Stochastic Programming in Enterprise-Wide Optimization

Andrew Schaefer
University of Pittsburgh
Department of Industrial Engineering
October 20, 2005

Outline

- What is stochastic programming?
- How do I formulate a stochastic program?
- Have these been used in industry?
- How can I solve a stochastic program?
- What about extensions?
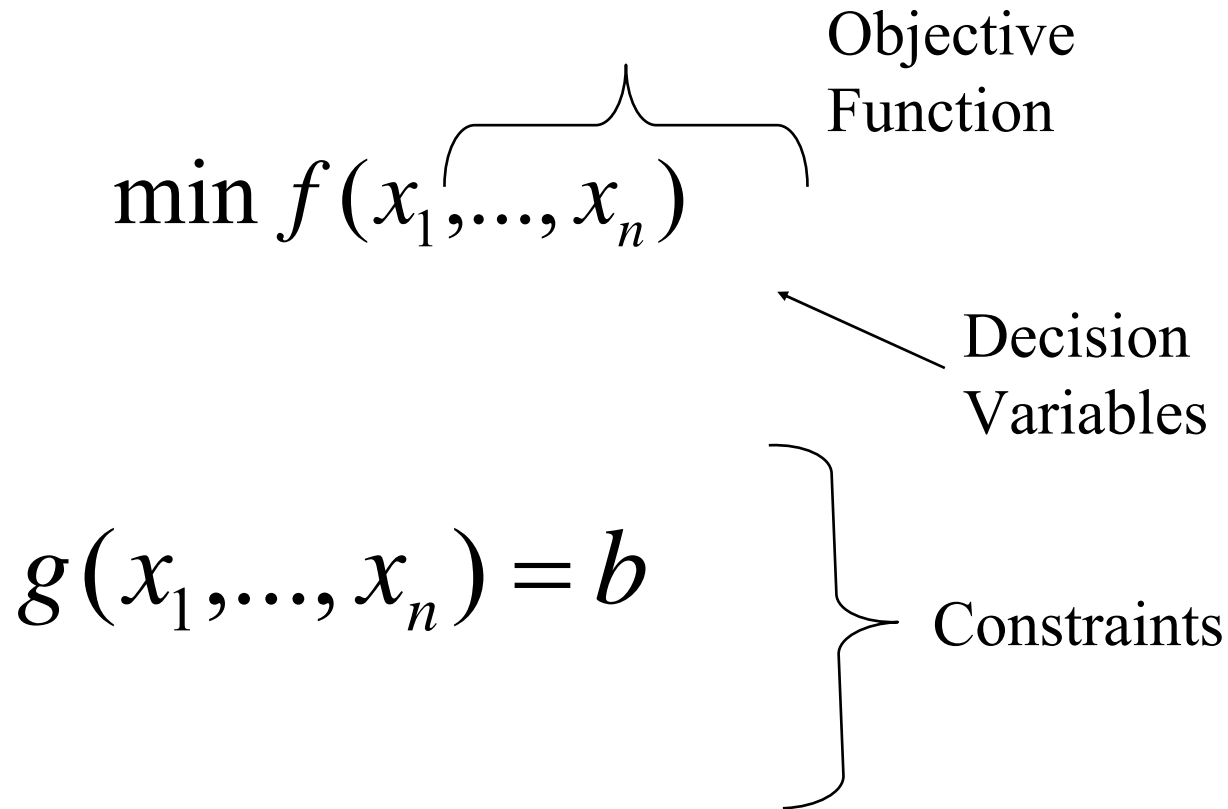
Deterministic Optimization

- Two broad categories of optimization models exist
  - **Deterministic**
    - Parameters/data are known with certainty
  - **Stochastic**
    - Parameters/data are known with uncertainty

  It may be helpful to think of deterministic models as a special case of stochastic models

Deterministic Optimization

- Components
  - Decision variables
  - Objective function
  - Constraints

General Model

$$\min f(x_1, \ldots, x_n)$$

Objective Function

Decision Variables

$$g(x_1, \ldots, x_n) = b$$

Constraints

Deterministic Optimization Models

- Linear Programs
- Nonlinear Programs
- Integer Programs
- Dynamic Programs

Shortcomings of Deterministic Optimization

- Deterministic optimization requires being certain about the parameters/data
  - In reality, we are almost never certain about the parameters/data of the problem.
  - Exact data is unavailable or expensive
- Tradeoff between validity and tractability
  - Stochastic models produce more valid results
  - Deterministic models are easier to solve; typically can have finer granularity than stochastic optimization models

What is Stochastic Programming (SP)?

• A Stochastic Program is a mathematical program in which some of the parameters defining a problem instance are random
• A stochastic linear program (SLP) is the simplest case of stochastic program

Assumptions in SP:

• The probability distribution for the uncertainty is known or can be approximated
• The probabilities are independent of the decisions that are taken (with few exceptions; see Goel dissertation 2005)

Formulating Stochastic Programs

Stages – time periods

The sequence of events and decisions in a two-stage SLP with recourse (2SLPwR):

$$x \; \rightarrow \; \text{Realization of } \omega \; \rightarrow \; y(\omega)$$

•<u>First-stage decisions</u> (x) are taken before the uncertainty ω is realized (known)

•<u>Second-stage decisions</u> (y(ω)) are taken as corrective actions after the actual value of ω becomes known

Standard Objective: minimize first-stage cost plus **expected** second-stage costs

Formulating Stochastic Programs

- Familiar Linear Program (LP):

$$\min \ c^T x$$
$$\text{s.t. } Ax = b,$$
$$x \geq 0$$

- 2SLPwR

$$\min \ c^T x + Q(x)$$
$$\text{s.t. } Ax = b$$
$$x \geq 0$$

where $Q(x)$, the *expected recourse function*, gives the expected cost of the optimal second-stage decision given first-stage decision x

What can be random?

- In the second stage, the objective, constraints and right-hand sides are permitted to be random. Also, the matrix describing the relationship between x and y

- **For scenario ω:**
  - Objective: q(ω)
  - Right-hand side: h(ω)
  - Constraint (*recourse*) matrix: W(ω)
  - Matrix relating x to y (*technology matrix)*: T(ω)

Expected Recourse Function

- How do we define $Q(x)$?

- For any scenario ω, Define $Q(x,\omega)$ by
  $Q(x,\omega) = \min q(\omega)^t y(\omega)$
  s.t. $W(\omega)y(\omega) = h(\omega) - T(\omega)x$
  $\qquad\qquad y(\omega) \geq 0$

And $Q(x) = E_\omega Q(x,\omega)$

Notice that the relationship between x and $y(\omega)$ must be described linearly

"Extensive-Form" Two-stage SLPs with Recourse

- We could solve a 2SLPwR as a (really big) linear program, called the *extensive form*

$$\text{Min } z = c^T x + E_\omega[\min q(\omega)^T y(\omega)]$$

s.t.

$$Ax = b$$
$$T(\omega)\, x + W(\omega)\, y(\omega) = h(\omega) \text{ for all } \omega$$
$$x \geq 0,\ y(\omega) \geq 0$$

Extensive Form of a Two-stage Stochastic Linear Program

$$
\begin{aligned}
\text{Min} \quad & c^\top x + p_1 q_1^\top y_1 + p_2 q_2^\top y_2 + \cdots + p_s q_s^\top y_s \\
\text{s.t.} \quad & Ax && = b \\
& T_1 x + W_1 y_1 && = h_1 \\
& T_2 x + \quad W_2 y_2 && = h_2 \\
& \vdots \ + \qquad\qquad \cdot && = \vdots \\
& \vdots \ + \qquad\qquad\quad \cdot && = \vdots \\
& \vdots \ + \qquad\qquad\qquad \cdot && = \vdots \\
& T_S x + \qquad\qquad W_s y_s && = h_s \\
& x \geq 0,\ y_1 \geq 0,\ y_2 \geq 0, \ldots\ldots,y_s \geq 0
\end{aligned}
$$

Properties of the Expected Recourse Function

- What do we know about Q(x)?

- Q(x) is convex (whew!)
  - But not if $y(\omega)$ must be integer-valued
  - This convexity is critical for solving SLPs

- If there are a finite number of scenarios, Q(x) is piece-wise linear

Expected Recourse Function
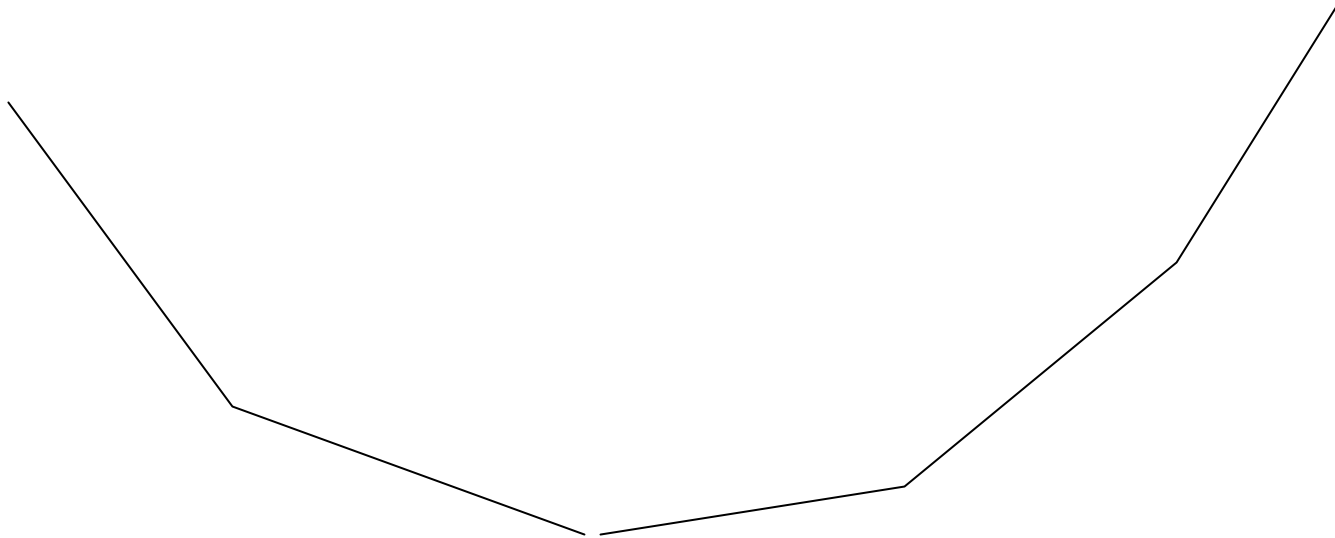
The expected recourse function Q(x) is convex and, if there are a finite number of scenarios, is also piece-wise linear

SP Applications

- SP deals with a class of optimization models and algorithms

  – some of the data may be subject to significant uncertainty

- Such models appropriate when

  – data evolve over time
  – decisions need to be made prior to observing the entire data stream

SP Applications - Examples

- Investment decisions in portfolio planning:
  - Portfolio planning Problems must be implemented before stock performance can be observed

- Power Generation:
  - Utilities must plan power generation before the demand for electricity is realized

Optimizing electricity distribution using two-stage integer recourse models

- Consider two planning problems faced by an (Dutch) electricity distributor
  - Every year a contract with the power plants.
  - Determine a supply schedule for each single day.

- Electricity:
  - 1- Power plant
  - 2- Small generators (i.e. hospitals, greenhouses, industrial consumers)

Willem K. Klein Haneveld and Maarten H. van der Vlerk (May 2000)

Optimizing electricity distribution using two-stage integer recourse models

- Future demand for electricity is uncertain
- Price
  - Depends on quota
  - Yearly contract
- Various constraints on switching generators on/off
- Two-stage integer recourse models.
  - Very hard to solve in general
  - Develop tailor-made solution methods
  - Use of valid inequalities and Lagrange relaxation

IBM Research in SP

- Duality and martingales: a stochastic programming perspective on contingent claims

  - The hedging of contingent claims in the discrete time, discrete state case
  - Modeling the hedging problem as an SP
  - Model easily extends to the analysis of options pricing
  - An extension: Incorporate pre-existing liabilities and endowments ⟹ why buyers and sellers trade in options

Uncertainty Modeling & Management in MRP Systems

- Examine an MRP environment with
  - Demand and supply uncertainty
  - Capacity limits
  - Service level requirements
- Static, finite horizon SP
- Feasible instances → Develop optimal / heuristic solutions
- Infeasible instances → Identify good, feasible combinations of capacity and service

Ramesh Bollapragada - Lucent Technology Bell Labs

Uday S. Rao   - Carnegie Mellon University

The Russell-Yasuda Kasai Model:

- Frank Russell Company and the Yasuda Fire and Marine Insurance Co. Ltd.

- An asset/liability model

- For a Japanese insurance company

- Multistage stochastic programming

David R. Carino; Terry Kent; David H. Myers; Celine Stacy; Mike Sylvanus; Andrew L. Turner; Kouji Watanabe; William T. Ziemba
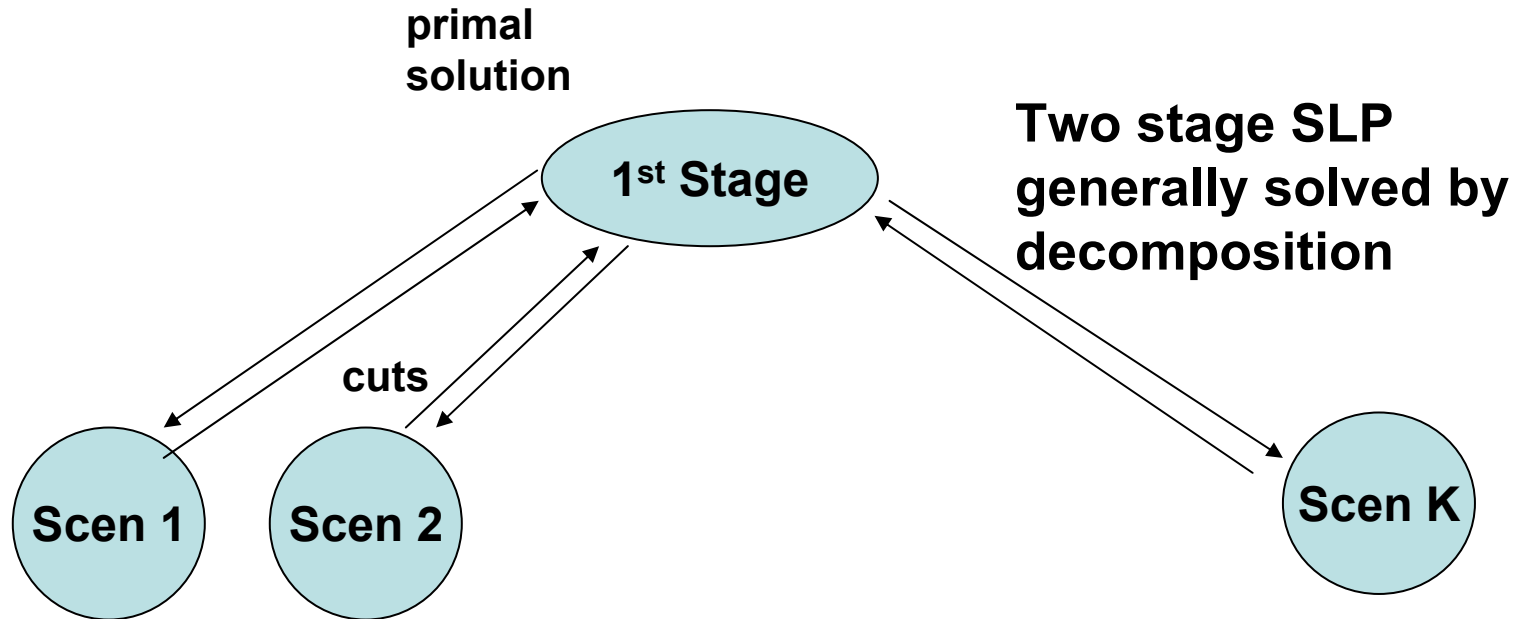
The Russell-Yasuda Kasai Model:

- Model ⟹ Optimal investment strategy

  – Integrates a multi period method

  – Allows decision makers to describe risks in tangible operational terms.

- Generate an extra revenue of 42 basis points (about $79 million during the first two years of its use).

Some Other Applications

- Electric power generation
  - Murphy, F.H., Sen, S., and Soyster, A.L. (1982). "Electric utility capacity expansion planning with uncertain load forecasts." AIIE Transaction 14, 52-59.

- Supply chain management
  - Fisher, M., J. Hammond, W. Obermeyer, and A. Raman [1997]. "Configuring a supply chain to reduce the cost of demand uncertainty," Production and Operations Management, 6,pp.211-225.

- Telecommunications network planning
  - Sen, S. R.D. Doverspike and S. Cosares [1994]. "Network Planning with Random Demand," Telecommunication Systems, 3, pp. 11-30.

How can I solve Two-Stage Stochastic Linear Programs?

primal solution

Two stage SLP generally solved by decomposition

1st Stage

cuts

Scen 1

Scen 2

Scen K

How do we solve the simplest form of Stochastic Programs?

- Linear Approaches

  - The L-Shaped Method (most common)
    - (Single cut & Multicut version)
  - Inner Linearization Methods (won't discuss)
  - Basis Factorization Methods (won't discuss)

The L-Shaped Method

- The most commonly used technique.
- Basic idea: To approximate the convex term in the objective function.
- Recourse function involves a solution of all second stage recourse linear programs, we want to avoid numerous function evaluations for it.
- Therefore, divide the problem into two stage:
  - Master Problem
  - Sub-problems
- Converges to the optimal solution in finite steps with adding 2 new type of constraints to the master problem called feasibility cuts and optimality cuts.
- Feasibility cuts – ensure that the second-stage problems are feasible
- Optimality cuts – relate the second-stage costs to first-stage constraints

L-shaped Restricted Master Problem

Master problem is as follows:

min $\quad z = c^T x + \theta$

s. $\quad Ax = b$

t $\quad D_\ell x \geq d_\ell \qquad \ell = 1, \ldots, s \qquad$ **Feasibility cuts**

$\quad E_\ell x + \theta \geq e_\ell \quad \ell = 1, \ldots, r \qquad$ **Optimality cuts**

$\quad x \geq 0 \quad \theta \in \Re$

L-shaped Optimality Subproblem

Recourse problem for scenario k given first-stage x is as follows:

$$\min \ w = q_k^T y$$

$$\text{s.t} \ \ Wy = h_k - T_k x$$

$$y \geq 0$$

Let π denote the duals. Then $E_l = \pi^T T_k$

and $e_l = \pi^T h_k$

Algorithm

At each iteration:

Step 1: Solve the master problem

Step 2: If the solution to the master problem (x*) leads to feasible recourse problems for all scenarios,
- Go to step 3
- else add a *FEASIBILITY CUT* and go to step 1 and solve master problem again.

Step 3: If the expected value for the optimal values of the recourse problems (w*) is greater than Θ obtained in step 1
- Stop the current solution is optimal,
- else add an *OPTIMALITY CUT* and go to step 1 and resolve master problem.

Expected Recourse Function
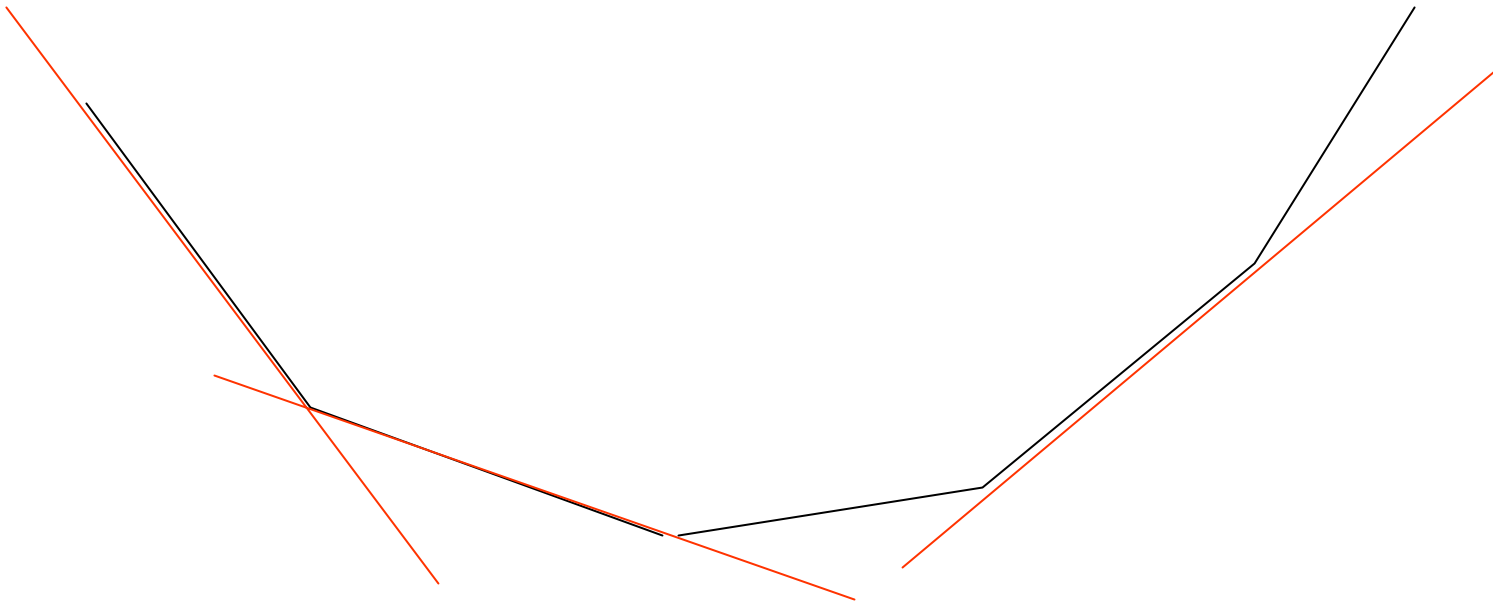
The expected recourse function Q(x) is convex and, if there are a finite number of scenarios, is also piece-wise linear

.L-shaped optimality cuts support Q(x) from below. If there are a finite number of scenarios, there are a finite number of optimality cuts

Multicut Version

- Instead of using Θ to represent $Q(x)$, we use Θ^k to represent $Q(x,\omega^k)$

- We may add as many as K cuts per iteration, but we get more information

- In general, multicut works well if there aren't "too many" scenarios

- For integer first stage, multicut is not competitive with single cut

Stochastic Integer Programs

- When recourse problem is an MIP, Q(x) is nonconvex and discontinuous
- The absence of general efficient methods reflects this
- Some techniques have been proposed that address specific problems or use a particular property
- Much work needs to be done to solve SIPs efficiently
- No industrial-sized problems with integer recourse have been solved thus far
- The field is expected to evolve a great deal in the future

How to solve SIPs?

- A set of valid feasibility cuts and optimality cuts, which are based on duality theory of linear programming, is known to exist in the continuous case and forms the basis of the classical L-shaped method

- They can also be used in the case where only the first-stage variables contain some integrality restrictions

- The most common method to solve SIPs is the so-called *Integer L-Shaped Method*

A little history

- The first application of the integer L-shaped method was proposed by Laporte and Louveaux (1993) for the case of binary first- and second-stage variables

- A full characterization of the method based general duality theory is given by Carøe and Tind (1996)

- A stochastic version of the branch and cut method (stochastic branch and bound) using statistical estimation of the recourse function instead of exact evaluation is given by Norkin, Ermoliev, and Ruszczyński (1997)

Simple Integer Recourse

- When a stochastic program has *simple recourse*, the only feasible solutions are to pay a shortage or surplus penalty

- With *simple integer recourse*, the recourse variables must take integer values

- These are the best understood SIP models

An SIP with a special structure (Simple integer recourse)

- 2-stage SP with simple recourse can be transformed into :

$$\min \quad c^T x + \sum_{i=1}^{m} \psi_i(\chi_i)$$

$$\text{s.t.} \quad Ax = b,$$

$$Tx = \chi,$$

$$x \in X \subset Z_+^{n_1},$$

where $\quad \psi_i(\chi_i) = q_i^+ u_i(\chi_i) + q_i^- v_i(\chi_i)$

with $\quad u_i(\chi_i) = E\lceil \xi_i - \chi_i \rceil^+ \quad$ defined as the expected shortage

$\quad\quad\quad v_i(\chi_i) = E\lceil \chi_i - \xi_i \rceil^+ \quad$ defined as the expected surplus

SIP with Stochastic Right-hand Sides

- Kong et al. considered two-stage SIPs where the randomness was on the r.h.s.

- They reformulated the problem using the superadditive dual of IPs in both stages

- They developed methods for finding these superadditive duals

- They were able to solve relatively large problems (D.E. ~ 10^{10} columns, 10^7 rows)

Solving SIPs via Global Optimization

- Ahmed, Tawarmalani and Sahinidis (2004) used a similar reformulation and applied global optimization techniques
- Their model could handle stochasticity in the objective function

Other approaches to solving SIPs

- Extensive forms & decomposition
- Asymptotic analysis / approximation
- Markov Chains
- Dynamic programming
- Lagrangean decomposition

Multistage SP with Recourse

- Involves a sequence of decisions that react to outcomes that evolve over time

- The extensive form of an N-stage fixed-recourse problem:

$$\min \quad z = c^1 x^1 + E_{\xi^2}[c^2(\omega)x^2(\omega^2) + ... + E_{\xi^N}[c^N(\omega)x^N(\omega^N)]...]$$

s.t.

$$W^1 x^1 = h^1$$

$$T^1(\omega)x^1 + W^2 x^2(\omega^2) = h^2(\omega)$$

$$\vdots$$

$$T^{N-1}(\omega)x^{N-1}(\omega^{N-1}) + W^N x^N(\omega^N) = h^N(\omega)$$

$$x^1 \geq 0, x^t(\omega^t) \geq 0, t = 2,...,N-1$$

where ω^t denotes the history up to time t

Multistage SP with Recourse

To obtain the deterministic equivalent form

- Let

$$Q^N(x^{N-1}, \xi^N(\omega)) = \min c^N(\omega) x^N(\omega^N)$$

$$\text{s.t.} \quad W^N x^N(\omega) = h^N(\omega) - T^{N-1}(\omega) x^{N-1}$$

$$x^N(\omega) \geq 0$$

- Also, let

$$\vartheta^{t+1}(x^t) = E_\xi [Q^{t+1}(x^t, \xi^{t+1}(\omega)]$$

for all t and obtain the recursion for $t=2,\ldots,N-1$

$$Q^t(x^{t-1}, \xi^t(\omega)) = \min c^t(\omega) x^t(\omega) + \vartheta^{t+1}(x^t)$$

$$\text{s.t.} \quad W^t x^t(\omega) = h^t(\omega) - T^{t-1}(\omega) x^{t-1}$$

$$x^t(\omega) \geq 0$$

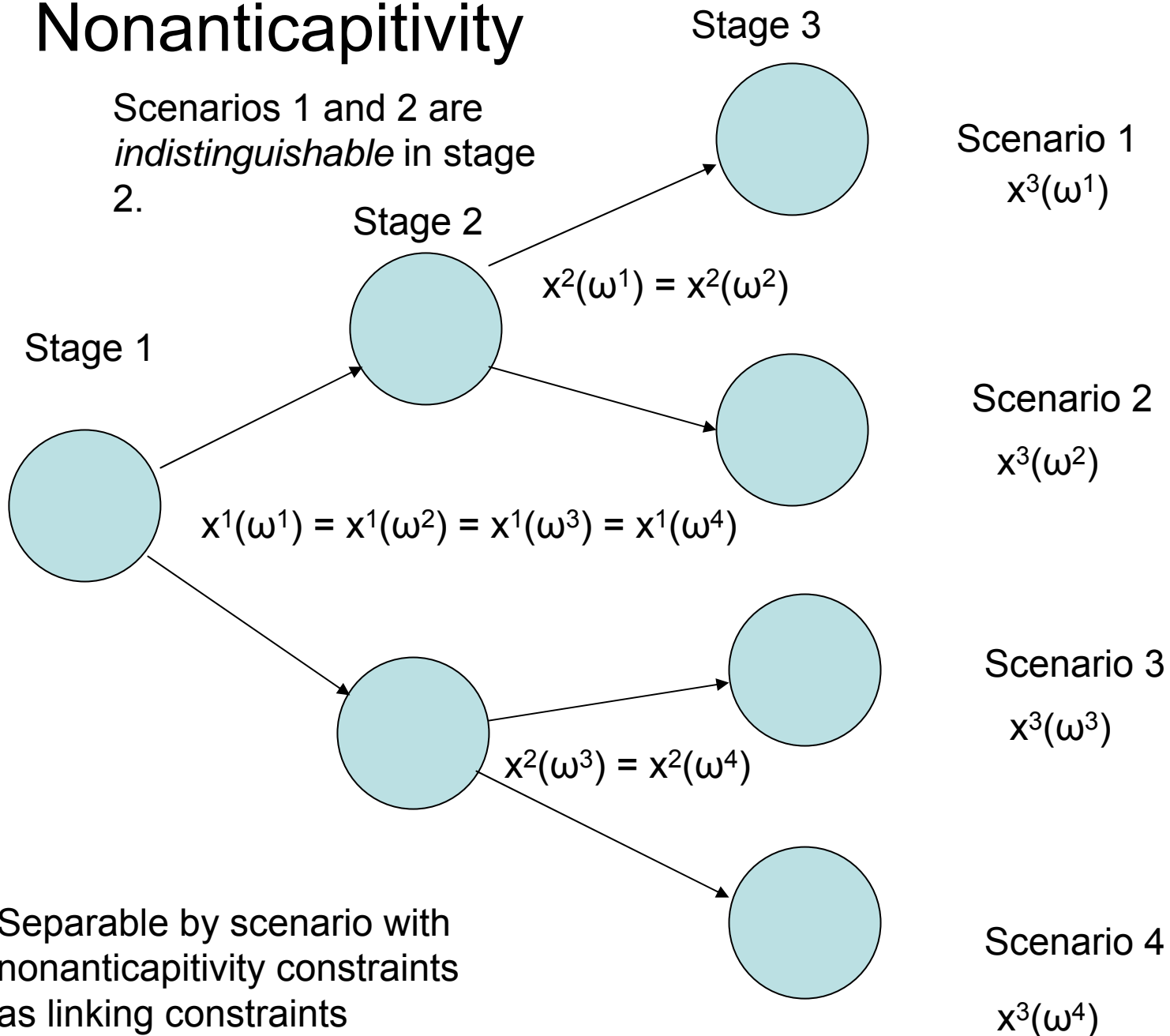# Multistage SP with Recourse

Deterministic Equivalent:

$$\min z = c^1 x^1 + \vartheta(x^1)$$
$$\text{s.t.} \quad W^1 x^1 = h^1$$
$$x^1 \geq 0$$

Nonanticapitivity

Scenarios 1 and 2 are *indistinguishable* in stage 2.

Stage 3

Stage 2

Stage 1

$x^2(\omega^1) = x^2(\omega^2)$

$x^1(\omega^1) = x^1(\omega^2) = x^1(\omega^3) = x^1(\omega^4)$

$x^2(\omega^3) = x^2(\omega^4)$

Scenario 1
$x^3(\omega^1)$

Scenario 2
$x^3(\omega^2)$

Scenario 3
$x^3(\omega^3)$

Scenario 4
$x^3(\omega^4)$

Separable by scenario with nonanticapitivity constraints as linking constraints

Solving SPs with Nonanticipativity

- Such a model is decomposable by scenario, where nonanticipativity constraints are linking constraints

- Lagrangian relaxation of linking constraints

- For reasonably large scenario trees, the number of possible nonanticipativity constraints is enormous

Solving Multistage SPs

Nested Decomposition

- Built on the two-stage L-shaped method

- Extended to the multistage case by Birge

- The idea is to place cuts on $\vartheta^{t+1}(x^t)$ and to add other cuts to achieve an x^t that has a feasible completion in all descendant scenarios

- Successive linear approximations of ϑ^{t+1}

- Due to the polyhedral structure of ϑ^{t+1}, the process converges finitely

Continuous Distributions

- How can stochastic programming handle continuous distributions?

- With a continuous distribution, there are infinitely many scenarios (Σ becomes $\int$)

- The extensive form formulation is infinite-dimensional

- However, there are ways to overcome these difficulties

Sampling Approach

- One approach is to sample from the continuous distribution (ω^1, …, ω^N)

- Assign each scenario probability 1/N

- As N $\rightarrow \infty$, get optimal solution

- How big should N be? Use statistical properties to estimate convergence to optimal solution

Stochastic Decomposition

- Works for 2SLPwR (Higle & Sen)
- Sample to create cuts on each iteration
- These cuts fall off (are given less weight) as the algorithm progresses

Conclusions

- Stochastic programming is an emerging filed that extends classical optimization to stochastic and dynamic settings

- SP has significant potential in EWO

- Stochastic programs are difficult to solve, so modeling becomes challenging

- Despite their inherent difficulty, industrial-sized stochastic programs can now be solved