MASTER OF DATA SCIENCE (SEMESTER 2 – 2022/2023)

FACULTY OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

WQD7011 NUMERICAL OPTIMIZATION

GROUP ASSIGNMENT

Title: Optimizing Dietary Planning Using the Simplex Method

Instructor: Dr Lim Chee Kau

Group Members	**Student ID**
Nasir Uddin Ahmed	**S2015449**
Yin QiXiang	**S2150692**
Ronjon Ahmed	**S2150527**
Alfian Rosisaputra	**S2148301**

1. Introduction

Maintaining a balanced, healthy diet is crucial for general well-being in today's fast-paced society. However, it can be difficult to provide optimal nutrition while taking into account different dietary restrictions. Here, mathematical optimization approaches can be used to solve difficult food planning issues in a methodical manner.

A popular linear programming approach called the Simplex Method offers a potent tool for enhancing decision-making and resource allocation. The Simplex Method can be used to formulate dietary planning as a linear programming problem in order to discover the best solution that satisfies nutritional needs while taking into account restrictions like price, and daily nutrient requirements.

The Diet_KSA_2020 (1) dataset from Kaggle will be used in this assignment to optimize a dietary planning problem using the simplex approach. The dataset includes details on the cost and nutritional value of 10 different foods. Finding meal combinations that satisfy a person's daily nutritional needs while being as inexpensive as possible is our aim.

By showing how the Simplex Method may be used in practice and offering insightful information on the advantages and difficulties of applying mathematical optimization approaches for dietary planning, we hope to make a positive impact on the field of dietary optimization.

In this assignment, we have three targets to achieve. The first one is understanding the basic concepts of the simplex method. The second one is building the objective function and identifying the constraints. And the third one is to implement the simplex method using octave programming.

By making dietary planning as effective as possible, we can enable people to make knowledgeable decisions about their nutrition, improving health outcomes and all-around well-being.

2. Dataset Overview

The "Diet_KSA_2020" dataset that we are using for our project is taken from Kaggle. This dataset offers thorough details on the most popular food items consumed in Saudi Arabia, including their nutrient profiles and price tags. The dataset is intended primarily for the use of linear programming methods in order to optimize meal preparation.

The "Diet_KSA_2020" dataset consists of a CSV (Comma-Separated Values) file, enabling easy access and manipulation of the data. It comprises several columns, each containing specific information about the food items. The following is a list of the columns in the dataset and a brief description of each

Table 1: Dataset Description

Attribute Names	Description
index	The index of the food in the dataset.
name	The name of the food.
max_serv	The maximum number of servings of the food that can be consumed in one day.
size	The size of one serving of the food in grams.
energy	The amount of energy in one serving of the food in calories.
protein	The amount of protein in one serving of the food in grams.
calcium	The amount of calcium in one serving of the food in milligrams.
magnesium	The amount of magnesium in one serving of the food in milligrams.
vitaminc	The amount of vitamin C in one serving of the food in milligrams.
iron	The amount of iron in one serving of the food in milligrams.
price	The price of one serving of the food in dollars.

3. Problem Statement

This assignment's goal is to use the Simplex Method to select the best food products from the "Diet_KSA_2020 (1)" dataset in order to satisfy a person's daily nutritional needs while minimizing the associated costs. The aim is to determine the number of food products to include in a diet plan that satisfies the individual's daily nutritional requirements, given the nutritional information and cost data for each food item.

4. Literature Review

Promoting health and well-being requires careful consideration of nutrition planning. With the growing concern for nutrition and sustainability, researchers have explored various approaches to enhance the effectiveness of dietary planning methods. One such approach involves the application of mathematical optimization techniques to find optimal combinations of foods that meet specific nutritional requirements while considering various constraints. This literature review explores the use of optimization methods, such as the Simplex Method, Evolutionary Algorithms, and Linear Programming (LP), to optimize dietary planning and minimize costs while considering various constraints and objectives.

The optimization of dietary planning has been addressed using various methods. Seljak (2006) introduced a multi-objective and multi-constrained evolutionary algorithm for multi-level menu planning, which quickly finds a diverse set of feasible solutions that are both nutritionally and gastronomically adequate. The study highlights the importance of optimization techniques in generating nutritionally adequate menus with the lowest objective function values.

Gazan et al. (2018) conducted a narrative review highlighting the use of mathematical diet optimization, specifically LP, in understanding the relations between different dimensions of diet sustainability. LP-based approaches have been used to identify sustainable diets that fulfill multiple objectives and constraints related to nutrition, cost, and environmental impact. The review emphasizes the importance of carefully choosing model parameters and input data to correctly interpret and communicate the results.

Jardim et al. (2013) applied a nonlinear optimization tool to balance diets with constant metabolizability. The study used a nonlinear programming problem to determine optimal combinations of foods that maintain consistent diet qualities across different planes of nutrition. The research demonstrated the ability of the optimization tool to provide solutions for maintaining dietary consistency while considering multiple nutrient intakes and the effect of the plane of nutrition.

Pe et al. (2022) focused on cost optimization for weekly meal planning of college students based on calorie constraints using the LP method. The study aimed to minimize costs while satisfying the nutrient requirements of college students. LP techniques were employed to develop cost-effective dietary plans considering different factors such as sex, physical activity, and budget constraints. The results showed that LP was able to obtain lower costs compared to other methods, providing well-balanced diet options within budget limitations.

By synthesizing the findings from these studies, this literature review will shed light on the efficacy of the Simplex method in optimizing dietary planning. It will provide valuable insights into how mathematical optimization techniques can assist in designing nutritionally balanced, culturally acceptable, and economically feasible diets. Additionally, the review will emphasize the importance of careful parameter selection, input data quality, and expertise in interpreting and communicating the results of optimization algorithms.

5. Objective Function Identification

The linear programming problem's purpose or objective is represented by the objective function in the simplex technique. Depending on the demands of the situation, it quantifies the quantity that needs to be minimized or maximized. A linear combination of choice variables, each multiplied by its corresponding coefficient, makes up the objective function in most cases. The significance or contribution of each variable to the overall goal is shown by these coefficients. Finding the choice variable values that optimize the objective function is the goal of the simplex approach.

Decision variables play a crucial role in formulating and solving optimization problems using the simplex method. They represent the unknown quantities or values that need to be determined in order to optimize the objective function while satisfying the given constraints.

In our dataset, we have 10 food items. So we can say that we have 10 decision variables. The decision variables are introduced below.

x1 corresponds to the quantity of "**Oatmeal**", **x2** corresponds to the quantity of "**Chicken**", **x3** corresponds to the quantity of "**Eggs**", **x4** corresponds to the quantity of "**Whole Milk**", **x5** corresponds to the quantity of "**Cherry pie**", **x6** corresponds to the quantity of "**Pork with beans**", **x7** corresponds to the quantity of "**Yogurt**", **x8** corresponds to the quantity of "**Rice**", **x9** corresponds to the quantity of "**Meat**", **x10** corresponds to the quantity of "**Potato**".

The objective of this problem is to minimize the cost of a diet plan while meeting the daily nutritional requirements. The goal is to determine the optimal quantities of food items (x1, x2, ..., x10) to include in the diet plan, subject to certain constraints. So in this case we formulate our objective function as mentioned below.

Min 11***x1** + 16***x2** + 2***x3** + 8***x4** + 25***x5** + 4***x6** + 10***x7** + 10***x8** + 40***x9** + 3***x10**

6. Identifying Constraints

In linear programming, constraints specify the restrictions or requirements that the decision variables must meet. These constraints aid in simulating the actual demands, limitations, or resource availability in the optimization problem. Aspects like resource availability, capacity, budget, or demand may be restricted by constraints. They specify the constraints that must be met by the decision variables in order to produce a workable solution.

The dataset's max_serv and size columns, which describe the maximum serving size and the size of each food item, respectively, are not taken into account while formulating the constraint. Since the constraints are more concerned with the overall nutritional content than with the precise serving sizes or sizes of the food items, these numbers are not explicitly employed in the formulation of the constraints. The goal is to satisfy dietary needs while making the best possible choice of food items based on nutritional content and price.

We formulated a total of 6 constraints which are given below

110***x1** + 205***x2** + 160***x3** + 160***x4** + 420***x5** + 260***x6** + 58***x7** + 500***x8** + 143***x9** + 323***x10** ≥ **Daily Energy Requirement**

4***x1** + 32***x2** + 13***x3** + 8***x4** + 4***x5** + 14***x6** + 10***x7** + 10***x8** + 26***x9** + 2***x10** ≥ **Daily Protein Requirement**

2***x1** + 12***x2** + 54***x3** + 285***x4** + 22***x5** + 80***x6** + 110***x7** + 28***x8** + 6***x9** + 12***x10** ≥ **Daily Calcium Requirement**

26***x1** + 1***x2** + 10***x3** + 0***x4** + 9***x5** + 31***x6** + 11***x7** + 2501***x8** + 29***x9** + 28***x10** ≥ **Daily Magnesium Requirement**

0***x1** + 2***x2** + 0***x3** + 0***x4** + 1***x5** + 1***x6** + 0***x7** + 9***x8** + 0***x9** + 19***x10** ≥ **Daily VitaminC Requirement**

6***x1** + 1***x2** + 1***x3** + 11***x4** + 2***x5** + 1***x6** + 0***x7** + 9***x8** + 1***x9** + 1***x10** ≥ **Daily Iron Requirement**

Here x1, x2, x3, x4, x5, x6, x7, x8, x9, x10 ≥ 0

Daily Energy Requirement, Daily Protein Requirement, Daily Calcium Requirement, Daily Magnesium Requirement, Daily VitaminC Requirement, Daily Iron Requirement; these things actually vary based on age, and gender. Health.gov is the official website of the U.S. Department of Health and Human Services (HHS). It serves as a trusted source of information and resources related to health, wellness, and healthcare in the United States. From their (Table E3.1.A4. Nutritional goals for each age/sex group used in assessing the adequacy of USDA Food Patterns at various calorie levels) we have assigned values for Daily Energy Requirement, Daily Protein Requirement, Daily Calcium Requirement, Daily Magnesium Requirement, Daily VitaminC Requirement, Daily Iron Requirement. And our constraints now look like

110***x1** + 205***x2** + 160***x3** + 160***x4** + 420***x5** + 260***x6** + 58***x7** + 500***x8** + 143***x9** + 323***x10** ≥ **2000**

4***x1** + 32***x2** + 13***x3** + 8***x4** + 4***x5** + 14***x6** + 10***x7** + 10***x8** + 26***x9** + 2***x10** ≥ **50**

2***x1** + 12***x2** + 54***x3** + 285***x4** + 22***x5** + 80***x6** + 110***x7** + 28***x8** + 6***x9** + 12***x10** ≥ **1000**

26***x1** + 1***x2** + 10***x3** + 0***x4** + 9***x5** + 31***x6** + 11***x7** + 2501***x8** + 29***x9** + 28***x10** ≥ **400**

0***x1** + 2***x2** + 0***x3** + 0***x4** + 1***x5** + 1***x6** + 0***x7** + 9***x8** + 0***x9** + 19***x10** ≥ **90**

6***x1** + 1***x2** + 1***x3** + 11***x4** + 2***x5** + 1***x6** + 0***x7** + 9***x8** + 1***x9** + 1***x10** ≥ **18**

In short, we have 6 constraints and 10 decision variables. This means the number of decision variables> constraints, which satisfies the condition of the simplex method.

7. Numerical Optimization Technique

Diet optimization is a crucial task in ensuring proper nutrition while managing costs. This assignment aims to leverage the Simplex Method to find the most cost-effective combination of foods that meet daily nutritional requirements. By utilizing linear programming techniques, this approach optimizes diet planning for improved health and affordability. A system diagram is given below to show how the simplex method is working on this problem.

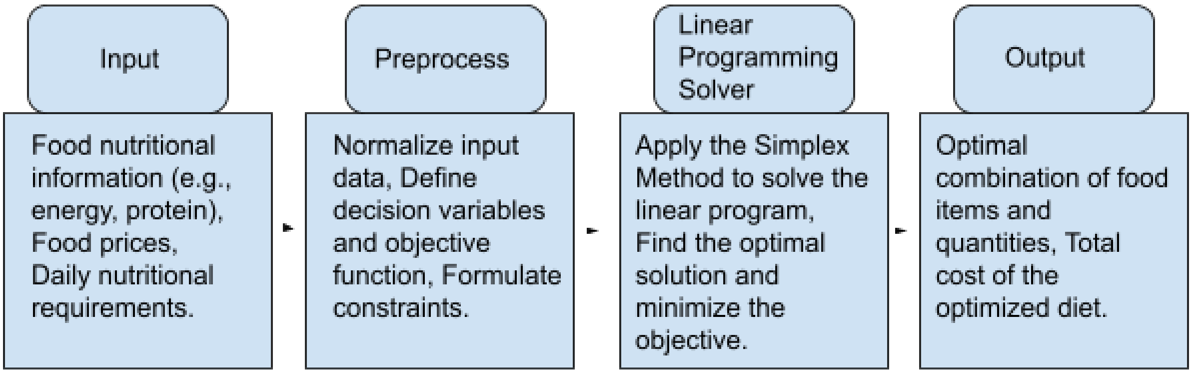

Figure 1: Numerical Optimization Technique Working Process

8. Modeling

For modeling, we have chosen the GNU octave with linear programming "glpk" function. GLPK (GNU Linear Programming Kit) is a powerful open-source software package designed for solving linear programming (LP) and mixed-integer programming (MIP) problems. It provides a flexible and efficient framework for modeling, solving, and analyzing optimization problems.

[xmin,fmin,status,extra] = glpk(C,A,b,lb,ub,cType,varType,sense); this is the modeling function. **C** is the Coefficient of the Objective Function, **A** is the Coefficient of the Constraints, **b** is the Column Array of the Constraints, **lb** is the Lower Bound, **ub** is the Upper Bound, **cType** is the type of the constraint, **varType** is the variable type, **sense = 1** signifies that the linear programming problem is a minimization.

Using the objective function and constraints all the parameters are set, and we ran the code. The vector **xmin** reflects the best solution for the choice variables in the context of linear programming. The vector's elements, x1 through x10, each represent a decision variable.

The values in xmin indicate the optimal values for each decision variable that maximize the objective function while satisfying the given constraints. Let's analyze the values

x1: The optimal value for x1 is 0, which means it is not included in the optimal solution. It indicates that the objective function does not benefit from increasing x1 and that the optimal solution lies at the boundary where x1 is 0. **x2**: Similarly, the optimal value for x2 is 0, indicating that it is also not included in the optimal solution. **x3**: The optimal value for x3 is approximately 1.1481. This value suggests that including x3 in the optimal solution contributes positively to minimizing the objective function. **x4**: The optimal value for x4 is approximately 3.0837. It indicates that x4 is an important decision variable in the optimal solution and contributes significantly to minimizing the objective function. **x5**: The optimal value for x5 is 0, indicating that it is not included in the optimal solution. **x6**: Similarly, the optimal value for x6 is 0, suggesting that it is not part of the optimal solution. **x7**: The optimal value for x7 is 0, indicating that it is not included in the optimal solution. **x8**: The optimal value for x8 is approximately 0.1029. This value suggests that including x8 in the optimal solution contributes positively to minimizing the objective function, although to a lesser extent compared to x3 and x4. **x9**: The optimal value for x9 is 0, indicating that it is not part of the optimal solution. **x10**: The optimal value for x10 is approximately 4.6881. This value suggests that including x10 in the optimal solution contributes significantly to minimizing the objective function.

The fmin value of approximately 42.059 represents the minimum value of the objective function achieved by the optimal diet plan.

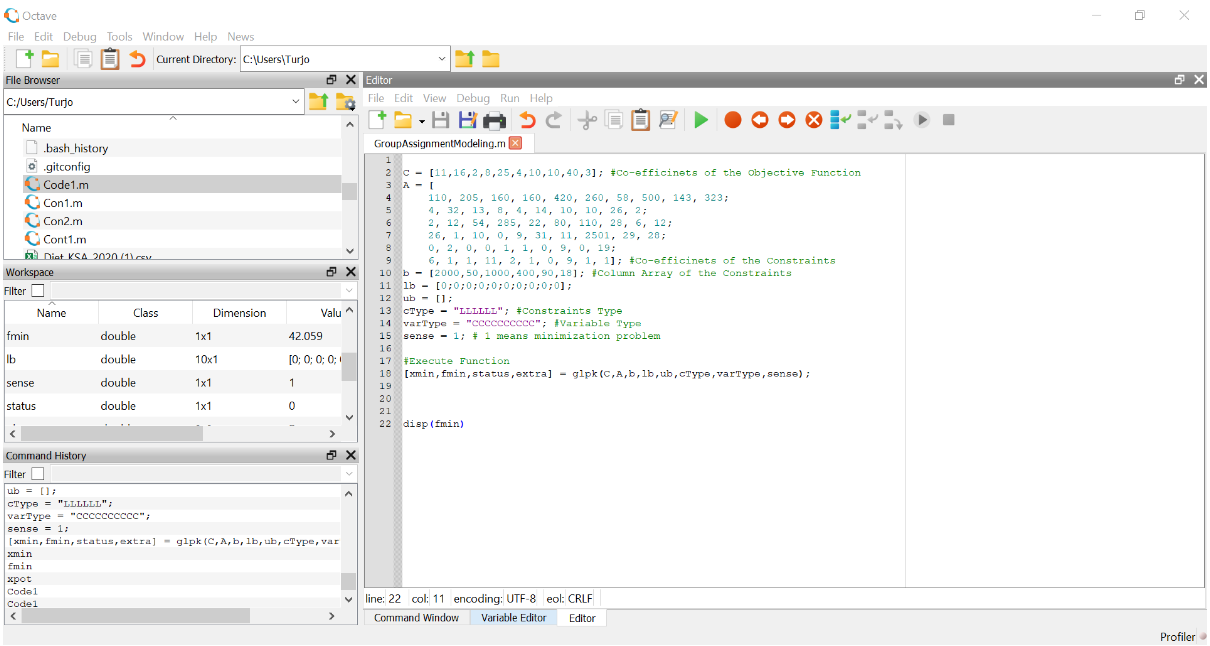

Figure 2: Implementation of the Modeling

9. Conclusion

In conclusion, the Simplex Method has demonstrated its effectiveness in optimizing dietary planning by finding the optimal combination of food items that satisfy nutritional requirements while considering various constraints. By leveraging this mathematical tool in conjunction with expert guidance, individuals can embark on well-informed and personalized dietary journeys that promote their health, performance, and overall well-being.

Moreover, the Simplex Method offers the advantage of flexibility and adaptability. As dietary needs or constraints change, the linear programming model can be easily adjusted to reflect these modifications. This dynamic approach enables continuous optimization and ensures that dietary plans remain relevant and effective over time.