MSDS 460 Decision Analytics
Linear Optimization
Implicit/Explicit Modeling Formulation
#1

Brandon O'Briant
Northwester University SPS
M.S. Data Science

January 25, 2018

Contents

Hospital Menu Planning System

The technical staff of a hospital wishes to develop a computerized menu-planning system. To start with, a lunch menu is sought. The menu is divided into three major categories: vegetables, meat, and dessert. At least one equivalent serving of each category is desired. The cost per serving of some suggested items as well as their content of carbohydrates, vitamins, protein, and fats are summarized below.

Categories	Items	Carbohydrates	Vitamins	Protein	Fats	Cost in \$/serving
Vegetables	Peas	1	3	1	0	0.10
	Green beans	1	5	2	0	0.12
	Okra	1	5	1	0	0.13
	Corn	2	6	1	2	0.09
	Macaroni	4	2	1	1	0.10
	Rice	5	1	1	1	0.07
Meat	Chicken	2	1	3	1	0.70
	Beef	3	8	5	2	1.20
	Fish	3	6	6	1	0.63
Dessert	Orange	1	3	1	0	0.28
	Apple	1	2	0	0	0.42
	Pudding	1	0	0	0	0.15
	Jello	1	0	0	0	0.12

Suppose that the minimal requirement of carbohydrates, vitamins, protein, and fats per meal are respectively 5, 10, 10, and 2. Formulate the menu-planning problem (minimizing cost) as a linear program: clearly define the variables, and state the objective function and constraints with proper justification. Be sure that your formulation is written in explicit form.

Solution:

Problem Definition

To be able to properly maintain a state of homeostasis one must meat a minimum requirements of nutrients. This especially true when a patient at a hospital due to the nature of the body trying to recover or survive a trauma. With the current meal plan system it is difficult to make sure every patient is getting the correct balance of nutrients while minimizing the cost associated. Thus, a digital meal plan is being develop to help make sure patients bodies will have the minimum necessary nutrients while minimizing the total monetary strain on the hospital. An attempt at providing a start to this problem a prototype lunch meal plan is developed via LP modeling and is as follows:

Variable Definitions

Due to at least one equivalent serving from each of the major categories is desired, the items (foods) in each of the categories will become the decision variables. Do make this easier to work with than writing out each items full name we will define the decision variables as such:

x_1 := Peas

x_2 := Green Beans

x_3 := Okra

$$x_4 := \text{Corn}$$

$$x_5 := \text{Macaroni}$$

$$x_6 := \text{Rice}$$

$$x_7 := \text{Chicken}$$

$$x_8 := \text{Beef}$$

$$x_9 := \text{Fish}$$

$$x_{10} := \text{Orange}$$

$$x_{11} := \text{Apple}$$

$$x_{12} := \text{Pudding}$$

$$x_{13} := \text{Jello}$$

Objective Function

We want to find amount of carbohydrates, vitamins, proteins, and fats that meet the minimal requirement per meal, while producing the lowest cost (minimizing cost) we can state the objective function as a linear combination of the decision variables multiplied by their associated cost per serving as such:

minimize

$$z = 0.10x_1 + 0.12x_2 + 0.13x_3 + 0.09x_4 + 0.10x_5 + 0.07x_6 + 0.70x_7 +$$

$$1.20x_8 + 0.63x_9 + 0.28x_{10} + 0.42x_{11} + 0.15x_{12} + 0.12x_{13}$$

Constraints

The constraints associated with this meal planning system is the minimal requirements of carbohydrates, vitamins, protein, and fats per meal. Therefore, each meal will have to contain a greater than or equal to amount of each:

Carbohydrates ≥ 5

Vitamins ≥ 10

Protein ≥ 10

Fats ≥ 2

Explicit Formulation

Given the decision variables and constraints formulated above we can now explicitly formulate the linear programming model as such:

$minimize$

$$0.10x_1 + 0.12x_2 + 0.13x_3 + 0.09x_4 + 0.10x_5 + 0.07x_6 + 0.70x_7 +$$

$$1.20x_8 + 0.63x_9 + 0.28x_{10} + 0.42x_{11} + 0.15x_{12} + 0.12x_{13}$$

$subject\ \ to:$

$$1x_1 + 1x_2 + 1x_3 + 2x_4 + 4x_5 + 2x_6 + 3x_7 + 3x_8 + 1x_9 + 1x_{10} + 1x_{11} + 1x_{12} + 1x_{13} \geq 5$$

$$3x_1 + 5x_2 + 5x_3 + 6x_4 + 2x_5 + 1x_6 + 1x_7 + 8x_8 + 6x_9 + 3x_{10} + 2x_{11} \geq 10,$$

$$1x_1 + 2x_2 + x_3 + 1x_4 + 1x_5 + 1x_6 + 3x_7 + 5x_8 + 6x_9 + 1x_{10} \geq 10,$$

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \geq 1,$$

$$x_7 + x_8 + x_9 \geq 1,$$

$$x_{10} + x_{11} + x_{12} \geq 1,$$

$$x_i, \ldots, x_n \geq 0 \quad for\ \ i = 1, \ldots, n$$

Problem 2

The quality of air in an industrial region largely depends on the effluent emission from n plants. Each plant can use m different types of fuel. Suppose that the total energy needed at plant j is bj British thermal units per day and that eij is the effluent emission per ton of fuel type i at plant j. Further suppose that fuel type i costs ci dollars per ton and that each ton of this fuel type generates ij British thermal units at plant j. The level of air pollution in the region is not to extend b micrograms per cubic meter. Finally, let j be a meteorological parameter relating emissions at plant j to air quality at the region. Formulate the problem of determining the mix of fuels to be used at each plant as a linear program: clearly define the variables, and state the objective function and constraints with proper justification. Be sure that your formulation is written in implicit form.

Solution:

Decision Variables

We are given that there are i types of fuels which varies from 1 to m. We are also given that the plants j vary from 1 to n. Since we need to determine the mix of fuels to be used at each plant we will define the quantity of i type of fuel at plant j as the following:

$$x_{ij}$$

Therefore, the decision variables are:

$$x_{ij} = \text{amount (in tons per day) of fuel } i \text{ used by plant } j \text{ for the mix.}$$

Objective Function

The objective is to determine the minimum total cost of fuels. We are given that the cost of fuel type i costs c_i dollars/ton we have:

$$c_i \cdot (X_{i1} + X_{i2} + X_{i3} + \cdots + X_{in})$$

Thus, for the total cost of all the type of fuels we can represent the objective function as:

$$minimize \;\; z = \sum \sum c_i * X_{ij}, \quad \forall i, j$$

Constraints

The constraints come in the form of pollution tolerance levels such that it is not to exceed levels of b micrograms per cubic meter. Let, P be the meteorological parameter relating to emissions at plant j to air quality in the region. Now, let b_j be the British thermal units at plant j. Further, we are given that e_{ij} is the effluent emissions per ton of fuel type i at plant j, thus we can form the total emission at plant j as

$$\sum_{i=1}^{m} e_{ij} * X_{ij} \leq P_j, \forall i, \forall j = 1, \ldots, n$$

Given that t_i represents thermal unite per ton of fuel type i, then we have the following for the total energy consumed at plant j

$$\sum_{i=1}^{m} t_i * X_{ij} = b_j, \forall i, \forall j = 1, \ldots, n$$

Further,

$$X_{ij} \geq 0, \forall i = 1, \ldots, m, \forall j = 1, \ldots, n$$

Now, we have to consider the constraint of b micrograms

$$\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} c_i * X_{ij},}{\sum t_i * X_{ij}} = b_j, \forall i, j$$

Implicit Formulation

$$minimize \quad {}_{i=1}^{m} \sum \sum_{j=1}^{n} c_i * X_{ij}, \quad \forall i, j$$

subject to

$$\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} c_i * X_{ij},}{\sum_{i=1}^{m} t_i * X_{ij}} = b_j, \forall i, j$$

$$\sum_{i=1}^{m} t_i * X_{ij}, \;\; = b_j, \forall j = 1, \ldots, n$$

$$X_{ij} \geq 0, \forall i = 1, \ldots, m, \forall j = 1, \ldots, n$$

Production Scheduling

A small foundry needs to schedule the production of four different castings during the next week. The production requirements of each casting are summarized in the following table:

Product	Unit production times (minutes)				
	Pouring	Cleaning	Grinding	Inspection	Packing
A	3	8	10	1	3
B	1	12	6	1	5
C	2	6	9	1	3
D	1	7	7	1	2

The unit profit for Products A, B, C, and D are 18,15, $13, and 14$, respectively. Current demands indicate that all castings that are made can be sold; however, contracts dictate that at least 200 units of Product A and 300 units of Product D must be produced. The estimated times available for each of the operations during the next week are:

Pouting	40 hours,
Cleaning	80 hours,
Grinding	80 hours,
Inspection	20 hours,
Packing	40 hours.

The decision problem is to determine how much of each of product should be produced next week to maximize the total profit. Formulate explicitly the linear program for this problem: clearly define the variables, and state the objective function and constraints with proper justification.

Solution:

Decision Variables

Since, the decision problem is to determine how much units of each product should be produced to maximize profit the decision variables are the products A, B, C, and D. We will let $x_1, \ldots, x_4$ represent these products as such:

x_1 := A

x_2 := B

x_3 := C

x_4 := D

Objective Function

Given that the profit per unit of product for A,B,C, and D is 18, 15, 13, and 14 respectively and we are concerned with maximizing the total profit we then set up the objective function using the decision variables and profit per unit:

$$maximize \quad z = 18x_1 + 15x_2 + 13x_3 + 14x_4$$

Constraints

The constraints for production come in the form of time limits. For each of the units there are stages (Pouring, Cleaning, Grinding, Inspection, Packing) with associated minutes per stage that differentiate per product. Further, there are only an alloted total amount of time per operation, as shown in the problem. Thus, to determine the constraints we need to find the linear combination of the total time taken per operation and set this less than or equal to the alloted time hence we have:

Pouring: $40hr = 2400minutes$

$$3x_1 + 1x_2 + 2x_3 + 1x_4 \leq 2400$$

Cleaning: $80hr = 4,800minutes$

$$8x_1 + 12x_2 + 6x_3 + 7x_4 \leq 4,800$$

Grinding: $80hr = 4,800minutes$

$$10x_1 + 6x_2 + 9x_3 + 7x_4 \leq 4,800$$

Inspection: $20hr = 1,200minutes$

$$1x_1 + 1x_2 + 1x_3 + 1x_4 \leq 1,200$$

Packing: $40hr = 2400minutes$

$$3x_1 + 5x_2 + 3x_3 + 2x_4 \leq 2400$$

Further,

$$x_1 \geq 200,$$
$$x_4 \geq 300,$$
$$x_i, \ldots, x_n \geq 0, \quad i = 1, \ldots, 4$$

Explicit Formulation

Thus, we can now express the LP model explicitly as:

$$\begin{aligned}
maximize \quad & z = 18x_1 + 15x_2 + 13x_3 + 14x_4, \\
subject\ \ to \quad & 3x_1 + 1x_2 + 2x_3 + 1x_4 \leq 2400, \\
& 8x_1 + 12x_2 + 6x_3 + 7x_4 \leq 4,800, \\
& 10x_1 + 6x_2 + 9x_3 + 7x_4 \leq 4,800, \\
& 1x_1 + 1x_2 + 1x_3 + 1x_4 \leq 1,200, \\
& 3x_1 + 5x_2 + 3x_3 + 2x_4 \leq 2400, \\
& x_1 \geq 200, \\
& x_4 \geq 300, \\
& x_i, \ldots, x_n \geq 0, \quad i = 1, \ldots, 4
\end{aligned}$$

Reformulation and Visualization

Fractional LP Model Reformulation

Reformulate the following problem as a linear program:

$$\text{maximize} \quad z = \frac{4x_1 + x_2 - 3x_4 + 1}{2x_1 + x_3 + 4x_4 + 3},$$

$$\begin{aligned} \text{subject to} \quad & x_1 - 2x_2 + x_3 + 2x_4 \leq 10, \\ & x_2 - x_3 + 5x_4 \leq 12, \\ & x_i \geq 0, \quad i=1,\ldots 4. \end{aligned}$$

Let

$$j_i = \frac{1}{2x_1 + x_3 + 4x_4 + 3} \cdot x_i, \quad for \ \ i = 1, \ldots, 4 \ \ and \ \ k = \frac{1}{2x_1 + x_3 + 2x_4 + 3}$$

Thus, by substitution we now have:
maximize

$$4j_1 + j_2 - 3j_4 + k$$

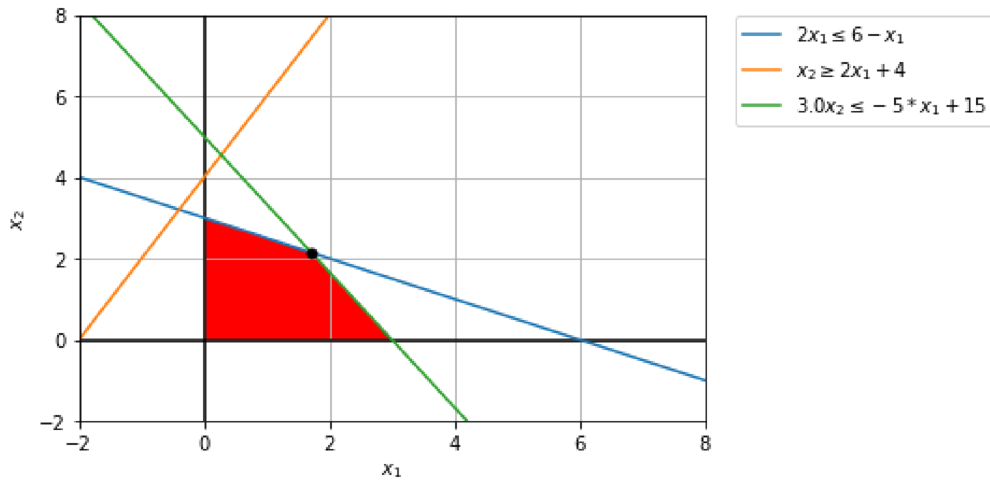
To get the appropriate constraints multiply by k on each side.
$\therefore$ we have the following reformulation of the maximization problem:

$$\begin{aligned} maximize \quad & 4j_1 + j_2 - 3j_4 + k \\ subject \ \ to \quad & j_i - 2j_2 + j_3 \cdot k + 2j_4 - 10k \leq 0, \\ & j_2 - j_3 + 5j_4 - 12k \leq 0 \\ & 2j_1 + j_3 + 4j_4 + 3 \cdot k = 1 \\ & j_i, \ldots, j_n \geq 0, \quad for \ \ i = 1, \ldots, 4 \\ & k \geq z, \quad where \ \ z \in \{n \in \mathbb{Z} : n > 0\} \end{aligned}$$

Note: If j^*, k^* solves the *Linear Programming* problem, then $x^* = \frac{j^*}{k^*}$ solves the original linear fractional program.

LP Graphical Solution (Python)

$$
\begin{aligned}
\text{maximize} \quad & 5x_1 + 4x_2, \\
\text{subject to} \quad & x_1 + 2x_2 \le 6, \\
& -2x_1 + x_2 \le 4, \\
& 5x_1 + 3x_2 \le 15, \\
& x_1, x_2 \ge 0.
\end{aligned}
$$



Then using the corner point of the feasible region we have:

$$z = 5x_1 + 4x_2$$
$$z = 5 * 0 + 4 * 0 = 0 \quad for \quad (0, 0)$$
$$z = 5 * 0 + 4 * 3 = 12 \quad for \quad (0, 3)$$
$$z = 5 * 3 + 4 * 0 = 15 \quad for \quad (3, 0)$$

$\therefore$ the optimal solution is:

$$x^* = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1.714286 \\ 2.142857 \end{pmatrix},$$
$$z = 5 * 1.714286 + 4 * 2.142857 = 17.142828$$

Portfolio Optimization

An approach to portfolio optimization is the scenario approach. In this, we identify a few scenarios (less than 10, say) that might occur during the next year. For each scenario we estimate the return on each investment. Then, we estimate the expected return and risk of the portfolio.
The Sentinel Finance Company, a small firm, wishes to invest in four stocks. The cost of each stock **($ per share)** and the forecasts of the return **($ per share)** for each stock made by the company's five analysts are given in the following table:

	Stock 1	Stock 2	Stock 3	Stock 4
Cost	$30.00	$45.00	$27.00	$53.00
Forecast 1	3.00	13.00	4.00	25.00
Forecast 2	1.00	4.50	.60	15.00
Forecast 3	2.75	1.75	2.75	20.00
Forecast 4	4.50	5.00	1.90	5.00
Forecast 5	3.25	2.75	3.75	35.00
Expected return ($/share)	2.90	5.40	2.60	20.00

Assume that all forecasts are equally likely. Additionally, the finance company would like to invest no more than $100,000. Sentinel has the following requirements for its investment portfolio:
1. Achieve an expected return of at least 10% of total amount invested.
2. Achieve a minimum risk [as measured by the absolute deviation from the expected return (a surrogate for variance)].
3. Invest at least 10% of the total investment in stock 4.

Formulate the above portfolio optimization problem as a linear program.

Solution:

Let the cost of stock j be C_j such that $C_1 = \$30, C_2 = \$45, C_3 = \$27, C_4 = \$53$ and the expected return of stock j be R_j, where

$$R_1 = \frac{\$2.90}{share}, R_1 = \frac{\$5.40}{share}, R_1 = \frac{\$2.60}{share}, R_4 = \frac{\$20.00}{share}$$

.

Let F_{ij} be the i^{th} forecast for stock j and μ_j for the mean forecast. Then the return on a $1.00 investment for stock $j = \frac{R_j}{C_j}$ gives us the following returns for $1.00 investment

$$r_1 = \frac{\$2.90}{\$30.00}, r_2 = \frac{\$5.40}{\$45.00}, r_2 = \frac{\$2.60}{\$27.00}, r_1 = \frac{\$20.00}{\$53.00}$$

The *mean absolute deviation* of stock j is

$$D_j = \frac{1}{n}\sum_{i=1}^{n} |F_{ij} - \mu_j|$$

From Table 1:

$D_1 = \$0.82$

$D_2 = \$3.06$

$D_3 = \$1.08$

$D_4 = \$8.00$

Table 1: Minimum Risk Calculations

	Stock 1	Stock 2	Stock 3	Stock 4
Mean Deviation	$\lvert x-\bar{x}\rvert$	$\lvert x-\bar{x}\rvert$	$\lvert x-\bar{x}\rvert$	$\lvert x-\bar{x}\rvert$
Forecast 1	$\lvert 3.00-2.90\rvert = 0.10$	$\lvert 13.00-5.40\rvert = 7.60$	$\lvert 4.00-2.60\rvert = 1.40$	$\lvert 25.00-20.00\rvert = 5.00$
Forecast 2	$\lvert 1.00-2.90\rvert = 1.90$	$\lvert 4.50-5.40\rvert = 0.90$	$\lvert 0.60-2.60\rvert = 2.00$	$\lvert 15.00-20.00\rvert = 5.00$
Forecast 3	$\lvert 2.75-2.90\rvert = 0.15$	$\lvert 1.75-5.40\rvert = 3.65$	$\lvert 2.75-2.60\rvert = 0.15$	$\lvert 20.00-20.00\rvert = 0.00$
Forecast 4	$\lvert 4.50-2.90\rvert = 1.60$	$\lvert 5.00-5.40\rvert = 0.40$	$\lvert 1.90-2.60\rvert = 0.70$	$\lvert 5.00-20.00\rvert = 15.00$
Forecast 5	$\lvert 3.25-2.90\rvert = 0.35$	$\lvert 2.75-5.40\rvert = 2.65$	$\lvert 3.75-2.60\rvert = 1.15$	$\lvert 35.00-20.00\rvert = 15.00$
$sum\lvert x-\bar{x}\rvert /n$	$(4.10)/5 = \$0.82$	$(15.20)/5 = \$3.04$	$(5.40)/5 = \$1.08$	$(40.00)/5 = \$8.00$

Decision Variables

Since we are seeking to find out which of the our stocks The Sentinel Finance Company wishes to invest in, these four stocks will be the decision variables:

X_1 := Stock 1

X_2 := Stock 2

X_3 := Stock 3

X_4 := Stock 4

Objective Function

Each stock has an associated cost of \$30.00 for X_1, \$45.00 for X_2, \$27.00 for X_3, and \$53.00 for X_4 and the firm, being risk adverse investors, want to know amount of these stocks to invest in that will give them the largest return for the portfolio of investments with the least amount of exposure to risk. Thus, this gives rise to the following objective function:

$$Minimize \;\; (risk) \quad Z = \sum_{j=1}^{m}\sum_{i=1}^{n}\left(\tfrac{D_j}{C_j}\right)X_i, \quad for \;\; i,j = 1\ldots,4$$

$$Minimize \;\; (risk) \quad Z = \$\frac{\$0.82}{\$30.00}X_1 + \frac{\$3.06}{\$45.00}X_2 + \frac{\$3.06}{\$45.00}X_3 + \frac{\$8.00}{\$53.00}X_4$$

$$Minimize \;\; (risk) \quad Z = 0.0273X_1 + 0.0680X_2 + 0.0400X_3 + 0.1509X_4$$

Constraints

The firm has specific requirements that must be met:
1: Amount invested must be less than $\$100,000.00$:

$$\sum_{j=1}^{n} X_j \leq 100000, \quad for \;\; j = 1,\ldots,4$$

2: Achieve an expected return of at least 10% of total amount invested.

$$\sum_{j=1}^{n} r_j X_j \geq 0.1 \sum_{j=1}^{n} X_j, \quad for \;\; j = 1, \ldots, 4$$

$$\$2.90X_1 + \$5.40X_2 + \$2.60X_3 + \$20.00X_4 \geq (\$30.00X_1 + \$45.00X_2 + \$27.00X_3 + \$53.00X_4) \cdot 10\%$$

$$\frac{\$2.90}{\$30.00}X_1 + \frac{\$5.40}{\$45.00}X_2 + \frac{\$2.60}{\$27.00}X_3 + \frac{\$20.00}{\$53.00}X_4 \geq (X_1 + X_2 + X_3 + X_4) \cdot 0.10$$

$$0.0033X_1 - 0.02X_2 + 0.0037X_3 - 0.2774X_4 \leq 0$$

3. Invest at least 10% of the total investment in stock 4.

$$X_4 \geq 0.10 \sum_{j=1}^{n} jX_j$$

4.Further, each investment in stocks $X_i, \ldots, X_n$ cannot be negative. Thus,

$$X_j \geq 0, \forall j$$

Therefore, we have the following constraints for the LP model:

Explicit Formulation:

$$Minimize \;\; (risk) \quad Z = 0.0273X_1 + 0.0680X_2 + 0.0400X_3 + 0.1509X_4$$

$$Subject \;\; to:$$

$$\sum_{j=1}^{n} X_j \leq 100000, \quad for \;\; j = 1, \ldots, 4$$

$$\sum_{j=1}^{n} r_j X_j \geq 0.10 \sum_{j=1}^{n} jX_j, \quad for \;\; j = 1, \ldots, 4$$

$$X_4 \geq 0.10 \sum_{j=1}^{n} jX_j$$

$$X_j \geq 0, \forall j$$

Non-Optimal Solution Proof

Consider the linear program: maximize $c\mathbf{x}$, subject to $Ax \leq b$,$x \geq \overline{0}$, where c is a nonzero vector. Suppose that the point x_0 is such that $Ax_0 < b$ and $x_0 > \overline{0}$. Show that x_0 cannot be an optimal solution.

Solution:

We are given $Ax_0 < b$ and $x_0 > \overline{0}$. Therefore, x_0 is inside the feasible region $Ax_0 \leq b$. Thus it is not an optimal solution, otherwise it would be on a constraint of the region, and not inside the feasible region.

Proof. Let $x_0 + \lambda c^T$. Choose $\lambda > z$ such that $z \in \{n \in \mathbb{Z} : n > 0\}$ where z is sufficiently small. Then we have

$$A\left(x_0 + \lambda c^T\right) < b$$

and

$$x_0 + \lambda_c^T > 0$$

Thus, $x_0 + \lambda_0 c^T$ is feasible. Additionally,

$$c\left(x_0 + \lambda_0 c^T\right) = cx_0 + \lambda_0 cc^T = cx_0 + \lambda_0 \|x\|^2 > cx_0$$

Thus, $c \neq \overline{0}$. Therefore, x_0 cannot be an optimal solution $\square$

Appendix

Sources

Luenberger, D. G. (2014). Investment science (2nd ed.). New York: Oxford University Press.

Ragsdale, C. T. (2018). Spreadsheet modeling and decision analysis: a practical introduction to business analytics (8th ed.). Boston, MA: Cengage Learning.

Python Code: LP Model Graphical Solution

import numpy from numpy import arange import matplotlib.pyplot from matplotlib.pyplot import *

figure() #Construct lines
#$x1 > 0$
x = arange(-10, 20, 1)
#$x_2 >= 0$
y = arange(-10, 20, 1)

#$2x1 <= 6 - x1$
y1 = (6-x)/2.0

#$x2 <= 2x1 + 4$
y2 = (2*x + 4.0)

#$x2 <= (-5 * x1 + 15)/3.0$
y3 = (-5*x + 15)/3.0

xlim(-2, 8)
ylim(-2,8)
hlines(0,-2,10, color = ”k”)
vlines(0,-2, 10, color = ”k”)
grid(True)

#Make plot
#plt.plot(x, c1, label=r'$x_1 \geq 0$')
plot(x, c2, label=r'$2x_1 \leq 6 - x_1$')
plot(x, c3, label=r'$x_2 \geq 2x_1 + 4$')
plot(x, c4, label=r'$3.0x_2 \leq -5 * x_1 + 15$')

xlabel(r'x_1')
ylabel(r'x_2')

legend(bbox_to_anchor=(1.05, 1), loc=2, borderaxespad=0.)
x = [0, 0 , 1.714286 ,3]
y = [0, 3,2.142857, 0]
plt.plot(1.714286, 2.142857, marker='o', markersize=5, color=”black”)
fill(x,y, 'r')
show()