IN7K2: Optimización Bajo Incertidumbre March 22, 2012

Handout 1: Stochastic Programming I

Lecturer: Daniel Espinoza, Fernando Ordóñez

1 A Multiproduct Assembly Example

A manufacurer produces n products out of a total of m different parts which are ordered. One unit of product i requires a_{ij} units of part j. The manufacturer orders parts, then we observe how many units of each product are demanded $d_1, \ldots, d_n$, and should decide how much of each demand to satisfy.

$$\begin{array}{ll} \min_{z,y} & \sum_{i=1}^n (l_i - q_i) z_i - \sum_{j=1}^n s_j y_j \\ \text{s.t.} & y_j + \sum_{i=1}^n a_{ij} z_i = x_j \\ & 0 \leq z_i \leq d_i \\ & y_j \geq 0 \end{array}$$

where what is the strategic problem to be decided by the manufacturer ahead knowing the

demand? What if there are finitely many demand scenarios?

1.1 Chance constraints

If the manufacturer is more concerned about not loosing demand (and not so much about the salvage value of unused items) then a more interesting constraint would be to enforce that $\sum_{i=1}^{n} a_{ij} d_i \le x_j$ holds with high probability. The problem would be transformed as:

This corresponds to the chance constrained model:

$$\begin{array}{rl} \min_x & c^T x \\ \text{s.t.} & \Pr\{A^T D \le x\} \ge 1 - \alpha \\ & x \ge 0 \end{array}$$

easy when:

but typically hard:

$$\Pr\{\xi x_1 + x_2 \ge 7\} \ge 1 - \alpha, \quad \text{with } \xi \sim U[0,1]$$

1.2 Multi-stage model

Assume now that the manufacturer has a planning horizon of T periods, where each period a demand $d_{t1}, \ldots, d_{tn}$ is observed, assume that unused parts can be stored. Each period decisions of what to order x_t, what to produce z_t, and what parts to store y_t are made.

If we denote $D_{[t]} = (D_1, \ldots, D_t)$ the uncertain demand vector from period 1 to period t and $d_{[t]} = (d_1, \ldots, d_t)$ a sample realization of this uncertain vector, express this multistage problem

1.3 Network Design Problem

A deterministic mathematical model for the supply chain design problem can be written as follows:

$$\begin{array}{rl} \min_{x,y} & c^T x + q^T y \\ \text{s.t.} & Ny = 0 \\ & Cy \geq d \\ & Sy \leq s \\ & Ry \leq Mx \\ & x \in \mathcal{X}, y \geq 0 \end{array}$$

write this problem as a two stage stochastic programming problem if at the time of the desing of the supply chain there is uncertainty about operational costs, demand, supplies, processing requirements and processing capacities.

what about a multistage stochastic problem.

2 Two Stage Problems

We now want to explore the two stage stochastic linear problem

$$\begin{array}{rl} \min_x & c^T x + E\{Q(x,\xi)\} \\ \text{s.t.} & Ax = b, x \geq 0 \end{array}$$

where $Q(x,\xi)$ represents the value of the second stage given the uncertain parameters $\xi = (T, W, h, q)$ given by

$$\begin{array}{rl} \min_y & q^T y \\ \text{s.t.} & Tx + Wy = h, y \geq 0\ . \end{array}$$

Property The function $Q(\cdot,\xi)$ is convex and either

- polyhedral
- $Q(x,\xi) = +\infty$
- $Q(\cdot,\xi) = -\infty$

and if $Q(x^0,\xi)$ is finite, then $\delta Q(x^0,\xi) = -T\pi^*$ where $\pi^* = \text{argmax}(h - Tx^0)^T\pi\ :\ W^T\pi \leq q$

Example What is the subdifferential of the second stage cost function for the 2 stage supply chain design problem if the uncertainty has finite support?

3 Cutting Plane Methods: Bender's Decomposition

Example: Consider the problem in n variables, with $n << m$:

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & a_i^T x \leq b_i \quad i = 1 \dots m \end{array}$$

Consider an algorithm that defines a restricted problem with only $k < m$ constraints, and then gradually adds constraints that are missing:

$$\begin{array}{ll} \max & c^T x \\ \text{s.t.} & a_i^T x \leq b_i \quad i = 1 \dots k \ . \end{array}$$

- If x^* optimal for (RM) is feasible for $i = 1 \dots m$ then
- Otherwise need to identify a violated inequality

What upper and lower bounds does the cutting plane method provide?

3.1 Bender's basics

Problems for which Bender's Decomposition (constraint generation) methods work best, are those that have a large number of constraints and the following structure

$$\begin{array}{llllll} \min\limits_{x, y_1, \dots, y_k} & c^T x + & q^T y_1 & \dots & q^T y_k & \\ \text{s.t.} & Ax & & & & = b \\ & T_1 x & W y_1 & & & = h_1 \\ & T_2 x & & W y_2 & & = h_2 \\ & \vdots & & \ddots & & = \vdots \\ & T_k x & & & W y_k & = h_k \\ & \multicolumn{5}{l}{\quad x, \ y_1, \ \dots, \ y_k \geq 0} \end{array}$$

This structure is exploited by doing each minimization separately:

$$\begin{array}{rl} \min_x & c^T x + \sum_{i=1}^k Q_i(x) \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad \text{with} \qquad Q_i(x) = Q(x, \xi^i) = \begin{array}{rl} \min_{y_i} & q^T y_i \\ \text{s.t.} & W y_i = h_i - T_i x \\ & y_i \geq 0 \end{array}$$

for every $i = 1, \dots, k$. What does the function $Q_i(x)$ look like?

The functions $Q_i(x)$ are constructed sequentially through cuts:

- Optimality cut: If $\gamma \geq Q_i(x)$ then

- Feasibility cut: If x makes some subproblem infeasible, $Q_i(x) = \infty$, then

The master problem that is solved is

$$\begin{array}{rl} \min_x & c^T x + \sum_{i=1}^k \gamma_i \\ \text{s.t.} & Ax = b \\ & x \geq 0 \\ & h_i^T z^k - x^T T_i z^k \leq \gamma_i \ \text{ for } z^k \text{ BFS of } W^T z \leq q \\ & h_i^T w^k - x^T T_i w^k \leq 0 \ \text{ for } w^k \text{ extreme ray of } W^T z \leq q \end{array}$$

BFS: basic feasible solution = extreme point. An important observation is that if W and q are stochastic, i.e. vary in each scenario, then we have the same master with extreme points and rays z_i^k and w_i^k for each $Q_i(\cdot)$.

3.2 Benders Decomposition Algorithm

1- Formulate a master problem (i.e. find a BFS for $W^T z \leq q$)
2- Obtain (x^*, γ^*) the optimal solution for the master problem
3- Solve the subproblem $Q_i(x^*)$ for every scenario i
4- If all subproblems have optimal solution $Q_i(x^*) \leq \gamma_i^*$
- STOP: (x^*, γ^*) is optimal as it is feasible for all cuts
- else if some i has $\infty > Q_i(x^*) > \gamma_i^*$
- add optimality cut $h_i^T z^k - x^T T_i z^k \leq \gamma_i$
- else if some i has $Q_i(x^*) = \infty$
- add feasibility cut $h_i^T w^k - x^T T_i w^k \leq 0$
5- Goto 2

- How do we start this algorithm?

- What if we can't wait for it to converge?

3.3 Example

An electric utility company faces the problem of satisfying demand at minimum cost. In the case of a thermal plant and a hydro plant, satisfying the demand over the next two periods can be written as:

$$\begin{aligned} \min \quad & 3x_1 + 3x_2 \\ & x_1 + h_1 \geq 10 \\ & x_2 + h_2 \geq 12 \\ & h_1 \leq 5 \\ & h_2 \leq V_2 \\ & V_2 + h_1 = 5 + r \\ & x_i, h_i \geq 0 \end{aligned}$$

Note that this is a production and inventory problem.

Suppose

- 2nd period demand can be 15 or 10 each with prob. 1/2
- 2nd period thermal cost can be 1 or 5 each with prob. 1/2
- rain can be $r = 0$ or $r = 10$ each with prob. 1/2

Question: Best strategy to satisfy 1st period demand considering uncertainty?

scenario	prob.	2nd demand	thermal cost	rain	best strategy
1	0.125	15	5	0	save water, x_1 high
2	0.125	10	3	10	use water, x_1 low
$\vdots$	$\vdots$				

Solution: Minimize the expected value: $\sum_{i=1}^{8} p_i z_i$. The z_i is the optimal solution for each scenario. Note that some variables have to be decided before the uncertainty (x_1 and h_1) and some after the uncertainty (x_2 and h_2). *Recourse.*

This leads to the following problem:

4 Multi-Stage Problems

4.1 Non-anticipativity- 2 stage

Consider the following equivalent formulation of the 2 stage stochastic programming problem with finite number of scenarios:

$$\begin{array}{ll} \min & \sum_{k=1}^{K} p_k c^T x_k + Q(x_k, \xi_k) \\ \text{s.t.} & Ax_k = b, \;\; x_k \geq 0 \\ & x_k = x_{k+1} \;\;\; k = 1, \ldots, K-1 \end{array}$$

The second set of constraints define the non-anticipativity constraints, all first stage (here-and-now) variables have to be coordinated prior to the uncertainty being revealed. What if these constraints did not exist?

An equivalent form of these non-anticipativity constraints (a linear subspace in nK dimensional space) is

$$Px = (\sum_{k=1}^{K} p_i x_i, \ldots, \sum_{k=1}^{K} p_i x_i) = (x_1, \ldots, x_K) = x$$

4.2 Linear Multistage Problems

Consider the problem

$$\begin{array}{lcccccccccccl}
\min & c_1^T x_1 & + & c_2^T x_2 & + & c_3^T x_3 & + & \ldots & + & c_T^T x_T & & \\
\text{s.t.} & A_1 x_1 & & & & & & & & & = b_1 \\
 & B_2 x_1 & + & A_2 x_2 & & & & & & & = b_2 \\
 & & & B_3 x_2 & + & A_3 x_3 & & & & & = b_3 \\
 & & & & & & \ddots & & & & \\
 & & & & & & & B_T x_{T-1} & + & A_T x_T & = b_T \\
 & x_1 \geq 0 & & x_2 \geq 0 & & x_3 \geq 0 & & \ldots & & x_T \geq 0 &
\end{array}$$

Where the decision x_t is taken knowing the outcome of the uncertainty from periods $1, 2, \ldots, t$ and the decisions taken in periods $1, 2, \ldots, t-1$. The objective is to minimize the expected value. Letting $\xi_{[t]} = (\xi_1, \xi_2, \ldots, \xi_t)$ denote the vector of uncertain variables up to period t we can express the Nested Formulation of the multistage problem as

$$\min_{\substack{A_1 x_1 = b_1 \\ x_1 \geq 0}} c_1^T x_1 + \mathbb{E}\left[\min_{\substack{B_2 x_1 + A_2 x_2 = b_2 \\ x_2 \geq 0}} c_2^T x_2 + \mathbb{E}\left[\cdots + \mathbb{E}\left[\min_{\substack{B_T x_{T-1} + A_T x_T = b_T \\ x_T \geq 0}} c_T^T x_T \right]\right]\right]$$

Where the expectations are taken with respect to what is uncertain up to each period. A dynamic programming formulation is as follows: Starting from period T Therefore at period

t the future cost function is

4.3 Scenario tree and non-anticipativity constraints

Given a finite uncertainty scenarios we can obtain represent the multistage optimization problem by repeating all variables in every scenario by:

$$\begin{array}{llllllll}
\min & \sum_{k=1}^{K} p_k \Big[(c_1^k)^T x_1^k & + & (c_2^k)^T x_2^k & + & (c_3^k)^T x_3^k & + & \dots & + & (c_T^k)^T x_T^k \Big] & \\
\text{s.t.} & A_1 x_1^k & & & & & & & & & = b_1 \\
 & B_2^k x_1^k & + & A_2^k x_2^k & & & & & & & = b_2^k \\
 & & & B_3 x_2^k & + & A_3^k x_3^k & & & & & = b_3^k \\
 & & & & & & & \ddots & & & \\
 & & & & & & & B_T^k x_{T-1}^k & + & A_T^k x_T^k & = b_T^k \\
 & x_1^k \geq 0 & & x_2^k \geq 0 & & x_3^k \geq 0 & & \dots & & x_T^k \geq 0 &
\end{array}$$

Where the uncertainty is represented by a scenario tree

The non-anticipativity constraints, not only have to match up all the first stage variables $x_1^l = \sum_{k=1}^{K} p_k x_1^k$, but also all variables that share a common history up to a certain period:

4.4 Progressive Hedging

We say that the solution $(x_1^k, \dots, x_T^k)_{k=1}^K$ is

- admissible: for each k $(x_1^k, \dots, x_T^k) \in C_k$ is feasible for scenario k, i.e.:

$$\begin{aligned} &A_1 x_1^k = b_1 \\ &B_t^k x_{t-1}^k + A_t^k x_t^k = b_t,\ t = 2 \dots T \end{aligned}$$

- implementable: when the value of solution at time t only depends on past information, that is

$$x_t^k = \frac{1}{\sum_{s \in A_t} p_s} \sum_{s \in A_t} p_s x_t^s = \mathbb{E}[x_t^s | A_t], \quad \forall k \in A_t$$

  Note that this is equivalent to a linear constraint $(I - J)x = Kx = 0$

The progressive hedging algorithm is an augmented Lagrangian method to solve our problem of interest:

$$\begin{array}{ll} \min & \sum_{k=1}^{\bar{K}} p_k (\sum_{t=1}^T (c_t^k)^T x_t^k)) \\ \text{s.t.} & (x_1^k, \dots, x_T^k) \in C_k \\ & Kx = 0 \end{array}$$

1- At iteration i: x^i an admissible (not implementable) solution
   W^i vector of multipliers

2- Let $\hat{x}^i = Jx^i$

3- Solve the subproblem $Q_i(x^*)$ for every scenario i

4- Solve the separable subproblem to obtain x^{i+1}, that is
   $(x^{i+1})^k = \text{argmin} \sum_{t=1}^T (c_t^k)^T x_t + (W^i)^T x_t + r/2 \|x_t - \hat{x}_t^{ik}\|_2^2 \mid x_t \in C_k$

5- $W^{i+1} = W^i + rKX^{i+1}$

6- $i \leftarrow i + 1$. STOP if $\|x^i - \hat{x}^i\| \leq \varepsilon$, otherwise Goto 2

5 Sample Average Approximation

Given x_i i.i.d. random variables with mean μ and standard deviation σ we have that the central limit theorem says that

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ \sim \ N(\mu, \frac{1}{n}\sigma^2) \text{ for large n}$$

we use this property to build confidence intervals (smallest interval guaranteed to contain a value with a given probability):

$$\mu \in [\bar{x} - \frac{1}{\sqrt{n}}\sigma\phi^{-1}(\alpha/2), \bar{x} + \frac{1}{\sqrt{n}}\sigma\phi^{-1}(\alpha/2)]$$

with probability $1-\alpha$

If the uncertainty of a stochastic programming problem is continuous:

$$\begin{aligned} v^* = \quad & \min \quad c^T x + \mathbb{E}\{Q(x,\xi)\} \\ & \text{s.t.} \quad Ax = b,\ x \geq 0 \end{aligned}$$

we can use statistics to obtain bounds on the optimal solution value by sampling the uncertainty and considering the approximate problem

$$\begin{aligned} v_N^* = \quad & \min \quad c^T x + \tfrac{1}{N}\textstyle\sum_{k=1}^{N} Q(x,\xi^k) \\ & \text{s.t.} \quad Ax = b,\ x \geq 0 \end{aligned}$$

Indeed we have that

$$\hat{g}_{N'}(\bar{x}) = c^T\bar{x} + \frac{1}{N'}\sum_{k=1}^{N'} Q(\bar{x},\xi^k)$$

is a statistical upper bound if $\bar{x}$ is feasible. That is with $1-\alpha$ probability

$$c^T\bar{x} + \mathbb{E}\{Q(\bar{x},\xi)\} \in [\hat{g}_{N'}(\bar{x}) - STD(\hat{g}_{N'}(\bar{x}))\phi^{-1}(\alpha/2), \hat{g}_{N'}(\bar{x}) + STD(\hat{g}_{N'}(\bar{x}))\phi^{-1}(\alpha/2)]$$

where $STD(\hat{g}_{N'}(\bar{x})) = \sqrt{\frac{1}{N'(N'-1)} \sum_{i=1}^{N'} (c^T \bar{x} + Q(\bar{x}, \xi^i) - \hat{g}_{N'}(\bar{x}))^2}$.

Similarly $\bar{v}_{N,M} = \frac{1}{M} \sum_{m=1}^{M} v_N^m$ is a statistical lower bound of v^* whose empirical variance is

$$\sigma^2_{\bar{v}_{N,M}} = \frac{1}{M(M-1)} \sum_{m=1}^{M} (v_N^m - \bar{v}_{N,M})^2$$