Carnegie Mellon University

ACCELERATING BENDERS STOCHASTIC DECOMPOSITION FOR THE OPTIMIZATION UNDER UNCERTAINTY OF THE PETROLEUM PRODUCT SUPPLY CHAIN

Fabricio Oliveira, Ph.D.c.
Industrial Engineering Department
Pontifícia Universidade Católica do Rio de Janeiro

(Visiting Researcher – Grossmann Research Group)
Advisors: Prof. Silvio Hamacher (PUC-Rio)
Prof. Ignacio Grossmann (CMU)

CAPD CENTER

Carnegie Mellon University

Introduction

CAPD CENTER

- For almost 50 years, companies in the oil and chemical industries have pushed the development and use of linear mixed-integer programming to support decision making at all levels of planning. However:
  - The predominant focus on the tactical and operational planning of refining activities
  - Logistics of petroleum products being represented in a simplified fashion
  - Most of the applications in the field have not considered uncertainty in their evaluations
- Aiming to fill this gap, we consider the strategic planning of petroleum products (downstream) distribution under uncertainty, deciding over:
  - the levels of investments in **logistics** infrastructure
  - the distribution of flows
  - inventory policies
  - external commercialization
- The source of uncertainty considered is the future demand levels for the petroleum products

---

Carnegie Mellon University

Introduction

CAPD CENTER

- *Our objective is:*
  - Propose a computational tool capable of providing support to decision making in the design process of the petroleum downstream distribution under uncertainty.
    - The model is based on two-stage stochastic programming for the optimization of oil product distribution that considers issues of tactical planning (second-stage) to evaluate decisions of a strategic nature (first-stage);
- *Main challenges:*
  - Complexity of the mathematical model;
  - Dealing with potentially huge number of scenarios;
  - Solving the problem efficiently in terms of available computational capacity and solution time;

Carnegie Mellon University

Mathematical Model

CAPD CENTER

First-Stage:

Node Investment Costs

Arc Investment Costs

$$\min_{\mathbf{w},\mathbf{y}\in\{0,1\}} \boxed{\sum_{j,p,t} W_j^t w_{jp}^t} + \boxed{\sum_{i,j,t} Y_{ij}^t y_{ij}^t} + \boxed{\mathcal{Q}(\mathbf{w},\mathbf{y})}$$

Second-stage cost

$$s.t. \quad \sum_t w_{jp}^t \leq 1 \quad \forall j,p$$

$$\sum_t y_{ij}^t \leq 1 \quad \forall i,j$$

Investment Constraint
(only once in the time horizon)

Mathematical Model

Second-Stage:

$$\min_{\mathbf{x},\mathbf{v},\mathbf{u}\in\mathbb{R}_+} \underbrace{\sum_{i,j,p,t} C_{ij}x_{ijp}^t(\xi)}_{\text{Freight Costs}} + \underbrace{\sum_{j,p,t} H_{jp}v_{jp}^t(\xi)}_{\text{Holding Costs}} + \underbrace{\sum_{j,p,t} S_{jp}u_{jp}^t(\xi)}_{\text{Shortfall Costs}}$$

s.t

$$\underbrace{\sum_i x_{ijp}^t(\xi) + v_{jp}^{t-1}(\xi) + u_{jp}^t(\xi) = \sum_l x_{jlp}^t(\xi) + v_{jp}^t(\xi) + D_{jp}^t(\xi)}_{\text{Material Balance}} \qquad \forall j \in \mathcal{B}, p, t$$

$$\underbrace{\sum_l x_{jlp}^t(\xi) \le O_{jp}^t}_{\text{Supply Limit}} \qquad \forall j \in \mathcal{S}, p, t$$

$$\underbrace{\sum_p x_{ijp}^t(\xi) \le A_{ij}^0 + A_{ij}\sum_{t'\le t} y_{ij}^{t'}}_{\text{Arc Capacity}} \qquad \forall i, j, t$$

$$\underbrace{v_{jp}^t(\xi) \le M_{jp}^0 + M_{jp}\sum_{t'\le t} w_{jp}^{t'}}_{\text{Node Capacity}} \qquad \forall j \in \mathcal{B}, p, t$$

$$\underbrace{\sum_{i,p} x_{ijp}^t(\xi) \le K_{jp}\left(M_{jp}^0 + M_{jp}\sum_{t'\le t} w_{jp}^{t'}\right)}_{\text{Throughput Limit}} \qquad \forall j \in \mathcal{B}, p, t$$

Cutting-Plane Approach

- The use of cutting-plane approaches, such as **Benders decomposition**, might present itself as an efficient framework for dealing with large mixed-integer optimization problems
- However, it is commonly reported in literature that the direct approach of such cutting-plane techniques might not be as efficient as one would expect (see, for example Rei et al., 2007, Saharidis et al., 2010)
  - Convergence problems due to the complexity of the subproblems (both Master and/or Slave problems);
  - Difficulties in solving large integer/mixed-integer problems;

Cutting-Plane Approach

- Some acceleration ideas has been proposed for deal with these drawbacks:
  - Relaxing the Master Problem (McDaniel and Devine, 1977)
  - Getting feasible solutions instead of the optimal MP solution (Cote and Laughton, 1984)
  - Pareto-optimal cuts (Magnanti and Wong , 1981; Papadakos, 2008)
  - Non-dominated cuts (Sherali and Lunday, 2011)
  - Maximum Feasible Subsystem cuts (Saharidis and Ierapetritou, 2010)
  - Combined feasibility and optimality cuts (Fischetti et al., 2010)
  - Local Branching cuts (Rei et al.2007)
  - …

Suits our case: mixed-integer with complete recourse

Stochastic Benders Decomposition

- Although there is several research investigations towards improving the efficient of Benders procedure, the number of papers that uses **these ideas within stochastic problems** are **very reduced**;
- Therefore, we seek to **combine** the recently developed **acceleration techniques** for the traditional **Benders decomposition** algorithm with its extension to the **stochastic version**;
- We propose a **decomposition framework** based on the Stochastic Benders decomposition (commonly referred as the L-Shaped Method (Van Slyke and Wets (1969))
  - Exploit the multi-scenario structure;
  - Traditionally used within the context of Supply Chain Design under uncertainty;

Carnegie Mellon University

CAPD CENTER

Stochastic Benders Decomposition

- We can state the Master Problem (MP) as follows:

$$\min_{\mathbf{w},\mathbf{y}\in\{0,1\}} \sum_{j,p,t} W_{jp}^t w_{jp}^t + \sum_{i,j,t} Y_{ij}^t y_{ij}^t + \boxed{M}$$

$$s.t.: \quad \sum_t w_{jp}^t \le 1 \qquad \forall j,p$$

$$\sum_t y_{ij}^t \le 1 \qquad \forall i,j$$

$$\boxed{M \ge \mathcal{Q}(\mathbf{w},\mathbf{y})}$$

Bi-level complete problem

- We need to reformulate this, in order to make it "solvable"…
  - We use dual theory for representing as its vertices…

Carnegie Mellon University

CAPD CENTER

Stochastic Benders Decomposition

- We can state the Master Problem (MP) as follows:

$$\min_{\mathbf{w},\mathbf{y}\in\{0,1\}} \sum_{j,t} W_{jp}^t w_{jp}^t + \sum_{i,j,t} Y_{ij}^t y_{ij}^t + M$$

$$s.t.: \quad \sum_t w_{jp}^t \le 1 \qquad \forall j,p$$

$$\sum_t y_{ij}^t \le 1 \qquad \forall i,j$$

$$M \ge \sum_{\xi\in\Omega} P(\xi)\left\{\sum_{j,p,t} D_{jp}^t(\xi)\alpha_{jp}^{t(b)}(\xi) + \sum_{j,p,t} O_{jp}^t\beta_{jp}^{t(b)}(\xi) + \sum_{i,j,t}\left(A_{ij}^0 + A_{ij}\sum_{t'\le t} y_{ij}^{t'}\right)\gamma_{ij}^{t(b)}(\xi)\right.$$

$$\left. + \sum_{j,p,t}\left(M_{jp}^0 + M_{jp}\sum_{t'\le t} w_{jp}^{t'}\right)\delta_{jp}^{t(b)}(\xi) + \sum_{j,p,t}\left[K_{jp}\left(M_{jp}^0 + M_{jp}\sum_{t'\le t} w_{jp}^{t'}\right)\right]\zeta_{jp}^{t(b)}(\xi)\right\} \quad \boxed{\forall b \in B}$$

set of all extreme points of the polyhedron of the Dual Slave Problem

- This can be potentially huge and, thus, impossible to solve…
  - However, since only a subset of would be binding at optimality, one could use an iterative process to compose this subset of B.

Carnegie Mellon University

CAPD CENTER

Stochastic Benders Decomposition

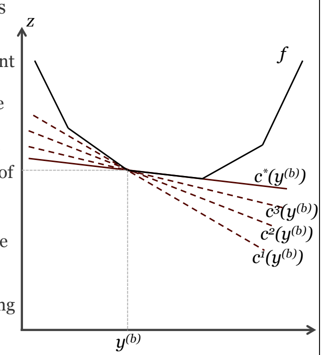
- A overview of the mentioned algorithm can be stated as follows:
  - *Step 1*: solve the MP to obtain a feasible solution (w,y). The value of this solution represents an lower bound for the value of the complete problem;
  - *Step 2*: plug the feasible solution into the Slave Problem (the second-stage problem, given a fixed first-stage solution) to obtain $\alpha^{(b)}$, $\beta^{(b)}$, $\gamma^{(b)}$, $\delta^{(b)}$, and $\zeta^{(b)}$. The value of this solution represents an upper bound to the complete problem
  - *Step 3*: If the lower and the upper bounds are close enough, then we end the procedure. Otherwise, we generate a cut representing this dual vertex, insert it into the MP, and go back to *Step 1*.
- The structure of stochastic programs allows us to add multiple cuts to the MP instead of one in each major iteration;
  - We consider the problem under **two decomposable dimensions**: stages AND scenarios;
- Birge and Louveaux (1988) show that the usage of such a framework may greatly speed up convergence;

Carnegie Mellon University

CAPD CENTER

Acceleration Techniques

- Generating stronger cuts
  - In case of degeneracy of the sub-problem, different cuts might be generated from the same first-stage optimal point
  - These cuts might loosely represent the true slope of the objective function segment
  - The main idea behind the acceleration technique proposed is is to choose the "strongest one" among these cuts.



Carnegie Mellon University

Acceleration Techniques

1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. **Stochastic Benders Decomposition**
   a. **Acceleration Technique**
   b. Our Proposition
5. Case Study
6. Numerical Experiments
7. Conclusions

- Based on interpreting Magnanti and Wong's (1981) pioneer ideas for generating non-dominated cuts, Sherali and Lunday (2011) proposed an strategy that uses small perturbations of the RHS of the Benders subproblem to generate **maximal non-dominated** Benders cuts.
  - Their ideas are based on the definition of *maximal cuts,* commonly used in the cutting-plane methods literature.

Carnegie Mellon University

Acceleration Techniques

1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. **Stochastic Benders Decomposition**
   a. **Acceleration Technique**
   b. Our Proposition
5. Case Study
6. Numerical Experiments
7. Conclusions

- Sherali and Lunday Maximal cuts

Positive Weight Vector (PWV)

$$\max_{\alpha,\beta,\gamma,\delta,\zeta} \sum_{\xi\in\Omega} P(\xi)\left\{\sum_{j,p,t} D_{jp}^{t}(\xi)\alpha_{jp}^{t}(\xi) + \sum_{j,p,t} O_{jp}^{t}\beta_{jp}^{t}(\xi) + \sum_{i,j,t}\left(A_{ij}^{0} + A_{ij}\sum_{t'\le t}\bar{y}_{ij}^{t'}\right)\gamma_{ij}^{t}(\xi) + \sum_{j,p,t}\left(M_{jp}^{0} + M_{jp}\sum_{t'\le t}\bar{w}_{jp}^{t'}\right)\delta_{jp}^{t}(\xi) + \sum_{j,p,t}\left[K_{jp}\left(M_{jp}^{0} + M_{jp}\sum_{t'\le t}\bar{w}_{jp}^{t'}\right)\right]\zeta_{jp}^{t}(\xi)\right\}$$

$$\boxed{+\mu}\left(\sum_{\xi\in\Omega} P(\xi)\left\{\sum_{j,p,t} D_{jp}^{t}(\xi)\alpha_{jp}^{t}(\xi) + \sum_{j,p,t} O_{jp}^{t}\beta_{jp}^{t}(\xi) + \sum_{i,j,t}\left(A_{ij}^{0} + A_{ij}\sum_{t'\le t}\boxed{\bar{\bar{y}}_{ij}^{t'}}\right)\gamma_{ij}^{t}(\xi) + \sum_{j,p,t}\left(M_{jp}^{0} + M_{jp}\sum_{t'\le t}\boxed{\bar{\bar{w}}_{jp}^{t'}}\right)\delta_{jp}^{t}(\xi) + \sum_{j,p,t}\left[K_{jp}\left(M_{jp}^{0} + M_{jp}\sum_{t'\le t}\boxed{\bar{\bar{w}}_{jp}^{t'}}\right)\right]\zeta_{jp}^{t}(\xi)\right\}\right)$$

$$\begin{aligned} s.t.:\ & \alpha_{jp}^{t}(\xi) - \alpha_{ip}^{t}(\xi) + \beta_{ip}^{t}(\xi) + \gamma_{ij}^{t}(\xi) + \zeta_{jp}^{t}(\xi) \le C_{ij} && \forall i,j,p,t,\xi \\ & \alpha_{jp}^{t}(\xi) - \alpha_{jp}^{t-1}(\xi) + \delta_{jp}^{t}(\xi) \le H_{jp} && \forall j\in\mathcal{B},p,t,\xi \\ & \alpha_{jp}^{t}(\xi) \le S_{jp} && \forall j\in\mathcal{B},p,t,\xi \\ & \boldsymbol{\alpha}\in\mathbb{R}, \boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta},\boldsymbol{\zeta} \le \mathbf{0} \end{aligned}$$

Perturbation magnitude

Carnegie Mellon University

Our Proposition

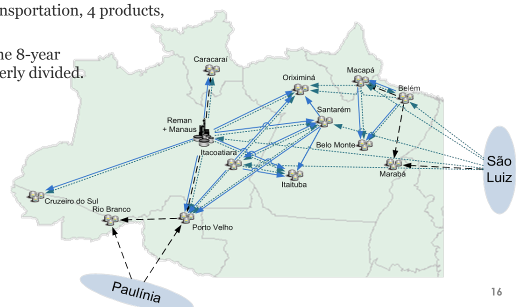
1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. **Stochastic Benders Decomposition**
   a. Acceleration Technique
   b. **Our Proposition**
5. Case Study
6. Numerical Experiments
7. Conclusions

- Near-optimal maximal non-dominated cuts with updated weights
  - The main difficulty related with Sherali and Lunday's method:
    - strongly relies on the definition of two parameters: the PWV and the perturbation magnitude term.
    - The authors proposed reliable values for both terms under the assumption of relatively well-scaled problems,
      - In fact, we observed that for some of our cases, the use of certain values for μ might even prevent the algorithm to reach the optimal solution.
  - What we propose:
    - Dynamically update policy for the weight parameter;
    - Strict proof of convergence in such case;
  - What we observed from our numerical experiments:
    - Consistently **provided better results** in terms of computational times;
    - It **did not seem to critically depend** on the exact fixed parameter value to perform better than the others acceleration approaches considered;
    - Seems to **describe more efficiently the second-stage cost curve** despite the scenario sampled, as it presents the smaller variance in the solution times observed;

Case Study

Carnegie Mellon University

- Refined products distribution in northern Brazil
  - Characteristics:
    - Transportation through waterways, subject to sazonalities
    - 3 means of transportation, 4 products, 13 locations;
    - Planning for the 8-year horizon, quarterly divided.



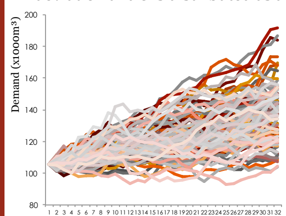
Case Study - Scenarios

1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. Stochastic Benders Decomposition
   a. Acceleration Technique
   b. Our Proposition
5. **Case Study**
6. Numerical Experiments
7. Conclusions

- Scenario Generation
  - The generation of scenarios is made according to a first order autoregressive model, defined as:

$$D_{lp}^{t} = D_{lp}^{t-1}\left[1 + \omega_t + \sigma\varepsilon_{lp}\right] \quad \forall l, p$$

where ω is the forecasted growth rate, σ is a estimated maximum deviation and ε is distributed as a standard normal.



Example of 50 demand scenarios for Diesel demand in Manaus

$$N = \left(\frac{z_{\alpha/2}\hat{\sigma}_N}{\beta\hat{z}(w,y)}\right)^2$$

		α			
		90%	95%	99%	100%
β	10%	61	87	150	244
	5%	244	347	599	978
	3%	977	1388	2397	3911
	2%	1527	2168	3745	6111
	1%	6108	8672	14978	24443

Minimum number of scenarios considering the significance level α and the acceptable margin β

Numerical Experiments

1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. Stochastic Benders Decomposition
   a. Acceleration Technique
   b. Our Proposition
5. Case Study
6. **Numerical Experiments**
7. Conclusions

- Size of the instances: 13 bases and 1 refinery and 2 external supply locations, 24 periods (quarterly);
- We compared 4 different solution procedure:
  - *Independent Pareto Optimal cuts* (Magnanti and Wong, 1986; Papadakos, 2008) - IPOc
  - *Maximal Nondominated cuts* (Sherali and Lunday, 2011); - MNc
  - *Near-optimal maximal non-dominated cuts with updated weights* (our proposition) - NMUWc
  - Solving the Deterministic Equivalent - CPLEX 11.2 - CPLEX

Numerical Experiments

- Results for 100 samples – computational time(s)

	NMUWCc		IPOc		MNc		CPLEX	
Scenarios	Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation	Average	Std. Deviation
20	22.3	2.2	56.5	19.2	55.5	15.0	9.6	0.6
40	43.2	2.7	104.7	37.7	102.1	31.3	33.7	3.3
60	64.0	4.5	162.4	66.5	140.7	52.0	51.5	5.0
80	83.7	5.8	219.5	78.6	221.5	111.8	104.0	20.9
100	109.1	12.7	287.2	124.3	272.5	139.2	185.6	34.5
120	128.3	9.2	369.6	150.5	313.9	152.4	287.8	46.3
140	151.2	10.8	387.6	163.3	381.8	170.9	473.0	122.3
160	176.7	11.6	483.8	208.8	372.3	135.8	601.1	140.8
180	203.8	13.9	612.7	270.3	477.1	189.4	734.9	165.6
200	230.1	12.2	631.4	252.7	521.8	196.1	972.6	161.0

Computational Time consistently small for instances with larger number of scenarios

19

Conclusions

1. Introduction
2. Mathematical Model
3. Cutting-Plane approach
4. Stochastic Benders Decomposition
   a. Acceleration Technique
   b. Our Proposition
5. Case Study
6. Numerical Experiments
7. **Conclusions**

- Main Conclusions:
  - As our computations suggest, our modified algorithm performed better for the problem considered under a sampling framework.
  - The experimental results showed that, for a greater number of scenarios, the proposed modification can achieve performances 4.5 times faster than the use of a traditional CPLEX solver.
  - Our algorithm also presented better results when compared to the other acceleration approaches recently appeared in the literature considered in this chapter.
- Main contributions:
  - The development of a mathematical model which contemplates the logistics of the downstream distribution in more comprehensive fashion;
  - The consideration of stochasticity in the demand and the proposition of a framework for modeling the uncertainty and obtaining statistically certified solutions, without the use of sophisticated scenario generation procedures;
  - The application of recently developed improvements for cutting plane solution procedures under the context of stochastic optimization;