VIETNAM NATIONAL UNIVERSITY, HO CHI MINH CITY
UNIVERSITY OF TECHNOLOGY
FACULTY OF COMPUTER SCIENCE AND ENGINEERING

MATHEMATICAL MODELING (CO2011)

Assignment (Semester: 231, Duration: 06 weeks)

"Stochastic Programming and Applications"

(Version 0.1, in Preparation)

Advisor: Nguyễn Tiến Thịnh, CSE-HCMUT

Students: Trần Đình Đăng Khoa - 2211649.
Trần Đặng Hiển Long - 2252449.
Nguyễn Hồ Phi Ưng - 2252897.
Nguyễn Hồ Đức An - 2252009.
Vũ Minh Quân - 2212828.

HO CHI MINH CITY, OCTOBER 2023

Contents

1 Abstract

When you have a problem that requires you to find the optimal solution to a goal, while taking into account the limitations of your resources and the trade-offs of your choices, you may have a **linear programming problem**. This type of problem can be expressed using linear functions of some variables for both the goal and the limitations. Linear programming problems are very useful for modeling many practical situations in different fields, such as:

- A farmer who wants to maximize the profit from planting crops, while considering the available land, water, seeds, and fertilizer.
- A manufacturer who wants to minimize the cost of producing goods, while meeting the demand and quality standards of the customers.
- A transportation company who wants to optimize the routes and schedules of its vehicles, while reducing the fuel consumption and travel time.

A **stochastic programming problem** is a linear programming problem in which some of the parameters or variables are **uncertain**. The uncertainty can be expressed using probability distributions. The goal of a stochastic programming problem is to find the optimal solution that maximizes the expected value of the goal function, while satisfying the constraints with a certain probability.

In this report, we will introduce the basic concepts of ...

2 Introduction to Stochastic Programming and Optimization

2.1 What is Stochastic Programming?

An optimization problem is said to be a **stochastic program** if it satisfies the following properties:

1. There is a unique objective function.
2. Whenever a decision variable appears in either the objective function or one of the constraint functions, it must appear only as a power term with an exponent of **1**, possibly multiplied by a constant.
3. No term in the objective function or in any of the constraints can contain products of the decision variables.
4. The coefficients of the decision variables in the objective function and each constraint are ***probabilistic*** in nature.
5. The decision variables are permitted to assume fractional as well as integer values.

These properties ensure, among other things, that the effect of any decision variable is proportional to its value.

2.2 Approach to solve a Stochastic Programming problem

As mentioned above, stochastic problems involve some parameter or variables which are uncertain and described by probability distributions. To yield the optimal result of a stochastic problem, it requires a combination of mathematical modeling, statiscal analysis and optimization techniques to effectively handle the uncertainty that lie within the problem.

There are different methods to solve this kind of problem that depend on the complexity, the common methods are **Stochastic Linear Programming (SLP)**, **Stochastic Integer Programming (SIP)**, **Stochastic Dynamic Programming (SDP)**. This report will discuss specifically about **SLP** problems.

3 Background - Basic concepts

3.1 The Simplex Method

The **Simplex Method**, developed by George Dantzig, incorporates both *optimality* and *feasibility* tests to find the optimal solution(s) to a linear program (if one exists).
Geometrically, the Simplex Method proceeds from an initial extreme point to an adjacent extreme point until no adjacent extreme point is more optimal.

To implement the Simplex Method we first separate the *decision* and *slack* variables into two non-overlapping sets that we call the **independent** and **dependent** sets. For the particular linear programs we consider, the original independent set will consist of the decision variables, and the slack variables will belong to the dependent set.

The Algorithm:

1. Tableau Format: Place the linear program in Tableau Format, as explained later.
2. Initial Extreme Point: The Simplex Method begins with a known extreme point, usually the origin (0, 0).
3. Optimality Test: Determine whether an adjacent intersection point improves the value of the objective function. If not, the current extreme point is optimal. If an improvement is possible, the optimality test determines which variable currently in the independent set (having value zero) should *enter* the dependent set and become nonzero.
4. Feasibility Test: To find a new intersection point, one of the variables in the dependent set must *exit* to allow the entering variable from Step 3 to become dependent. The feasibility test determines which current dependent variable to choose for exiting, ensuring feasibility.
5. Pivot: Form a new, equivalent system of equations by eliminating the new dependent variable from the equations do not contain the variable that exited in Step 4. Then set the new independent variables to zero in the new system to find the values of the new dependent variables, thereby determining an intersection point.
6. Repeat Steps 3 - 5 until the extreme point is optimal.

3.2 Place holder

4 PROBLEM I. [Industry - Manufacutring]

4.1 Summary of the Problem

This problem revolves around minimizing the production output of an industrial firm F. Specifically, the firm is tasked with manufacturing n distinct products. Each product necessitates varying quantities of components, sourced from m external suppliers, each carrying a unit cost of b. Before the product demands are determined, the firm must make decisions regarding the quantity of components to procure. Following this procurement phase, the production commences and the product demands are disclosed. Corresponding to each product, there exists a selling price, denoted as q. In the event of an excess in component procurement beyond the production requirements, salvage components must be resold at a price of s to mitigate costs. Additionally, any unmet product demands incur an extra cost of l per unit to fulfill.

This problem constitutes a 2-stage stochastic linear programming (2-SLP) challenge, where the product demands introduce uncertainty. The primary aim across the two stages involves minimizing the expenses linked to both pre-ordering components and actual production. However, as stipulated in the assignment, this problem necessitates recourse action, transforming it into a 2-stage stochastic linear programming problem with recourse (2-SLPWR) with the request of finding the optimal value of $x, y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$. The comprehensive formulation process and subsequent solution will be detailed further below.

4.2 Model formulation

4.2.1 Notations

For ease of reading, the tables below summarize the notation of parameters and decision variables used in the formulation of this report.

Symbol	Definition
n	Number of products
m	Number of parts
i, j	Index of products and parts, respectively
b	The set of preorder cost of parts
q	Selling price per unit of product i
l	The set of additional cost to satisfy a demand per unit
A	The matrix representing parts needed for each product
a_{ij}	Number of part j needed for product i
s	The set of salvage part selling price
D	The set of the demand of the product
S	Number of scenarios
p_s	Probability of scenario s

Symbol	Definition
x	The set of preordering quantities
z	The set of manufactured products
y	The set of salvage parts

4.2.2 First Stage

The initial stage primarily focuses on ensuring the preordered quantity of parts is sufficient to accommodate future, yet uncertain, demands while avoiding surplus or deficit, which could adversely impact the firm's profitability. The decision variable, denoted as x, constitutes a *here-and-now* determination that must be made prior to the revelation of actual demand. The preordering cost is represented as a linear function:

$$f(x) = b^T x$$

However, as the problem at hand pertains to a 2-SLPWR (2-stage stochastic linear programming problem with recourse), probability functions linked to distinct scenarios are incorporated to fine-tune the original constraints. Consequently, the quantities denoted by x can be determined by solving the following optimization problem:

$$\text{minimize } g(x, y, z) = b^T x + Q(x) = b^T x + E[Z(z)]$$

Here, $Q(x) = E_\omega[Z] = \sum_{i=1}^{n} p_i c_i z_i$ is taken with respect to the probability distribution of $\omega = D$. $Q(x)$ represents the expected value of the optimal production cost (the optimal solution to the second-stage problem detailed below). Therefore, $Q(x) = \sum_{k=1}^{S} p_k(c^T z_k - s^T y_k)$, where z_k and y_k denote the set of manufactured products and the quantity of salvage parts in scenario k, respectively.

4.2.3 Second Stage

Subsequent to the initial stage, the actual demand $d = (d_1, d_2, ..., d_n)$ for the products is realized. Consequently, the optimal production plan is derived by solving the stochastic linear program involving decision variables x and y:

$$\begin{aligned}
\text{minimize } Z &= \sum_{i=1}^{n} (l_i - q_i) z_i - \sum_{j=1}^{m} s_j y_j \\
\text{subject to } y_j &= x_j - \sum_{i=1}^{n} a_{ij} z_i, \quad j = 1, ..., m \\
0 &\le z_i \le d_i, \quad i = 1, ..., n \\
y_j &\ge 0, \quad j = 1, ..., m
\end{aligned}$$

4.2.4 Grand model

It can be observed that to solve the first stage problem, the optimal solution of the second stage is required. However, the optimal solutoin of the second stage again, needs the value of x of the first stage. Here we counter the interdependence between the two stages. Therefore, to solve this problem, it is essential to combine the two stages into one *grand objective function* and let the unknown demand follows the binomial distribution specificied in the description of the assignment. Then, the quantities x can be decided and use that to solve the second stage again once the actual demand is realized. The grand objective function is described below:

$$\min\ g(x, y, z) = b^T x + \sum_{k=1}^{S} p_k (c^T z_k - s^T y_k)$$
$$\text{s.t. } y_k = x - A^T z_k, \quad k = 1, ..., S$$
$$x \geq 0$$
$$0 \leq z_k \leq d_k, \quad k = 1, ..., S$$
$$y_k \geq 0, \quad k = 1, ..., S$$

4.2.5 Solution

In the description of the assignment, specific value of the parameters are given as follows:

- n = 8
- m = 5
- S = 2
- p_s = 1/2
- Values of b, l, q, s and A are randomized
- Random demand vector D and its density p_i follows the binomial distribution $\mathbf{Bin(10,\frac{1}{2})}$

To solve this 2-SLPWR problem, our group decided to employ Gurobi - a powerful mathematical optimization solver in Python programming language with Gurobi API. The source code has been included in our assignment submission.

For the values of b, l, q, s and A, a function to randomize integer is used to generate the set of said values. To be specific, the ranges for each one is listed below:

- Preordering cost b: 50 -> 100
- Additional producing cost l: 100 -> 200
- Selling price per unit q: 1000 -> 2000
- Salvage part selling price s: 30 -> price of the corresponding part (for part j ranges from 1 to m, $s_j < b_j$ as specificied in the assignment)
- Number of parts required for each product A: the number of each part required for a product will ranges from 0 -> 5.

Then, a model for the first stage are defined identically to the grand model in section 4.2.4 above. The model is then optimized with demand vector D randomized following the binomial distribution to find quantities x. After that, a second model for the second stage is defined according to section 4.2.3. With that, a vector of d is hard-coded to represent the fact that the actual demand has been revealed, y and z then can be found by solving the optimizing problem.

For example, an instance of our model having the following parameters:

- b = [74 50 56 66 51]
- l = [138 131 196 188 169 127 164 101]

- $q = [1399\ 1122\ 1639\ 1410\ 1420\ 1572\ 1261\ 1377]$
- $s = [32\ 32\ 41\ 42\ 32]$
- $A = \begin{bmatrix} 1 & 4 & 0 & 2 & 4 \\ 3 & 1 & 3 & 2 & 0 \\ 3 & 0 & 4 & 2 & 4 \\ 3 & 1 & 2 & 1 & 4 \\ 3 & 2 & 3 & 1 & 1 \\ 2 & 0 & 3 & 0 & 2 \\ 0 & 4 & 0 & 3 & 1 \\ 3 & 0 & 2 & 2 & 1 \end{bmatrix}$
- $D = \begin{bmatrix} 5 & 4 & 5 & 5 & 4 & 6 & 4 & 6 \\ 6 & 2 & 6 & 6 & 4 & 4 & 6 & 5 \end{bmatrix}$

With those randomzed value, the optimal vector x is found: [89 64 84 66 95]. Then second stage is solved, getting the result of z and y as follows: [5 3 5 8 4 4 6 5], [1 1 5 0 0].

5 Member list & Workload

No.	Fullname	Student ID	Contribution	Percentage of work
1	Trần Đình Đăng Khoa	2211649	- Report, problem 1	20%
2	Trần Đặng Hiển Long	2252449	- Problem 2	20%
3	Nguyễn Hồ Phi Ưng	2252897	- Problem 1	20%
4	Nguyễn Hồ Đức An	2252009	- Report, problem 1	20%
5	Vũ Minh Quân	2212828	- Problem 2	20%

References

[1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 3rd edition, 2009.

[2] F.R. Giordano, W.P. Fox, and S.B. Horton. *A First Course in Mathematical Modeling*. Cengage Learning, 2013.

[3] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory, Second Edition*. Society for Industrial and Applied Mathematics, USA, 2014.

[4] Stein W. Wallace and William T. Ziemba. *Applications of Stochastic Programming (Mps-Siam Series on Optimization) (Mps-Saimseries on Optimization)*. Society for Industrial and Applied Mathematics, USA, 2005.

[5] Li Wang. *A two-stage stochastic programming framework for evacuation planning in disaster responses*, volume 145. 2020.