Stochastic Programming Models with Decision Dependent Probabilities

David L. Woodruff
Graduate School of Management
University of California, Davis

October 2003

Outline

- **Overview of the modeling issues**
- **Beyond the state of the art**
- **A more complicated situation**
- **(if there is time) A more abstract view**

The Issues

- **In some settings, the time at which information is obtained depends on the decisions.**

- **(This is *not* the case in most financial markets, for example.)**

- **The classic example is oil exploration, but other examples involve costs and capabilities of new technologies.**

- **Although extension to real variables is possible, all work to date has focused on integer decisions that effect discovery timing.**

The Modeling Issues

1. **Consider the case where a binary variable, in addition to other effects, determines the timing of information discovery.**
2. **Semantically, all that is needed for modeling is to indicate which random elements are resolved at which times as a function of the decision variables.**
3. **Example: Production costs become known when an item is first produced.**

For an abstract view, consider a few words from Jonsbråten, Wets and Woodruff, “A Class of Stochastic Programs with Decision Dependent Random Elements”

‘Standard’ Stochastic Programming:

$$\min_{x \in R^n} \; E\{f(\boldsymbol{\xi}; x)\} = Ef(x) \tag{1}$$

where $f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}} = [-\infty, \infty]$ is the ‘cost’ associated with a decision x when the random variable $\boldsymbol{\xi}$ takes on the value ξ; $\boldsymbol{\xi}$ is a $\mathbb{R}^k$-valued random variable with possible values in $\Xi \subset \mathbb{R}^k$, which is the support of the distribution, μ, of the random variable; $Ef : \mathbb{R}^n \to \overline{\mathbb{R}}$, the function to be minimized, is defined by

$$Ef(x) = \int_{\Xi} f(\xi; x)\, \mu(d\xi).$$

One can recast multi-stage stochastic programs with recourse so that they are seen as special cases of the problem just formulated.

However, there are important decision making problems that do not fit in this mold, namely cases when the distribution of the random quantities will be affected by the decision selected. This can happen in many ways, but it seems the following formulation would cover all such cases:

$$\min \boldsymbol{E}^{\mu} f(x) = \int_{\Xi} f(\xi; x)\, \mu(d\xi)$$

such that

$$(\mu, x) \in \mathcal{K} \subset M \times \mathbb{R}^n$$

where M is a subset of the probability measures on Ξ and $\mathcal{K}$ are the constraints linking the decision x to the choice of μ.

In the literature devoted to Discrete Events Dynamical Systems, the dependence of the probability measure on the decision(s) has often received the following formulation:

$$\min_{x \in R^n} \int_{\Xi} f(\xi, x)\, \mu_x(d\xi).$$

In such a situation, the set $\mathcal{K}$ is the graph of the mapping $x \mapsto \mu_x$, i.e.,

$$\mathcal{K} = \left\{ (\mu, x) \,\middle|\, x \in I\!R^n, \mu = \mu_x \right\}$$

Simple SMPS Modification

- **The usual stoch file serves as the default case**

- **Simple bounds to define decision value sets corresponding to additional stoch files are given in the header of these files.**

- **Solvers exist only for very special cases. (See, e.g., Jonsbråten et al and also Vikas Goel and Ignacio E. Grossmann, “A Stochastic Programming Approach to Planning of Offshore Gas Field Developments under Uncertainty in Reserves”)**

One advantage of our more general formulation is that it allows for a better classification of problems of this type based on the properties of the set $\mathcal{K}$ of the linking constraints.

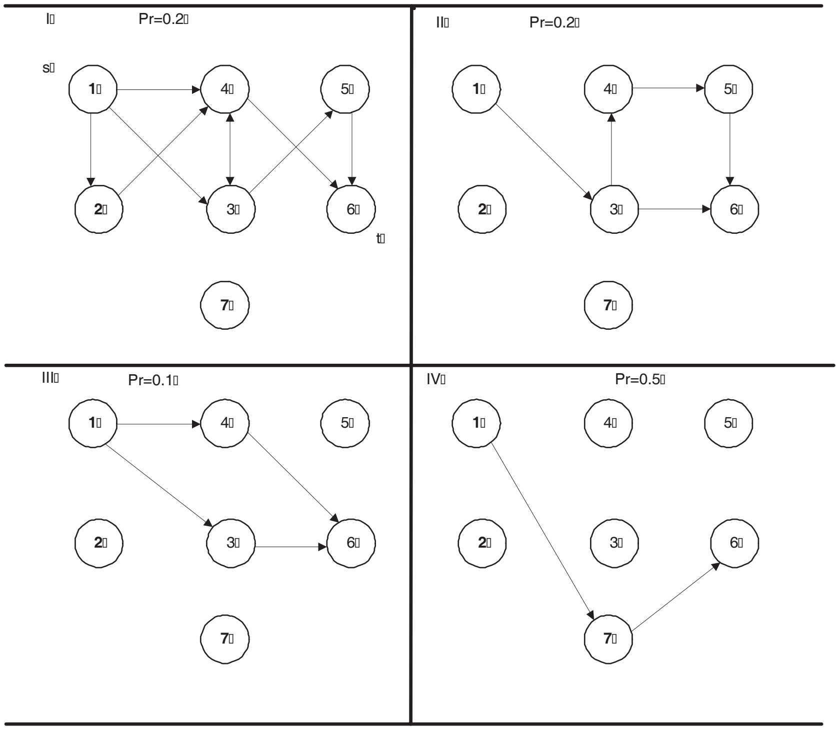
Multi-stage Network Interdiction

- **Interdict**
- **Observe flow**
- **Interdict**

- **Decision dependent random elements!!**
  (And the linkage is unusual.)

Notation

- **Problem: interdict the flow of information or goods in a network (N, A) with uncertain characteristics (for example computer, terrorist or drug transportation networks)**
- **goal: maximize the minimum distance between a node s and a node t**


I
Pr=0.2
s
1
4
5
2
3
6
t
7
II
Pr=0.2
1
4
5
2
3
6
7
III
Pr=0.1
1
4
5
2
3
6
7
IV
Pr=0.5
1
4
5
2
3
6
7

Conclusions So Far

- **Just beyond the state-of-the-art frontier in stochastic programming lie problems with decision dependent random elements.**

- **Some classes of problems can be expressed in fairly straightforward manner, at least in theory.**

Given data:

- node-arc incidence matrix G which includes an artificial arc (t, s)
- a vector c which contains the arc distances
- a vector d which contains the rates an arc is lengthened if chosen for interdiction
- binary decision variables $x \in X$ (system of linear budget constraints)

One Stage Deterministic Version

$$\max_{x \in X} \; \min_{y} \sum_{k \in A} (c_k + d_k x_k) y_k$$

subject to:

$$\begin{aligned} Gy &= 0, \\ y_{ts} &= 1, \\ y_k &\geq 0, \; k \in A. \end{aligned}$$

Or, using the dual of the inner problem:

$$\max_{x \in X} \; \max_{\pi} \pi_t - \pi_s$$

subject to:

$$\begin{aligned} -\pi_i + \pi_j &\leq c_{ij} + d_{ij} x_{ij}, \;\; (i,j) \in A \\ \pi_s &= 0. \end{aligned}$$

Stochastic Data

Let $\boldsymbol{\Omega}$ be a set of scenarios with probabilities $\boldsymbol{Pr(\omega)}\ \forall \boldsymbol{\omega} \in \boldsymbol{\Omega}$.

- the notation $(\boldsymbol{N}, \boldsymbol{A})$ is used to refer to

  $$\boldsymbol{N} = \bigcup_{\omega \in \Omega} \boldsymbol{N}(\boldsymbol{\omega}) \quad \text{and} \quad \boldsymbol{A} = \bigcup_{\omega \in \Omega} \boldsymbol{A}(\boldsymbol{\omega})$$

- for arcs $\boldsymbol{k} \notin \boldsymbol{A}(\boldsymbol{\omega})$ we set $\boldsymbol{c_k}(\boldsymbol{\omega}) = \infty$ and $\boldsymbol{d_k}(\boldsymbol{\omega}) = \boldsymbol{0}$

Stochastic Formulation

Maximize the probability that the minimum length path from $\boldsymbol{s}$ to $\boldsymbol{t}$ exceeds $\boldsymbol{\varphi}$.

Resulting problem:

$$\boldsymbol{\max_{x \in X} \; Pr([\max_{\pi} \pi_t(\omega) - \pi_s(\omega)] \geq \varphi)}$$

subject to:

$$\boldsymbol{\begin{aligned} -\pi_i(\omega) + \pi_j(\omega) \; &\leq \; c_{ij}(\omega) + d_{ij}(\omega) x_{ij}, (i,j) \in A, \omega \in \\ \pi_s(\omega) \; &= \; 0, \; \omega \in \Omega. \end{aligned}}$$

We can solve this.

With two interdiction attempts the formulation becomes a bit longer:

$$\max_{x^{(1)}\in X^{(1)}} \ \mathbb{P}(\{\omega\in\Omega:\ \min_{y^{(2)}}\sum_{k\in A}[c_k(\omega)+d_k(\omega)[x_k^{(1)}+x_k^{(2)}(y^{(1)})]]y_k^{(2)}(\omega)\geq\varphi \tag{2}$$

subject to

$Gy^{(2)}(\omega)=0,\ \omega\in\Omega,$
$y_{ts}^{(2)}(\omega)=1,\ \omega\in\Omega,$
$y_k^{(2)}(\omega)\geq 0,\ k\in A,\ \omega\in\Omega\});$

$$y^{(1)}=y^{(1)}(x^{(1)},\omega)\in \ \operatorname*{argmin}_{y} \ \sum_{k\in A}(c_k(\omega)+d_k(\omega)x_k^{(1)})y_k \ \forall\omega\in\Omega \tag{3}$$

subject to

$Gy=0,$
$y_{ts}=1,$
$y_k\geq 0,\ k\in A;$

$$x^{(2)}(y^{(1)})\in \ \operatorname*{argmax}_{x\in X^{(2)}} \ \mathbb{P}^{(2)}_{y^{(1)}}(\{\omega\in\Omega:\ \min_{y}\sum_{k\in A}[c_k(\omega)+d_k(\omega)[x_k^{(1)}+x]]y(\omega)\geq\varphi \tag{4}$$

subject to

$Gy(\omega)=0,\ \omega\in\Omega,$
$y_{ts}(\omega)=1,\ \omega\in\Omega,$
$y_k(\omega)\geq 0,\ k\in A,\ \omega\in\Omega\}).$

With

$$\mathbb{P}^{(2)}_{y^{(1)}}(\omega)=\frac{1}{\sigma}\mathbb{P}(\omega)\chi_{y^{(1)}}(\omega),\ \forall\omega\in\Omega$$

$$\sigma:=\sum_{\omega\in\Omega}\mathbb{P}(\omega)\chi_{y^{(1)}}(\omega),\ \omega\in\Omega,$$

$$\chi_{y^{(1)}}(\omega):=\left\{\begin{array}{ll} 0, & \text{if}\ \ y_a^{(1)}=1 \ \ \text{for an arc}\ \ a\notin A(\omega),\\ 1, & \text{otherwise}\end{array}\right.,\ \omega\in\Omega.$$

First Stage Enumeration Algorithm:

Assume that the budget $B^{(1)}$ for the first-stage decision equals 1. For each node n do:

- **Consider the network resulting after interdicting node n.**
- **Compute a shortest path $y^{(1)} = y^{(1)}(\omega)$ for each scenario $\omega \in \Omega$; (i.e., "observe" the flow in each scenario.**
- **Obtain a third-stage decision $x^{(2)}(y^{(1)})$ for each such $y^{(1)}$ by solving the two-stage problem with budget $B^{(2)}$ and maybe a smaller set of scenarios (i.e. if a component of $y^{(1)}$ corresponding to an arc a is 1 but a does not exist in another scenario $\tilde{\omega}$, $\tilde{\omega}$ can be removed from Ω).**
- **Compute the shortest path lengths $L(\omega)$ for each scenario ω through the network that results after interdicting node n and nodes corresponding to $x^{(2)}$.**
- **Compute**

$$z(n) := \mathbb{P}(\{\omega \in \Omega : \; L(\omega) \geq \varphi\}.$$

The node n that yields the largest value $z(n)$ corresponds to the optimal $x^{(1)}$.

This generalizes easily to larger first stage budgets.

Technical issue: multiple solutions to inner problems

Consider an optimizing network operator?

The Final Words

1. **Although some classes of decision dependent stochastic optimization can be expressed in a straightforward manner, others seem more complicated.**

2. **Development must be iterative as modeling and solution methods evolve together, thereby enabling a model taxonomy relevant for solvers.**