Assignment 7 Prashant Kubsad 06/29/2020

Question 15.2.1 1. Formulate an optimization model (a linear program) to find the cheapest diet that satisfies the maximum and minimum daily nutrition constraints, and solve it using PuLP.

Answer: As mentioned in the office hours video, it is a good idea to solve the problem with Math equation first and then convert it to python program. Lets assume the following variables for the given data set:

```
\begin{split} &C_i = Cost \ of \ food \ i \\ &a_{ji} = Amount \ of \ nutrients \ j \ for \ each \ food \ i. \\ &Min_j = minimum \ nutrients \ needed \ for \ food. \\ &Max_j = maximum \ value \ for \ nutrient \ j \ that \ the \ optimal \ food \ solution \ cannot \ exceed. \\ &X_i = amount \ of \ food \ i \end{split}
```

Aim is to reduce the cost, we can write objective function as:

• Minimize $\sum_i C_i X_i$

We can write the constraints as:

- $\sum_{i} a_{ji} x_{i} \ge = \min_{j}$
- $\sum_{i} a_{ji} x_{i} \leq max_{j}$
- xi>=0 (python program errors out if this is not there and rightly so)

The same logic I have written it in the python using PyCharm. Thanks to office hours session, it was good starting point for the assignment.

Steps followed:

- 1) Read the given spreadsheet raw-data
- 2) Convert the raw-data to list which we can navigate through/manipulate.
- 3) Extract each column into its own variable. The resulting list will have elements in the format <food>:<nutrientvalue> for all the food items.
 - ex.: For calories 'Frozen Broccoli': 73.8, 'Carrots, Raw': 23.7
- 4) Extract min and max constraints from the spreadsheet at the locations.
- 5) Start the optimization problem of minimizing the cost
- 6) Declare the objective variable that needs to be determined: amount of food to be served.
- 7) Write objective function
- 8) Define all the constraints using the list created in step 4.
- 9) Solve the problem and print the output.

Code:

```
from pulp import *;
import pandas as pa;
import xlrd as xlrd;
#1.read the excel raw data
foodDataRaw = pa.read_excel("diet.xls")
```

```
totalFats = \{x[0]: float(x[5]) \text{ for } x \text{ in } foodData\}
carbohydrates = \{x[0]: float(x[7]) for x in foodData\}
minCholestrol = foodDataRaw['Cholesterol mg'].values[65]
maxTotalFat = foodDataRaw['Total Fat g'].values[66]
minSodium = foodDataRaw['Sodium mg'].values[65]
maxSodium = foodDataRaw['Sodium mg'].values[66]
minCarbohydrates = foodDataRaw['Carbohydrates q'].values[65]
maxCarbohydrates = foodDataRaw['Carbohydrates g'].values[66]
minFiber = foodDataRaw['Dietary Fiber g'].values[65]
maxFiber = foodDataRaw['Dietary Fiber g'].values[66]
maxCalcium = foodDataRaw['Calcium mg'].values[66]
minIron = foodDataRaw['Iron mg'].values[65]
print("mins: ", minCalories, minCholestrol, minTotalFat, minSodium, minCarbohydrates,
orint("max:", maxCalories, maxCholestrol, maxTotalFat, maxSodium, maxCarbohydrates,
```

```
maxFiber, maxProtien,
     maxVitaminA, maxVitaminC, maxCalcium, maxIron)
problem = LpProblem("Diet Problem", LpMinimize)
problem += lpSum([carbohydrates[i] * amountVars[i] for i in foods]) >= minCarbohydrates,
problem += lpSum([iron[i] * amountVars[i] for i in foods]) <= maxIron, 'max iron'</pre>
problem.solve()
```

```
varsDictionary = {}
#print the output: food combination to meet the constraints and keep cost down
for v in problem.variables():
    varsDictionary[v.name] = v.varValue
    if(v.varValue>0):
        print(v.name , ": ", v.varValue)

#print the cost
print("Total cost: ",pulp.value(problem.objective))
```

Output:

```
Run: Assignment7 ×

/Library/Frameworks/Python.framework/Versions/3.8/bin/python3 /Users/prashantkubsad/Edx/homeworks/7/PythonPulp/Assignment7.py
mins: 1500.0 30.0 20.0 800.0 130.0 125.0 60.0 1000.0 400.0 700.0 10.0
max: 2500.0 240.0 70.0 2000.0 450.0 250.0 100.0 10000.0 5000.0 1500.0 40.0

Amounts_Celery,_Raw : 52.64371

Amounts_Frozen_Broccoli : 0.25960653

Amounts_Lettuce,Iceberg,Raw : 63.988506

Amounts_Oranges : 2.2929389

Amounts_Poached_Eggs : 0.14184397

Amounts_Popcorn,Air_Popped : 13.869322

Total cost: 4.3371167974

Process finished with exit code 0
```

Based on the output, the optimized combination is:

```
Amounts_Celery,_Raw: 52.64371
Amounts_Frozen_Broccoli: 0.25960653
Amounts_Lettuce,Iceberg,Raw: 63.988506
Amounts_Oranges: 2.2929389
Amounts_Poached_Eggs: 0.14184397
Amounts_Popcorn,Air_Popped: 13.869322
Total cost: 4.3371167974
```

Question 15.2.a. Please add to your model the following constraints (which might require adding more variables) and solve the new model: a. If a food is selected, then a minimum of 1/10 serving must be chosen. (Hint: now you will need two variables for each food i: whether it is chosen, and how much is part of the diet. You'll also need to write a constraint to link them.)

Answer: As mentioned in the homework question, we have to extract 2 more variables to identify if a food item has been chosen. This will be a binary type variable. Add one more constraint to the constraints list that we had in the solution for first part of this problem. As explained in the office hours, there is no maximum constraint defined for the amount of food, we can put in arbitrary large value for maximum constraint. With this, any realistic number will be less than that and will satisfy the constraint. We can write the new constraints as:

```
a_i= Amount of food chosen for food i c_i= whether food "i" is chosen or not a_i >= 0.1 * c_i
a_i <= 10000 * c_i
```

Question 15.2.b. Many people dislike celery and frozen broccoli. So at most one, but not both, can be selected.

Answer: We have to add one more constraint to the above list. Lets assume a_{celery} and a_{fbroccoli} are the amounts of celery and frozen broccoli. We can write the constraint as:

```
a_{celery} + a_{fbroccoli} \le 0;
```

The above equation will force us to pick either one or not pick both.

Question 15.2.b. To get day-to-day variety in protein, at least 3 kinds of meat/poultry/fish/eggs must be selected.

Answer: We have to add more constraint to the optimization problem. Lets assume the food a1, a2, a3, a4,a5 are all amounts of protein foods. We can write the above constrain as

```
\sum_{i} (a_i) >= 3. Translates to a1+a2+a3+a4+a5 >3
```

Added all the above mentioned constraints to the program:

```
foodData = foodDataRaw[0:64].values.tolist()
carbohydrates = \{x[0]: float(x[7]) for x in foodData\}
```

```
minCarbohydrates = foodDataRaw['Carbohydrates g'].values[65]
maxCarbohydrates = foodDataRaw['Carbohydrates g'].values[66]
maxProtien = foodDataRaw['Protein g'].values[66]
minProtien = foodDataRaw['Protein g'].values[65]
minVitaminC = foodDataRaw['Vit C IU'].values[65]
maxIron = foodDataRaw['Iron mg'].values[66]
print("max:", maxCalories, maxCholestrol, maxTotalFat, maxSodium, maxCarbohydrates,
maxFiber, maxProtien,
amountVars = LpVariable.dicts("Amounts", foods, 0)
problem2a += lpSum([calories[i] * amountVars[i]    for i in foods]) >= minCalories,
problem2a += lpSum([calories[i] * amountVars[i]    for i in foods]) <= maxCalories,
```

```
problem2a += lpSum([carbohydrates[i] * amountVars[i] for i in foods]) >=
minCarbohydrates, 'min carbohydrates'
problem2a += lpSum([protein[i] * amountVars[i] for i in foods]) >= minProtien,
problem2a += lpSum([vitaminA[i] * amountVars[i] for i in foods]) <= maxVitaminA,
problem2a += lpSum([calcium[i] * amountVars[i] for i in foods]) >= minCalcium,
problem2a += lpSum([iron[i] * amountVars[i] for i in foods]) <= maxIron, 'max iron'</pre>
problem2a += chosenFoods['Frozen Broccoli'] + chosenFoods['Celery, Raw'] <= 1, 'broccoli</pre>
 chosenFoods['Scrambled Eggs'] + chosenFoods['Frankfurter, Beef'] + \
 chosenFoods['Kielbasa,Prk'] + chosenFoods['Hamburger W/Toppings'] + \
    if (v.varValue > 0):
            print(str(v.varValue) + " units of " + str(v.name))
```

```
#print the cost
print("Total cost 2a: ", pulp.value(problem2a.objective))
```

Output:

```
Run: Assignment7_2 ×

/Library/Frameworks/Python.framework/Versions/3.8/bin/python3 /Users/prashantkubsad/Edx/homeworks/7/PythonPulp/Assignment7_2.py
mins: 1500.0 30.0 20.0 800.0 130.0 125.0 60.0 1000.0 400.0 700.0 10.0
max: 2500.0 240.0 70.0 200.0 800.0 150.0 1000.0 10000.0 5000.0 1500.0 40.0

### 2500.0 240.0 70.0 200.0 20.0 800.0 1500.0 10000.0 5000.0 1500.0 40.0

### 2500.0 240.0 70.0 200.0 100.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10000.0 10
```

As per the above output the optimal combination is:

```
42.399358 servings of Amounts_Celery,_Raw
0.1 servings of Amounts_Kielbasa,Prk
82.802586 servings of Amounts_Lettuce,Iceberg,Raw
3.0771841 servings of Amounts_Oranges
1.9429716 servings of Amounts_Peanut_Butter
0.1 servings of Amounts_Poached_Eggs
13.223294 servings of Amounts_Popcorn,Air_Popped
0.1 servings of Amounts_Scrambled_Eggs
Total cost 2a: 4.512543427
```

Extra question: If you want to see what a more full-sized problem would look like, try solving your models for the file diet_large.xls, which is a low-cholesterol diet model (rather than minimizing cost, the goal is to minimize cholesterol intake).

The solution to this question is basically extending the previous solution to include a larger data set, lot more constraints and objective function is to minimize cholesterol instead of cost.

Assumptions:

- 1) Any empty cell will be treated as 0.
- 2) Any columns/nutrients which do not have constraints are not significant to the problem. I confirmed this assumption with TA in the office hrs.
- 3) Column D and Column F both are marked as energy, I have changed them to energy1 and energy2 when I am extracting them in the code.

Steps Followed: In addition to the steps mentioned for the first problem, these are the extra changes that I have done to the program:

- 1) Loop through all the data cell and convert any empty cell to 0. Pandas package utility to parse the spread sheet marks all empty cells as Nan.Numpy package has a function isnan which detects if any cell is Nan.
- 2) Instead of reading constraints one by one like before, I have ran a loop from first column to last column on the rows where constraints are present. This results in a 2 dimensional array. The first element and last 3 elements are Nan.
- 3) Remove the Nan columns from the constraints.
- 4) Change the variable in the objective function to cholesterol instead of cost.

Code:

```
import numpy as np;
phosphorous = \{x[0]: float(x[9]) for x in foodData\}
potassium = \{x[0]: float(x[10]) for x in foodData\}
copper = \{x[0]: float(x[13]) for x in foodData\}
pantothenic = \{x[0]: float(x[23]) for x in foodData\}
```

```
minimumConstraints = foodDataRaw[7147:7148].values.tolist()
del minimumConstraints[0][0]
float (maximumConstraints[0][0]))
problem += lpSum([energy[i] * amountVars[i] for i in foods]) >= float(minimumConstraints[0][2])
problem += lpSum([energy[i] * amountVars[i] for i in foods]) <= float(maximumConstraints[0][2])</pre>
problem += lpSum([water[i] * amountVars[i] for i in foods]) <= float(maximumConstraints[0][3])</pre>
problem += lpSum([iron[i] * amountVars[i] for i in foods]) >= float(minimumConstraints[0][6])
problem += lpSum([iron[i] * amountVars[i] for i in foods]) <= float(maximumConstraints[0][6])</pre>
problem += lpSum([potassium[i] * amountVars[i] for i in foods]) >= float(minimumConstraints[0][9])
problem += lpSum([potassium[i] * amountVars[i] for i in foods]) <= float(maximumConstraints[0][9])</pre>
```

```
problem += lpSum([copper[i] * amountVars[i] for i in foods]) >= float(minimumConstraints[0][12])
problem += lpSum([riboflavin[i] * amountVars[i] for i in foods]) >=
problem += lpSum([niacin[i] * amountVars[i] for i in foods]) >= float(minimumConstraints[0][21])
problem += lpSum([pantothenic[i] * amountVars[i] for i in foods]) >=
    if(var.varValue>0):
```

output:

```
Run: Assignment7_large ×

/Library/Frameworks/Python.framework/Versions/3.8/bin/python3 /Users/prashantkubsad/Edx/homeworks/7/PythonPulp/Assignment7_large.py
minium constraints list: [[nan, 56, 130, 2400, 3700, 2400.0, 1000, 8, 270, 700, 4700, 1500, 11, 0.9, 2.3, 55, 900, 15, 200, 90, 0.0012, imaximum constraints list: [[nan, 1000000, 1000000, 1000000, 1000000, 2500, 45, 400, 4000, 1000000, 2300, 40, 10, 11, 400, 300]

Amounts_Leavening_agents,_baking_powder,_low_sodium : 0.20190712

Amounts_Oil,_vegetable,_sheanut : 1.5105448

Amounts_Peanuts,_all_types,_cooked,_boiled,_with_salt : 1.7091785

Amounts_Peanuts,_all_types,_cooked,_boiled,_with_salt : 1.7091785

Amounts_Seeds,_breadfruit_seeds,_roasted : 0.8512066

Amounts_Soybeans,_mature_seeds,_roasted : 0.8512066

Amounts_Water,_bottled,_non_carbonated,_EVIAN : 36.104705

Total cholestrol: 0.0
```

Based on the above output, the optimal combination of foods that meets all the constraints and keep cholesterol to zero is:

```
Amounts_Leavening_agents,_baking_powder,_low_sodium: 0.20190712

Amounts_Oil,_vegetable,_sheanut: 1.5105448

Amounts_Peanuts,_all_types,_cooked,_boiled,_with_salt: 1.7091785

Amounts_Puddings,_KRAFT,_JELL_O_Brand_Fat_Free_Sugar_Free_Instant_Reduc: 0.64694195

Amounts_Seeds,_breadfruit_seeds,_roasted: 0.8512066

Amounts_Soybeans,_mature_seeds,_raw: 1.1410814

Amounts_Water,_bottled,_non_carbonated,_EVIAN: 36.104705

Total cholestrol: 0.0
```