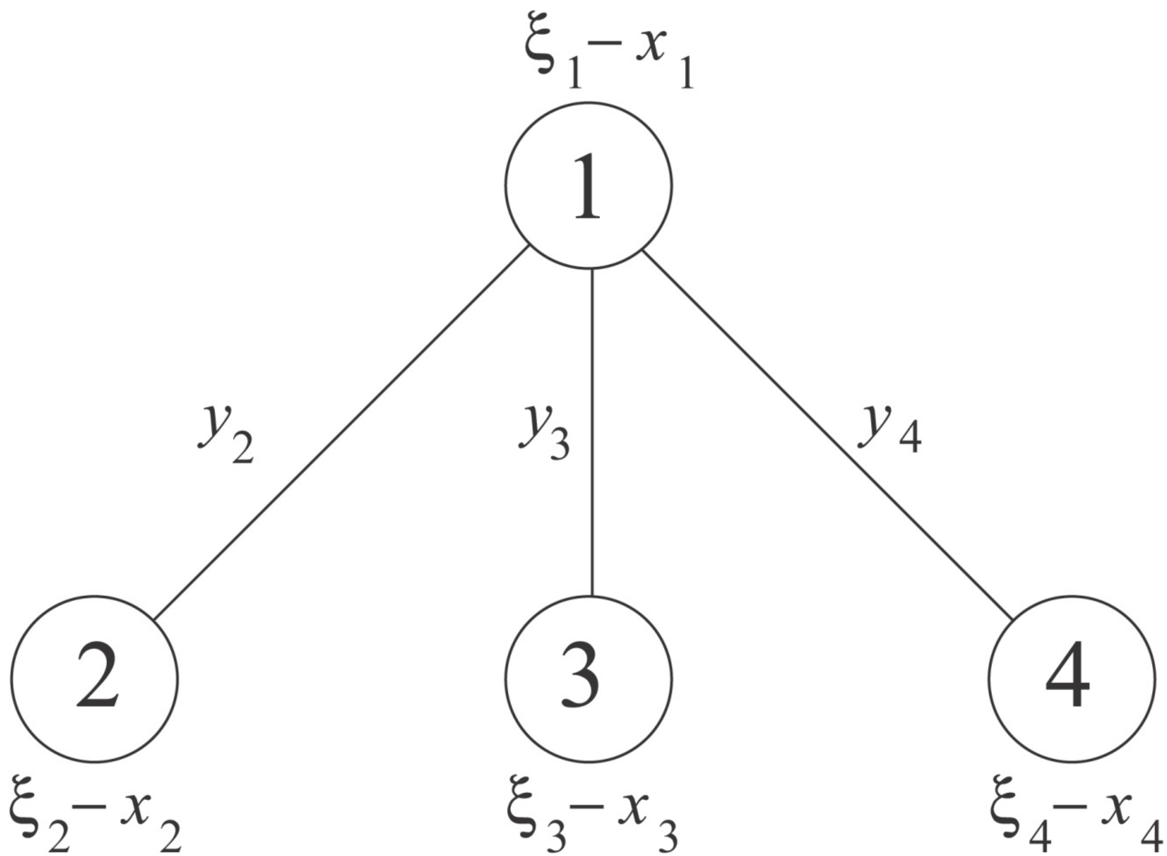
Each Gale–Hoffman inequality corresponds to a subset $S \subset N$.
Let (S) designate that inequality.

Theorem (Prékopa–Boros, 1991). *Let* S_1, S_2 be subsets of N, $S_1 S_2 = \emptyset$ *and suppose that there are no arcs between* S_1 *and* S_2. *Then for every* $S_3 \subset S_1 \cup S_2$ *with* $S_3 S_1 \neq \emptyset$, $S_3 S_2 \neq \emptyset$ *, the inequality* (S_3) *is a consequence of* ($S_3 S_1$) *and* ($S_3 S_2$).

Wallace and Wets (1993) proved also the only if statement. That statement, however, follows from the proof of the Prékopa–Boros theorem.

Theorem (Wallace, Wets, 1993). The inequality (S) is redundant, among the Gale–Hoffman inequalities, iff at least one of the subgraphs $G(S)$, $G(\bar{S})$ is not connected. In that case the inequality $d(S) \leq y(\bar{S}, S)$ is the sum of other Gale–Hoffman inequalities.

Other elimination procedures also apply, for details see Prékopa, Boros (1989).


$\xi_1 - x_1$
1
y_2
y_3
y_4
2
3
4
$\xi_2 - x_2$
$\xi_3 - x_3$
$\xi_4 - x_4$

Gale–Hoffman inequalities:

$$
\begin{aligned}
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 + \xi_4 - x_4 &\le 0 \\
\xi_1 - x_1 &\le y_2 + y_3 + y_4 \\
\xi_2 - x_2 &\le y_2 \\
\xi_3 - x_3 &\le y_3 \\
\xi_4 - x_4 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 &\le y_3 + y_4 \\
\xi_1 - x_1 + \xi_3 - x_3 &\le y_2 + y_4 \\
\xi_1 - x_1 + \xi_4 - x_4 &\le y_2 + y_3 \\
\xi_2 - x_2 + \xi_3 - x_3 &\le y_2 + y_3 \\
\xi_2 - x_2 + \xi_4 - x_4 &\le y_2 + y_4 \\
\xi_3 - x_3 + \xi_4 - x_4 &\le y_3 + y_4 \\
\xi_2 - x_2 + \xi_3 - x_3 + \xi_4 - x_4 &\le y_2 + y_3 + y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_4 - x_4 &\le y_3 \\
\xi_1 - x_1 + \xi_3 - x_3 + \xi_4 - x_4 &\le y_2 .
\end{aligned}
$$

Inequalities 9, 10, 11, 12 are sums of others, hence they are redundant.

The Remaining Gale–Hoffman Inequalities in Case of the Four Node Network, After Elimination by Graph Structure

$$
\begin{aligned}
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 + \xi_4 - x_4 &\le 0 \\
\xi_1 - x_1 &\le y_2 + y_3 + y_4 \\
\xi_2 - x_2 &\le y_2 \\
\xi_3 - x_3 &\le y_3 \\
\xi_4 - x_4 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 &\le y_3 + y_4 \\
\xi_1 - x_1 + \xi_3 - x_3 &\le y_2 + y_4 \\
\xi_1 - x_1 + \xi_4 - x_4 &\le y_2 + y_3 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_3 - x_3 &\le y_4 \\
\xi_1 - x_1 + \xi_2 - x_2 + \xi_4 - x_4 &\le y_3 \\
\xi_1 - x_1 + \xi_3 - x_3 + \xi_4 - x_4 &\le y_2.
\end{aligned}
$$

Here no inequality is the sum of others but on the left hand side there are four lines (2, 3, 4, 5), where $\xi_i - x_i$ stands alone. All left hand sides are sums of the left hand sides of these four inequalities.

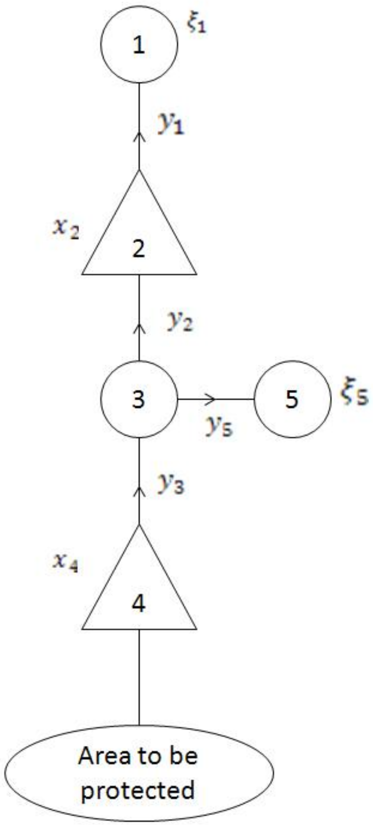
Other examples for networks, from the point of view of application:

- Flood control hydraulic networks.
- Evacuation network
- Transportation networks with parking facilities.

In these cases

ξ_i = demand for freeboard, parking place, shelter room

x_i = capacity for the same at node .



Flood control hydralic network.

Source: Prékopa–Szántai,1978.

ξ_1, ξ_2 flood amounts to be retained or demands for freeboard

$-x_4, -x_5$ reservoir capacities

x_4, x_5 freeboard supply values

y1. y2, y3, y5 arc capacities

Gale-Hoffman Inequalities 2^5-1=31

1. $S = \emptyset$, trivial
2. $\bar{S} = N$, $\xi_1 + \xi_5 \leq x_2 + x_4$
3. $\bar{S} = 1$, $\xi_1 \leq y_1$
4. $\bar{S} = 2, 3, 4, 5$, $\xi_5 - x_2 - x_4 \leq 0$
5. $\bar{S} = 2$, $-x_2 \leq y_2$
6. $\bar{S} = 1, 3, 4, 5$, $\xi_1 + \xi_5 - x_4 \leq y_1$
7. $\bar{S} = 3$, $0 \leq y_3$
8. $\bar{S} = 1, 2, 4, 5$, $\xi_1 + \xi_5 - x_2 - x_4 \leq y_2 + y_5$
9. $\bar{S} = 4$, $-x_4 \leq 0$
10. $\bar{S} = 1, 2, 3, 5$, $\xi_1 + \xi_5 - x_2 \leq y_3$
11. $\bar{S} = 5$, $\xi_5 \leq y_5$
12. $\bar{S} = 1, 2, 3, 4$, $\xi_1 - x_2 - x_4 \leq 0$
13. $\bar{S} = 1, 2$, $\xi_1 - x_2 \leq y_2$
14. $\bar{S} = 3, 4, 5$, $\xi_5 - x_4 \leq 0$
15. $\bar{S} = 1, 3$, $\xi_1 \leq y_1 + y_3$
16. $\bar{S} = 2, 4, 5$, $\xi_5 - x_2 - x_4 \leq y_2 + y_5$
17. $\bar{S} = 1, 4$, $\xi_1 - x_4 \leq y_1$
18. $\bar{S} = 2, 3, 5$, $\xi_5 - x_2 \leq y_2 + y_5$
19. $\bar{S} = 1, 5$, $\xi_1 + \xi_5 \leq y_1 + y_5$
20. $\bar{S} = 2, 3, 4$, $-x_2 - x_4 \leq 0$
21. $\bar{S} = 2, 3$, $-x_2 \leq y_3$
22. $\bar{S} = 1, 4, 5$, $\xi_1 + \xi_5 - x_4 \leq y_1 + y_5$

Gale-Hoffman Inequalities 2^5-1=31 (cont')

23. $\bar{S} = 2, 4,\ -x_2 - x_4 \leq y_2$

24. $\bar{S} = 1, 3, 5,\ \xi_1 + \xi_5 \leq y_1 + y_3$

25. $\bar{S} = 2, 5,\ \xi_5 - x_2 \leq y_2 + y_5$

26. $\bar{S} = 3, 4,\ -x_4 \leq 0$

27. $\bar{S} = 1, 2, 5,\ \xi_1 + \xi_5 - x_2 \leq y_2 + y_5$

28. $\bar{S} = 3, 5,\ \xi_5 \leq y_3$

29. $\bar{S} = 1, 2, 4,\ \xi_1 - x_2 - x_4 \leq y_2$

30. $\bar{S} = 4, 5,\ \xi_5 - x_4 \leq y_5$

31. $\bar{S} = 1, 2, 3,\ \xi_1 - x_2 \leq y_3.$

Remaining Inequalities After Elimination by Network Topology

2. $\xi_1 + \xi_5 \le x_2 + x_4$

3. $\xi_1 \le y_1$

10. $\xi_1 + \xi_5 \le y_3 + x_2$

11. $\xi_5 \le y_5$

13. $\xi_1 \le x_2 + y_2$

14. $\xi_5 \le x_4$

28. $\xi_5 \le y_3$.

More Concise Form

$$\xi_1 \le \min(y_1, x_2 + y_2)$$

$$\xi_5 \le \min(y_3, y_5, x_4)$$

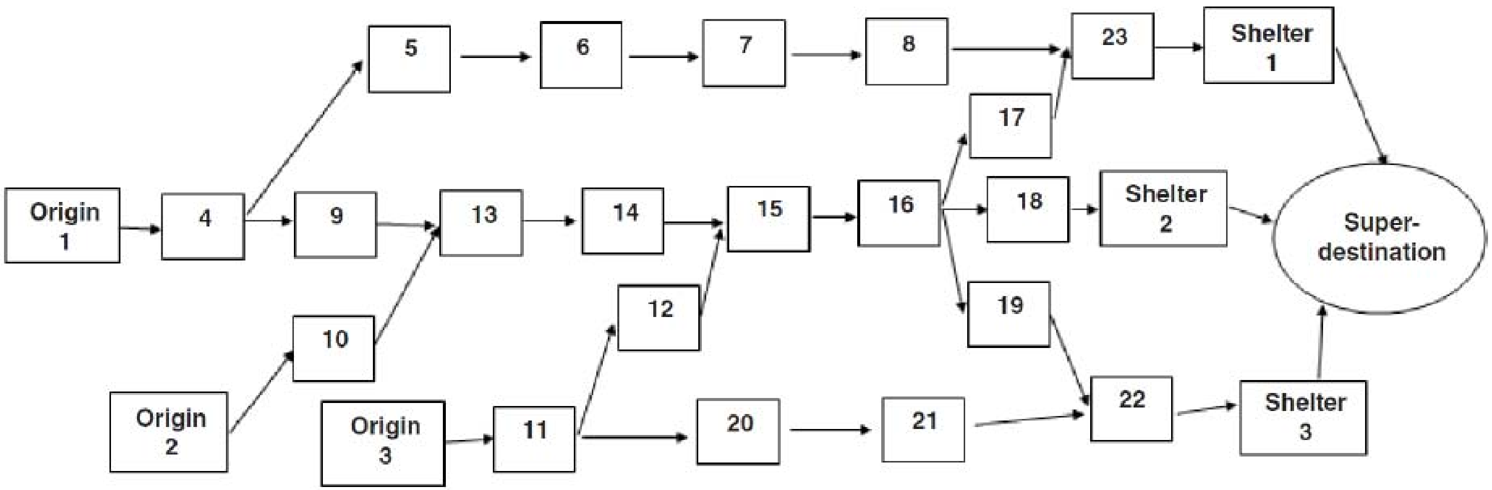
$$\xi_1 + \xi_5 \le \min(x_2 + x_4, y_3 + x_2).$$

If $y_1 = y_2 = y_3 = \infty$, taken

$$\xi_1 + \xi_5 \le x_2 + x_4$$

$$\xi_5 \le x_4.$$

Simplified Cell Representation of Cape May Evacuation Network



p-efficient Points of a Multivariate Discrete Probability Distribution

$$\xi = (\xi_1,\ldots,\xi_n)^T, \qquad F(z) = P(\xi \le z),\ \ z \in R^n.$$

Assumption. Each ξ_i has finite support Z_i. Let $Z = Z_1 \times \cdots \times Z_n$.

Definition. The point $z \in Z$ is a p-efficient point of the distribution of ξ, if

$$F(z) \ge p$$

and there is no $y < z$ such that $F(y) \ge p$.

Algorithms to enumerate all *p*-efficient points:

Prékopa, Vizvári, Badics (1996, 1998)

Boros, Elbassioni, Gurvich, Khachiyan, Makino (2003).

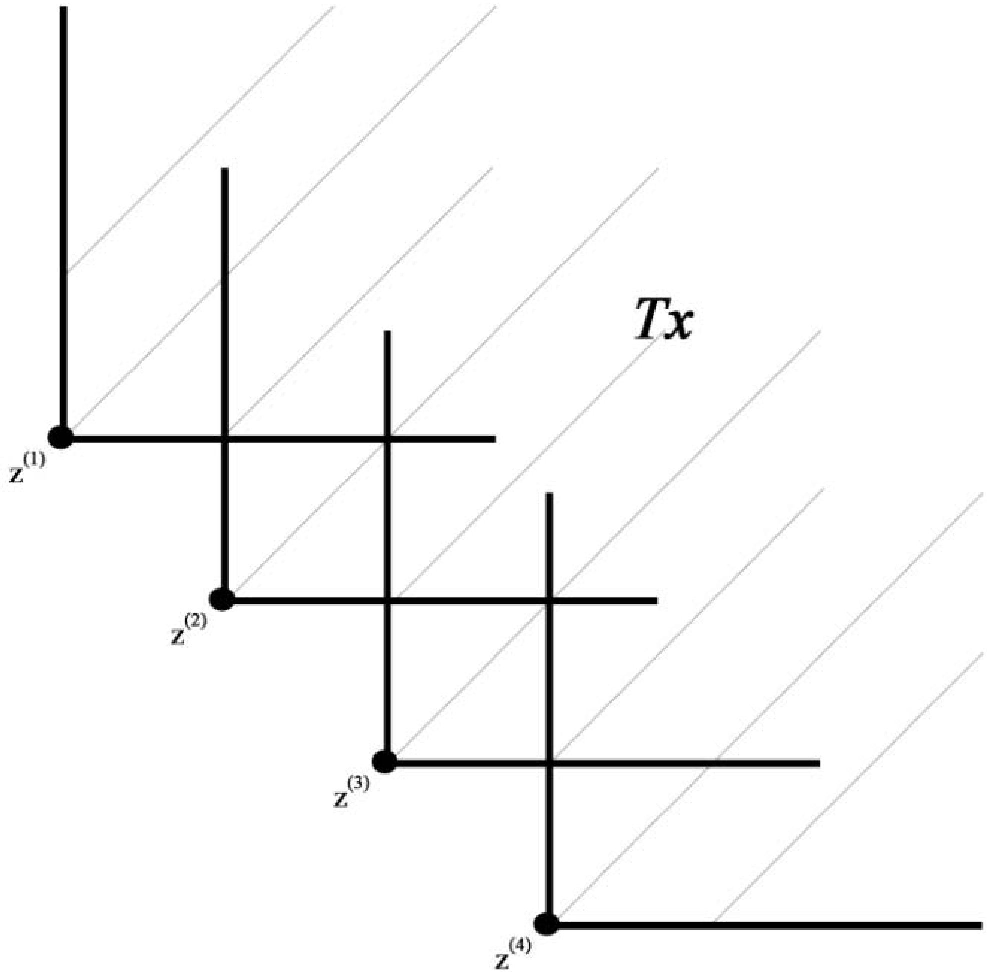
The concept of a p-efficient point has successfully been applied in programming under probabilistic constraint with discrete right hand side random vector:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&P(Tx \geq \xi) \geq p\\&Ax = b,\ \ x \geq 0.\end{aligned}$$

If $z^{(1)}, \ldots, z^{(M)}$ are the p-efficient points of ξ, then the probabilistic constraint is equivalent to the following:

$$Tx \geq z^{(i)},\ \text{for at least one } i = 1, \ldots, M.$$

The Constraint $P(Tx \geq \xi) \geq p$ is equivalent to the requirement that Tx is an element of the shaded set, where $z^{(1)}, z^{(2)}, z^{(3)}, z^{(4)}$ are the p-efficient points



The original problem is a disjunctive problem:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&Tx \ge z^{(i)},\ \text{ for at least one } i = 1,\dots, M\\&Ax = b,\ x \ge 0.\end{aligned}$$

Through "convexification'' we obtain a relaxation of it:

$$\begin{aligned}&\text{Min } c^T x\\&\text{subject to}\\&Tx \ge \sum_{i=1}^{M} \lambda_i z^{(i)}\\&\sum_{i=1}^{M} \lambda_i = 1\\&Ax = b,\ x \ge 0,\ \lambda \ge 0,\end{aligned}$$

where the decision variables are the components of x and λ.

A Basic Theorem on *p*-efficient Points

Let $\xi = (\xi_1, \ldots, \xi_n)^T$ be a random vector, where the support of ξ_i is a finite set Z_i, $i = 1, \ldots, n$. Let

$$Z = Z_1 \times \cdots \times Z_n$$

and

$$z^{(i)} = (z_1^{(i)}, \ldots, z_n^{(i)})^T, \quad i = 1, \ldots, M$$

the *p*-efficient points of the random vector ξ.

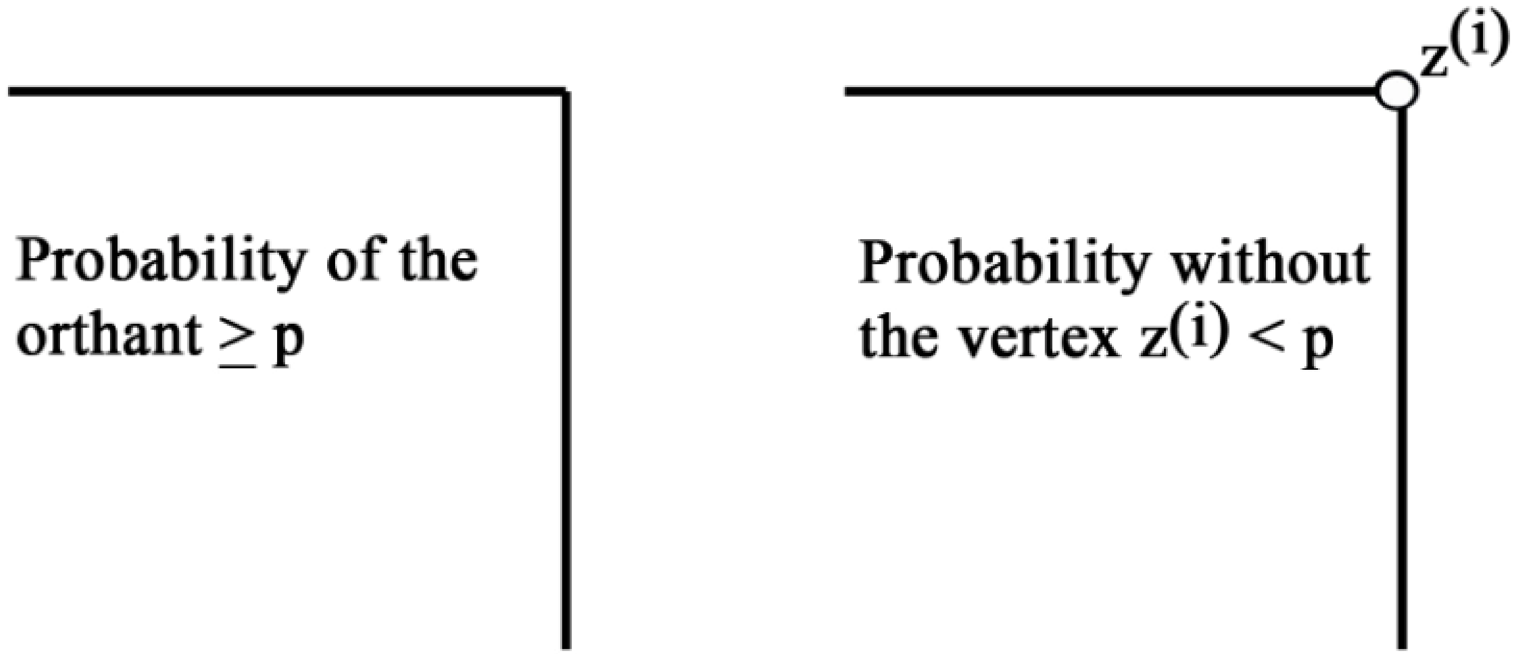
Let $B \geq 0$ be an n ´ M matrix that has positive entry in each row.

Assertion: If, $P(\{z \in Z \mid z \leq z^{(i)}\} \setminus \{z^{(i)}\} < p \quad i = 1, \ldots, M$, then the *p*-efficient points of the random vector

$$\begin{pmatrix} \xi \\ B\xi \end{pmatrix}$$

are

$$\begin{pmatrix} z^{(i)} \\ Bz^{(i)} \end{pmatrix}, \quad i = 1, \ldots, M.$$



- Minor restriction from practical point of view. Slight perturbation of the probability distribution can make the condition satisfied.

Example.

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \qquad B = (1,1)$$

ξ has p-efficient points, taken with positive probabilities:

$$\begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(2)} \\ z_2^{(2)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(3)} \\ z_2^{(3)} \end{pmatrix}.$$

Then the random vector

$$\begin{pmatrix} \xi \\ B\xi \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_1 + \xi_2 \end{pmatrix}$$

has p-efficient points:

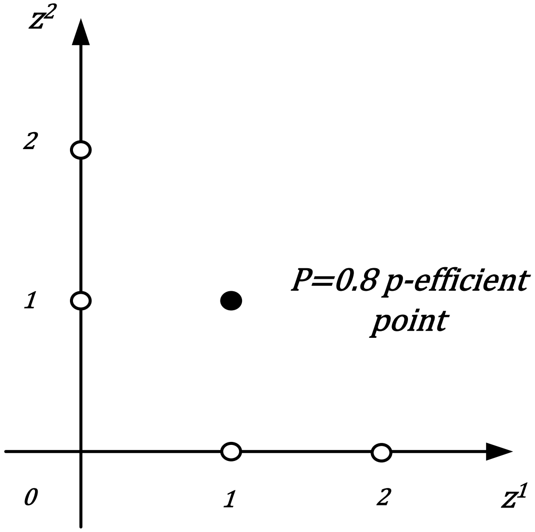
$$\begin{pmatrix} z_1^{(1)} \\ z_2^{(1)} \\ z_1^{(1)} + z_2^{(1)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_1^{(2)} + z_2^{(2)} \end{pmatrix}, \quad \begin{pmatrix} z_1^{(3)} \\ z_2^{(3)} \\ z_1^{(3)} + z_2^{(3)} \end{pmatrix}.$$

The theorem is very important from the point of view of network reliability calculation. If the number of nodes is n, then the number of Gale–Hoffman inequalities is $2^n - 1$. Still, it is enough to determine the p-efficient points of the n random demands at the n nodes because then we can easily generate the p-efficient points for the entire collection of the random variables.

The condition $P(\xi = z^{(i)}) > 0$, $i = 1, \dots, N$ is essential.

Example: Let $r = 2$, $z_1 = z_2 = \{0, 1, 2\}$

$$p_{ij} = P(\xi_1 = i, \xi_2 = j),$$
$$p_{00} = 0.6,\ \ p_{01} = p_{02} = p_{10} = p_{12} = 0.1,$$
$$p_{11} = p_{12} = p_{21} = p_{22} = 0,\ \ p = 0.8$$



ξ has one 0.8-efficient point: (1,1) but (1,1,2) is not 0.8-efficient for $(\xi_1, \xi_2, \xi_1 + \xi_2)$. In fact, $P(\xi_1 \leq 1, \xi_2 \leq 1, \xi_1 + \xi_2 \leq 1) = 0.8$.

The Stochastic Network Design Problem

Decision variables: $x_1, \ldots, x_n$, node capacities, $n = |N|$.

$$\text{Min} \sum_{i=1}^{n} c_i x_i$$
$$P(d(S) \leq y(\bar{S}, S), (S) \text{ non-eliminated}) \geq p$$
$$l_i \leq x_i \leq u_i, \ i = 1, \ldots, n,$$
$$\text{where } d(S) = \sum_{i \in S} (\xi_i - x_i).$$

Let $z^{(i)}$, $i = 1, \ldots, n$ be the p-efficient points of the random vector $\xi = (\xi_1, \ldots, \xi_n)$.

Form of The General Problem

The arc capacities are assumed to be constants, for simplicity.

$$\text{Min} \sum_{i=1}^{n} c_i x_i$$

subject to

$$P(Tx \geq \eta) \geq p$$

$$Ax \geq b.$$

Let $v^{(j)}, \ j \in J$ be the set of p-efficient points of the random vector ξ.

Relaxation of the Problem

$$\text{Min} \sum_{i=1}^{n} c_i x_i$$

subject to

(P)

$$Tx \geq \sum_{j \in J} \lambda_j v^{(j)}$$

$$Ax \geq b$$

$$\sum_{j \in J} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j \in J.$$

Solution of the Relaxed Problem by the Dentcheva–Prékopa–Ruszczyński Algorithm (Column Generation). The *p*-efficient Points are Simultaneously Generated with the Solution Algorithm.

In each iteration we generate a new *p*-efficient point and solve an LP. In iteration k we solve the problem

(PR)

$$\begin{aligned}
&\text{Min } c^T x \\
&\text{subject to} \\
&\quad Tx \geq \sum_{j \in J_k} \lambda_j v^{(j)} \\
&\quad Ax \geq b \\
&\sum_{j \in J_k} \lambda_j = 1, \quad \lambda_j \geq 0, \quad j \in J_k.
\end{aligned}$$

Introduce slack variables u into problems (P) and (PR) to obtain

(P)

$$
\begin{aligned}
&\text{Min}\, c^T x \\
&\text{subject to} \\
&Tx - u - \sum_{j \in J} \lambda_j v^{(j)} = 0 \\
&Ax = b \\
&\sum_{j \in J} \lambda_j = 1 \\
&x \ge 0, \quad u \ge 0, \quad \lambda \ge 0.
\end{aligned}
$$

(PR)

$$
\begin{aligned}
&\text{Min}\, c^T x \\
&\text{subject to} \\
&Tx - u - \sum_{j \in J_k} \lambda_j v^{(j)} = 0 \\
&Ax = b \\
&\sum_{j \in J_k} \lambda_j = 1 \\
&x \ge 0, \quad u \ge 0, \quad \lambda_j \ge 0, \quad j \in J_k.
\end{aligned}
$$

Solve optimally (PR) by a method that produces optimal dual vector

$$(w_1^T, w_2^T, w_3)$$

and check if the optimality condition holds in (P) with this.

Yes: stop, optimal solution has been obtained.

No: In this case the inequalities

$$-w_1^T v^{(j)} + w_3 \le 0, \quad j \in J_k$$

hold with equality for at least one j, hence

$$w_3 = \operatorname*{Min}_{j \in J_k} w_1^T v^{(j)}$$

but there is at least one $j \in J$ such that

$$w_3 > w_1^T \, v^{(j)} \, .$$

In other words, we have the inequality

$$\operatorname*{Min}_{j \in J_k} w_1^T v^{(j)} > \operatorname*{Min}_{j \in J} w_1^T v^{(j)}.$$

The optimal solution of the problem $\operatorname*{Min}_{j\in J} w_1^T v^{(j)}$ is the same as

$$\text{Min } w_1^T v$$

$$\text{subject to}$$

$$P(\eta \le v) \ge p.$$

The optimal solution to this last problem can be obtained by the use of the solution of the much simpler problem.

$$\text{Min } \gamma^T v$$

$$\text{subject to}$$

$$P(\xi \le z) \ge p.$$

Recall that the p-efficient points of

$$\eta = (\xi_1, \ldots, \xi_n, partial\ \ sums)^T$$

are

$$v = (z_1, \ldots, z_n, partial\ \ sums)^T$$

where $(z_1, \ldots, z_n)^T$ is a p-efficient point of $(\xi_1, \ldots, \xi_n)^T$

This implies that

$$w_1^T v = \gamma^T z,$$

where the components of γ are partial sums of w_1.

When we apply this algorithm for the solution of the network design problem, then the random variable η has size r equal to the number of non-eliminated Gale–Hoffman inequalities.

Some of the components of η are ξ_1 , …. , ξ_n , the local demands at the nodes of the network, the others are partial sums of them.

By the basic theorem of p-efficient points, it is enough to formulate and solve the knapsack problem based on $\xi = (\xi_1 , \ldots , \xi_n)$, where $n = |N|$, rather than based on η, the number of components of which (after the elimination) may still be very large.

This fact contributes greatly to the efficiency of the solution of the network design problem.

Finding new p-efficient Point

The problem to be solved is:

$$\min \gamma^T z$$

$$\text{subject to}$$

$$F(z) \geq p$$

$$z \in Z, \;\; l_i \leq z_i \leq u_i, \;\; i = 1, \ldots, n,$$

where l_i, u_i are known bounds.

If $\xi_1,\ldots,\xi_n$ are independent, then:

$$\min \gamma^T z$$

subject to

$$\sum_{i=1}^{n} (-\log F_i(z_i)) \leq -\log p$$

$$z \in Z, \;\; l_i \leq z_i \leq u_i, \;\; i = 1, \ldots, n,$$

We also assume that $\xi_1,\ldots,\xi_n$ are integer valued.

Equivalent Formulation of the Problem

write: $a_{ik} = -\log F_i(k), \underset{\cdot}{=} -\log p, \ z_i = \sum_{k=l_i}^{u_i} k\delta_{ik}$

$$\min \sum_{i=1}^{m} \sum_{k=l_i}^{u_i} w_{2i} k \delta_{ik}$$

subject to

$$\sum_{i=1}^{m} \sum_{k=l_i}^{u_i} a_{ik} \delta_{ik} \leq d$$

$$z \in D$$

$$\sum_{i=1}^{m} \delta_{ik} = 1, \ \ i = 1, \ldots, m$$

$$\delta_{ik} \in \{0, 1\}, \ \ \text{all } i, k,$$

Multiple Choice Knapsack Problem, MCKP

Solution of the MCKP

Relaxation, allowing $0 \leq \delta_{ik} \leq 1, \text{ all } i, k$

The problem is called Linear Multiple Choice Knapsack Problem, LMCKP

$$\min \sum_{i=1}^{n} \sum_{k=l_i}^{u_i} w_{2i} k \delta_{ik}$$

$$\text{subject to}$$

$$\sum_{i=1}^{n} \sum_{k=l_i}^{u_i} a_{ik} \delta_{ik} \leq d$$

$$z \in D$$

$$\sum_{k=1}^{n_i} \delta_{ik} = 1, \;\; i = 1, \ldots, n$$

$$0 \leq \delta_{ik} \leq 1, \;\; \text{all } i, k.$$

For efficient solution see Pisinger (1995). However, we have our own, more efficient solution, using ideas from stochastic programming

Introduce slack variable u in the inequality constraint, then split the sum into m terms, each term corresponds to a component of ξ . Changing the summation range to 1 through m_i, $i=1,\dots,m$ and designate the coefficients in the new sums by $h_{i1}, \dots, h_{im_i}$, $i=1,\dots,n$.

The new problems is:

$$\min\left\{0u + 0u_1 + \cdots + 0u_n + h_{11}\delta_{11} + \cdots + h_{1m_1}\delta_{1m_1} + \cdots + h_{n1}\delta_{n1} + \cdots + h_{nm_n}\delta_{nm_n}\right\}$$

$$\text{subject to}$$

$$\begin{array}{lllll}
u + u_1 + \cdots + u_n & & & & = d \\
\quad u_1 & -a_{11}\delta_{11} - \cdots - a_{1m_1}\delta_{1m_1} & & & = 0 \\
\quad\quad \ddots & & & & \vdots \\
\quad\quad\quad u_n & & -a_{n1}\delta_{n1} - \cdots - a_{nm_n}\delta_{nm_n} & & = 0 \\
 & \delta_{11} + \cdots + \delta_{n1} & & & = 1 \\
 & \quad\quad \ddots & & & \vdots \\
 & & \delta_{n1} + \cdots + \delta_{nm_n} & & = 1
\end{array}$$

$$u_i \geq 0,\ \ i = 1, \dots, n,\ \ \delta_{ik} \geq 0,\quad \text{all } i, k.$$

Simple recourse type problem. Fast algorithms and fast bounds in Prékopa (1990, 1995), Fábián, Prékopa, Ruff-Fiedler (1995). In the optimal solution one δ or a pair appears in each block. In the latter case a simple cost-efficient argument given the solution to the MCKP.

Solution by the Prékopa–Vizvári–Badics (1998) Cutting Plane Algorithm

Suppose that all p-efficient points of the random vector η have been generated (in the network design problem we may simplify their enumeration by the use of the basic theorem of p-efficient points) and they are $v^{(j)}$, $j=1, \ldots, M = |J|$. We solve the problem:

$$\text{Min}\, c^T x$$

subject to

(P)

$$\begin{aligned}
Tx - u - \sum_{j=1}^{M} \lambda_j v^{(j)} &= 0 \\
Ax &= b \\
\sum_{j=1}^{M} \lambda_j &= 1
\end{aligned}$$

$$x \ge 0, \quad u \ge 0, \quad \lambda \ge 0.$$

Preprocessing if $v^{(j)}$, $j = 1, \ldots, M$ are in a lower dimensional manifold

$$\overline{v} = \frac{1}{M}\sum_{j=1}^{M} v^{(j)}.$$

Find maximum number of linearly independent solutions of the equations

$$w^T(v^{(j)} - \overline{v}) = 0, \quad j = 1, \ldots, M$$

and let them be $w_1, \ldots, w_h$.

We will append the constraints

$$w_l^T(Tx - u - \overline{v}) = 0, \quad l = 1, \ldots, h$$

to the constraints $Ax = b$.

Example: In the two-dimensional case the p-efficient points may be on a line

Auxiliary Problem in the *k^th^* Iteration

$$\text{Min}\, e^T \mu = \alpha$$

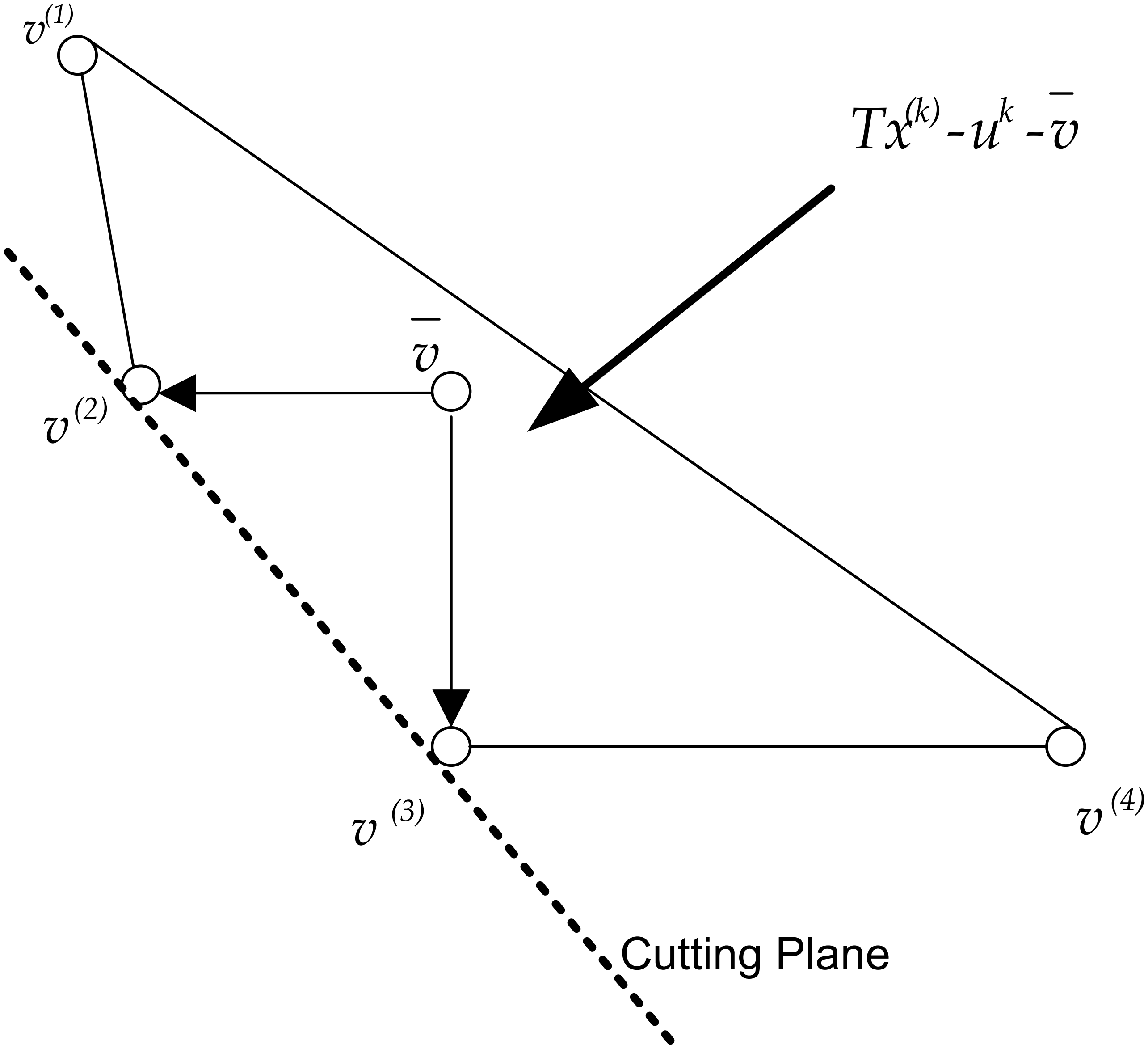
subject to

$$\sum_{j=1}^{M} (v^j - \overline{v})\mu_i = Tx^k - u^k - \overline{v}$$

$$\mu \geq 0,$$

where e^T = (1, …, 1). Here m_1, ..., m_M are the decision variables

If a £ 1, Stop, optimal solution has been found.



The auxiliary problem generates new cutting plane

If $\alpha > 1$ then let

$$v^{(j_1)} - \overline{v}, \ldots, v^{(j_r)} - \overline{v}$$

be an optimal basis of the auxiliary problem. Find $w \neq 0$ such that

$$\begin{aligned} w^T w_j &= 0, \quad j = 1, \ldots, h \\ w^T \left(v^{(j_i)} - v^{(j_1)} \right) &= 0, \quad j = 2, \ldots, r - h. \end{aligned}$$

These equations determine w up to a constant. If

$$w^T (Tx^k - u^k - \overline{v}) < 0,$$

then define $w^{(k+1)} = w$ and introduce the cut

$$\left(w^{(k+1)} \right)^T (Tx - u - \overline{v}) \geq 0.$$

Algorithm PVB

Step 1. Enumerate the p-efficient points of η. Initialize $k \leftarrow 0$.

Step 2. Solve the LP

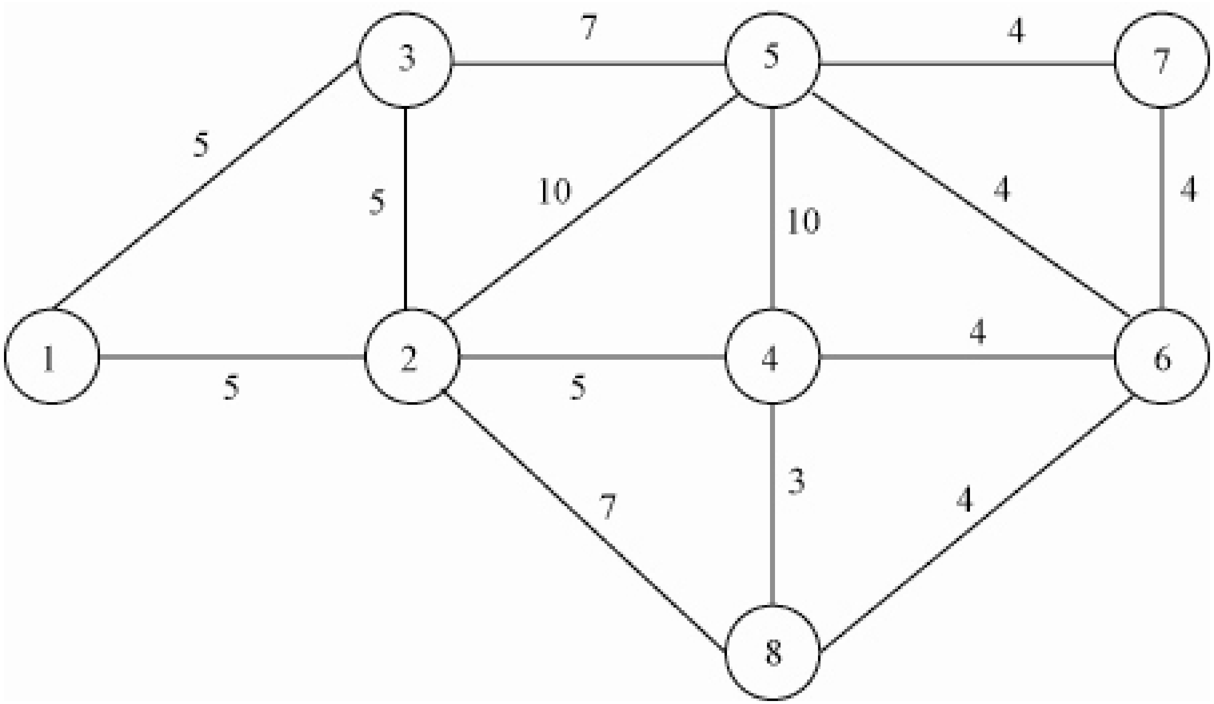
$$
\begin{aligned}
&\text{Min}\, c^T x \\
&\text{subject to} \\
&\qquad Ax = b \\
&\left(w^{(i)}\right)^T (Tx - u - \overline{v}) \geq 0, \quad i = 1,\ldots,k \\
&\qquad w_l^T (Tx - u - \overline{v}) = 0, \quad l = 1,\ldots,h \\
&\qquad x \geq 0, \quad u \geq 0.
\end{aligned}
$$

If $k = 0$, then ignore the constraint involving the cuts. Let $(x^{(k)}, u^{(k)})$ be an optimal solution.

Step 3. Solve the auxiliary problem. If $\alpha \leq 1$, Stop, optimal solution has been found. If $\alpha > 1$, go to Step 4.

Step 4. Create the new cut, append it to the existing cuts and go to Step 2.

Numerical Example



Eight-node Network from Prékopa–Boros (1989)

Possible Values and Corresponding Probabilities of the Random Demands in the 8-Node Network

ξ_1	34	39	44	49	54	59	64	69	74	79
ξ_2	33	38	43	48	53	58	63	68	73	78
ξ_3	17	22	27	32	37	42	47	52	57	62
ξ_4	33	38	43	48	53	58	63	68	73	78
ξ_5	15	20	25	30	35	40	45	50	55	60
ξ_6	10	15	20	25	30	35	40	45	50	55
ξ_7	15	20	25	30	35	40	45	50	55	60
ξ_8	25	30	35	40	45	50	55	60	65	70

$p_1,\dots,p_8$ are the Distribution of the Random Variables $\xi_1,\dots,\xi_{8,}$ Respectively

p_1	n=9 p=0.4	0.010078	0.060466	0.161243	0.250823	0.250823	0.167215	0.074318	0.021234	0.003539	0.000262
p_2	n=9 p=0.45	0.004605	0.033912	0.110986	0.211881	0.260036	0.212757	0.116049	0.040693	0.008323	0.000757
p_3	n=9 p=0.5	0.001953	0.017578	0.070313	0.164063	0.246094	0.246094	0.164063	0.070313	0.017578	0.001953
p_4	n=9 p=0.6	0.000262	0.003539	0.021234	0.074318	0.167215	0.250823	0.250823	0.161243	0.060466	0.010078
p_5	n=9 p=0.48	0.00278	0.023095	0.085272	0.183664	0.254303	0.234742	0.144456	0.057148	0.013188	0.001353
p_6	n=9 p=0.35	0.020712	0.100373	0.216188	0.271621	0.219386	0.118131	0.042406	0.009786	0.001317	7.88E-05
p_7	n=9 p=0.42	0.007428	0.048408	0.140216	0.236916	0.25734	0.18635	0.089962	0.027919	0.005054	0.000407
p_8	n=9 p=0.38	0.013537	0.074672	0.183068	0.261806	0.240693	0.147521	0.060278	0.015833	0.002426	0.000165

The Lines Represent $p = 0.95$ Level Efficient Points

64	68	57	73	55	45	55	65
64	73	62	78	50	45	50	65
69	68	57	78	50	40	55	60
69	68	62	78	60	40	50	55
69	73	57	73	55	40	55	55
69	78	52	78	55	55	50	55
69	78	57	73	50	40	50	70
74	68	57	73	50	50	50	60
74	73	62	73	55	40	55	55
79	73	52	73	55	40	50	60
79	78	52	73	60	40	50	60
79	78	62	78	60	55	60	70

l = Vector of Lower Bounds

u = Vector of Upper Bounds

$l = (5, 5, 10, 5, 15, 5, 5, 5)^T$

$u = (10, 10, 15, 10, 20, 10, 10, 10)^T$

Number of Gale–Hoffman Inequalities	$2^8 - 1 = 255$
After Elimination by Network Structure	161
By Upper Bounds	116
By Lower Bounds	63
By Linear Programming	16

The Optimization Problem

$$\text{Min} \sum_{i=1}^{8} x_i$$

subject to

$$P\left(\sum_{i \in I_j} \left(\xi_i - x_i\right) \le a_j, \quad j = 1, \ldots., 16 \right) \ge 0.95$$

$$l_i \le x_i \le u_i, \quad i = 1, \ldots., 8$$

I_j = Set of Subscripts of $\xi_i - x_i$ in jth Remaining Inequality

a_j = Sum of Capacities in the jth Remaining Inequality

Seven 0.95 – Level Efficient Points Have Been Generated

The Optimal Solution is

$x^* = (8, 6, 14, 9, 16, 8, 6, 8)^T$,

None of the Bounds is Binding.

References

1. G. Boole (1854). Laws of Thought. American Reprint of 1854 Edition, Dover, New York.
2. E. Boros, A. Prékopa (1989). Closed Form Two-Sided Bounds for Probabilities that Exactly *r* and at Least *r* out of *n* Events Occur. *Math. Oper. Res*. **14**, 317–342.
3. E. Boros, K. Elbassioni, V. Gurvich, L. Khachiyan, K. Makino (2003). An Intersection Inequality for Discrete Distributions and Related Generation Problems. Lecture Notes in Computer Science 2719, Springer, Berlin-Heidelberg, 543–555.
4. D. A. Dawson, A. Sankoff (1967). An Inequality for Probabilities. *Proc. Amer. Math. Soc*. **18**, 504–507.
5. D. Dentcheva, A. Prékopa, A. Ruszczynski (2000). Concavity and Efficient Points of Discrete Distributions in Probabilistic Programming. *Math, Prog. Series A*, **89**, 55–77.
6. D. Gale (1957). A Theorem on Flows in Networks. *Pacific J. of Math*. **7,** 1073–1082.

7. Th. Hailperin (1965). Best Possible Inequalities for the Probability of a Logical Function of Events. *Amer. Math. Month*. **72**, 343–359.
8. A. J. Hoffman (1960). Some Recent Applications of the Theory of Linear Inequalities to Extremal Combinatorial Analysis. In: *Proceedings of Symposia of Applied Math*. Vol. X. American Math. Soc., 113–117.
9. D. Hunter (1976). An Upper Bound for the Probability of a Union. *J. Appl. Probab*. **13**, 597–603.
10. S. M. Kwerel (1975a). Most Stringent Bounds on Aggregated Probabilities of Partially Specified dependent Probability Systems. *J. Amer. Statist. Assoc*. **70**, 472–479.
11. S. M. Kwerel (1975b). Bounds on Probability of a Union and Intersection of *m* Events. *Adv. Appl. Probab*. **7**, 431–448.
12. A. Prékopa (1980). Network Planning Using Two-Stage Programming Under Uncertainty. Lecture Notes in Economics and Math. Systems **179**, 216–237.
13. A. Prékopa (1988). Boole–Bonferroni Inequalities and Linear Programming. *Oper. Res*. **36**, 145–162.
14. A. Prékopa (1995). *Stochastic Programming*. Kluwer.

15. A. Prékopa, E. Boros (1989). On the Existence of a Feasible Flow in a Stochastic Transportation Network. *Oper. Res*. **39**, 119–129.
16. A. Prékopa, T. Szántai (1978). Flood Control Reservoir System Design Using Stochastic Programming. *Math. Prog. Study* **9**, 138–151.
17. A. Prékopa, T. Rapcsák, I. Zsuffa (1978). Serially Linked Reservoir System Design Using Stochastic Programming. *Water Resources Research* **14**, 672–678.
18. A. Prékopa, B. Vizvári, T. Badics (1998). Programming under Probabilistic Constraint with Discrete Random Variable. In: *New Trends in Mathematical Programming* (F. Giannessi et al. eds.), Kluwer, 235–255. See also: RUTCOR Research Report 10–96.
19. Term project works by Ünüvar, Yazici (RUTCOR), Z. Csizmadia (ELTE, Budapest).
20. S. W. Wallace, R. J.-B. Wets (1993). The Facets of the Polyhedral Set Determined by the Gale–Hoffman Inequalities. *Math. Prog*. **62**, 215–222.
21. M. A. Yazici, K. Özbay (2007). Determination of Hurricane Evacuation Shelter Capacities and Locations with Probabilistic Road Capacity Constraints. 86th Annual Meeting of the Transportation Research Board, Washington, D. C., 2007.