UNIVERSITÀ DEGLI STUDI DI PERUGIA
Department Of Engineering

Bachelor's Degree in

Engineering Management

Bachelor's Thesis in

The Diet Problem: An Optimization Approach

Advisor
Prof. Giuseppe Saccomandi

Author
Alessia Cecchini

Academic Year 2024-2025

The Diet Problem: An Optimization Approach

Alessia Cecchini

2025

Contents

Abstract

The Great Depression, experienced by the U.S. in the middle of the I and II World Wars, was a period during which everyone paid attention to their personal expenses, trying to minimize costs. Among them, also the U.S. Army was interested in finding a solution to the problem of feeding the soldiers in the proper way – that is, having them in optimal shape and forces while minimizing the diet costs.
At that time, two people were working simultaneously, even if not aware of that, on what will later be known as the diet problem, one of the first optimization problems in the economic and social field.
After few years, to the diet problem was associated the name of George Stigler[1]. He started by collecting data from the National Research Council on the daily intake required for a "moderately active man", in terms of calories, proteins, vitamins and other essential substances. Then, from the Bureau of Labor Statistics, he analyzes the list of commodities, presented in terms of costs and nutritive values, so to select only those foods able to provide the optimum nutritive substances at minimum cost. Stigler, with a completely heuristic method, selected few of these foods (as wheat flour, cabbage, dried navy beans. . .) that ensure a healthy diet for only 39.93$ per year (referred to 1939). Stigler solution, as he himself acknowledged, cannot be effectively adopted: there is a lack of palatability!
Another important aspect of which Stigler was aware is that, especially at the time, does not exist a function describing health trends, and it is usually something "ad personam", specific for each individual needs. Moreover, he highlighted the fact that the same type of food can provide different nutritive values, depending on a lot of factors: temperature, maturity, way it is cooked, time of conservation exc. Also, for a single food, there exists a lot of varieties whose content of fundamental nutrients changes. In terms of limitations of this solution, can also be identified the difference in terms of eaten food's main nutrients and those that are actually absorbed and used by the human body.[1]

The other who was working on this problem was Jerome Cornfield[2], whose working

[1]Economist (Washington, 1911 – Chicago, 1991), professor at Minnesota University, Brown University, Columbia University, Chicago University. His research was focused on industrial organization; for this he received the Nobel Prize in 1982. (Treccani)

[2]Stathistic (New York, 1912 – Great Falls, 1979), professor and director of the Biostatistic Center at

activity was stranded until his friend George Dantzig[3], the father of Linear Programming and creator of the Simplex Algorithm, proposed to test is invention on the diet problem. The resolution was managed by Laderman[4], who modeled the problem as 9 linear equations in 77 unknowns; so, the simplex problem was divided between nine clerks who, after 120 man working days, were able to find the optimal solution: 39.69$ (only 0.24$ less than the Stigler's solution!).
Over the years, the diet problem was revisited multiple times for multiple needs. Dantzig himself decided to implement a revisited diet problem, with the intention to test it on himself, based on the selection criteria of those foods that ensure the feeling of feeling full; the solutions (obtained by iterating the problem multiple times due to unfeasible solutions proposed, like drinking 1892L of vinegar per day) were not implemented. [2]

From the Stigler's solution a lot of updated and revisited case studies have been proposed, according to the needs of modern society. Indeed, some criteria slightly different from those of minimizing cost subsistence, such as reducing environmental impact, ensuring a great palatability, considering a specific world region - and consequently, a specific diet, or a specific category of people - have been added as constraints to better model reality. For the same purpose, also changes in the mathematical tools have been introduced: linear programming was joined by quadratic programming, integer programming, multi-objective nonlinear programming and even statistics. *We'll dive into them soon.*

Washington University. His studies provided a great contribution to the statistical inference and biostatistics. (Oxford Academy)

[3]Mathematician (Portland, 1914 – Palo Alto, 2005), professor at Berkeley and Stanford Universities. He's considered the father of the linear programming and developed the famous Simplex algorithm. (Treccani)

[4]Colleague of George Dantzig at the Mathematical Tables Project of the National Bureau of Standards

Chapter 1

Linear Programming

1.1 Stigler and Dantzig solution

The diet problem has been under the attention of the main American researchers in the economic, mathematical or statistical field during the Great Depression of the 1930s-1940s. As a problem born in the U.S. Army, the "diet problem" refers to the determination of a set of foods that can ensure the right nutrient intake while minimizing total cost.
George Stigler, one of the first approaching the diet problem, looked out for a solution even if he did not have a rigorous mathematically-based method that could have helped him. He started by collecting data from the National Research Council (U.S.) about the recommended daily intake for a "moderately active man" in terms of calories, proteins, vitamins and other essential substances (Table 1.1). The list of foods, from the Bureau of Labor Statistics, as Stigler himself recognized, is pretty short and not including fresh fruit, vegetables and fish – whose presence could have substantially reduced the diet's annual cost. In developing his study, Stigler looked forward the optimal solution for two years: 1939 and 1944 – also to show the impact of the World War II.
Before showing the result proposed by Stigler, is interesting to notice the limitations that such adopted model has – and that the economist himself has recognized. Indeed, the lack of accuracy both regards constraints and food's information. While Stigler only takes into account nine nutrients' daily allowance, the human body requires many more – even if in small quantities. The information related to the considered foods is rough and highly subjected to variations, that depends on temperature, maturity, way of cooking, variety of the product (e.g., an apple's amount of ascorbic acid depends on the apple variety: Jonathan, McIntosh, Ontario…), time of storage and lot of other factors that can't be monitored. Taking into account those aspects, let's analyze the solution proposed by Stigler.
Through a completely heuristic method, Stigler identified a solution for both the years under analysis – 1939 and 1944 – with the first being only 24 cents far from the optimum obtained, few years later, applying the new Simplex algorithm. The solutions are obtained starting

Table 1.1: Daily Nutrient Allowances for a Moderately Active Man

Nutrient	Allowance
Calories	3000 calories
Protein	70 grams
Calcium	0.8 grams
Iron	12 milligrams
Vitamin A	5000 International Units
Thiamine B_1	1.8 milligrams
Riboflavin B_2	2.7 milligrams
Niacin	18 milligrams
Vitamin C	75 milligrams

from the 77 foods, eliminating at first those that do not provide a great contribution in nutrients with respect to others – reducing the number to 15 alternatives. At this point, Stigler analyzed the possible linear combinations (not all of the 510) proposing a solution with 5 commodities (Figure 1.1).

Minimum Cost Annual Diets, August 1939 and 1944

Commodity	August 1939		August 1944	
	Quantity	Cost	Quantity	Cost
Wheat Flour	370 lb.	$13.33	535 lb.	$34.53
Evaporated Milk	57 cans	3.84	—	—
Cabbage	111 lb.	4.11	107 lb.	5.23
Spinach	23 lb.	1.85	13 lb.	1.56
Dried Navy Beans	285 lb.	16.80	—	—
Pancake Flour	—	—	134 lb.	13.08
Pork Liver	—	—	25 lb.	5.48
Total Cost		$39.93		$59.88

Figure 1.1: Stigler solutions considering data for 1939 and 1944

It is important to notice that, as Stigler himself highlighted, the solution proposed is not necessarily the optimum among the combinations; however, such optimum could differ of a non-relevant amount. The commodities of the 1939 solution are: wheat flour, evaporated milk, cabbage, spinach, dried navy beans – the solution referred to 1944, obtained with the same methodology, replace evaporated milk and dried navy beans with pancake flour and pork liver – with an annual cost of 39.93$ (1939), 59.88$ (1944).
In both cases, such a solution is a mathematical optimum, compliant with nutritional requirements, but cannot be adopted by anyone for multiple reasons – starting from palatability.
Later on, George Dantzig introduced the Simplex Algorithm (A.1), then used by Jack Laderman in 1947 to solve Stigler's problem – after 120 working days, with 9 people cooperating

– and his solution revealed the almost correctness of Stigler's one. Indeed, the difference between the two optimal solutions is of only 0.24$.

1.2 Modern Diet Problem

From the Stigler and Dantzig solutions' lot has changed. Indeed, while the diet problem represents still today a challenge from multiple points, it has been implemented with the knowledge acquired over the years about nutritional requirements and cost of foods. Moreover, it has been adapted to different requirements and scopes – such as, focusing on specific region's diet, on the environmental impact or to prevent risk of diseases.

Some studies have started by rearranging the Stigler problem, tailoring the solution to the period during which the study was made. In the revisited Stigler problem [3], Garille and Gass leveraged the greater amount of available information about lifestyle, nutritional food contents and foods' cost data. In fact, they introduced foods according to how they are usually eaten (that is, instead of "raw eggs", eggs are introduced in the different ways they can be cooked) and took the costs' information from two sources - one of which is the 17th version of the Stigler's source - matching them to pick the cheapest cost for same alternative. Moreover, the nutritional recommendations were considered both from those used by Stigler and by taking those of 1998 - the year when article was written.

In problem setting, two distinctions were made: gender and constraint's type - that is, solutions were proposed separately for male and female, and allowing or not the excess of upper bound for nutrients intake.

The solutions are way far from the Stigler's optimum - mainly due to the money's value change over the years. Indeed, allowing going above upper-bounds, for a 25-50 years old man an annual diet cost of 412.26$, for a 25-50 years old woman the diet cost is of 354.05$. Also, differently than the 1939 solution, the maintained food is only wheat flour, together with carrot, orange and milk. Instead, if no excess is allowed - so the constraints are equalities - for male is 522.30$ and 461.72$ for females, with a diet plan characterized of wider choice (e.g. bacon, roasting chicken, lard, peanut butter) even if some nutritional minimum requirements are not satisfied (for carbohydrates, dietary fiber, zinc, copper, iodine etc.). In another proposed version the updated nutritional values (that is the Recommended Daily Allowances of 1998s) and recommended intake were considered, but 1939 costs. The solution was of 40.92$ with respect to 1939, that become 481.16$ in 1998. This shows that, even if the daily allowances recommendations have been updated, including also other nutrients, the difference from the Stigler's solution is not too far.

The habits related to nutrition represent part of human culture, and indeed they vary across multiple regions of the world according to population's religion, to its needs depending on the area where they live, and to the products offered by the internal activities as agriculture and breeding - just as examples. Moreover, there are multiple scientific studies proving how an health nutrition can have impact on individual's health, both physical and psychical, but also on environment and economics. Thus, with the objective to preserve individuals' diet plan, a lot of studies arise, inspired by Stigler, with the goal to define a nutritional satisfactory plan maintaining habits.

Among these, Maillot *et al.* conducted a study focused on eating habits of French population and on the necessary changes to achieve both the Estimated Average Requirements (EAR) and Recommended Daily Allowances (RDA) [4], [5]. To consider possible variations from the usually consumed foods, Malliot *et al.* takes data from a survey conducted by INCA in 1999, involving 1171 French adults. Since the objective was to modify the personal diet plan of each individual, the starting input data were those declared by each adult participating in the survey, excluding all beverages, and grouped at different levels (starting from 7 main food categories, and going into detail to 41 sub-categories).

The constraints were applied at four levels: energy, weight, nutritional, social acceptability.

1. **energy**: the new plan's energy was set equal to the previous diet, not having any insights about each adult's lifestyle.
2. **weight**: total diet weight was imposed a limit of 115% with respect to current diet, to ensure the possibility to perform restrictions on the energy per weight of food, as some dietary guidelines suggest.
3. **nutritional**: macronutrients' intake recommendation were tailored according to EAR and RDA, while the others were adapted to individual parameters, as sex and gender.
4. **social acceptability**: the 95th percentile of the consumer intake distribution has been calculated for each food and considered as an upper-bound.

The objective function was designed to limit as much as possible the introduction of new foods into personal diet plan, to overcome the resistance to change that may arise when it is about dietary habits, and, of course, giving preference to those foods consumed by French population. That is, whenever an increase in a specific nutrient is required, the model suggests a greater consume of already eaten foods containing such a nutrient, rather than introduce a food not habitually eaten by the individual – or that the individual may have difficulties to insert into the diet plan, as foods that do not belong to the French characteristic foods.

The linear problem is modeled as:

$$min\ F_i = \sum_{j \in \mathbb{R}} |D_{ij}| \ + \sum_{j \notin \mathbb{R}} w_j Q_j^{opt}$$

where the first term refers to repertoire foods, the second to additional ones. Specifically:

$|D_{ij}|$ express the negative relative deviation between the optimized and the observed quantity of repertoire's food j for individual i;

Q_j^{opt} is the additional quantity of food j in the optimized diet of individual i;

$w_j = (1 + \frac{1}{NC_j})/Q_j^{med}$ are the weights to preferentially select the most consumed foods among additional foods;

Q_j^{med} is the median quantity of food j observed in the sample;

NC_j is the number of individuals who consumed food j.

The modeled diet plans were so compared with the original ones, evaluating the difference between the two through median, first and third quantile. Interesting outcomes can be noticed through the box-plot representation:

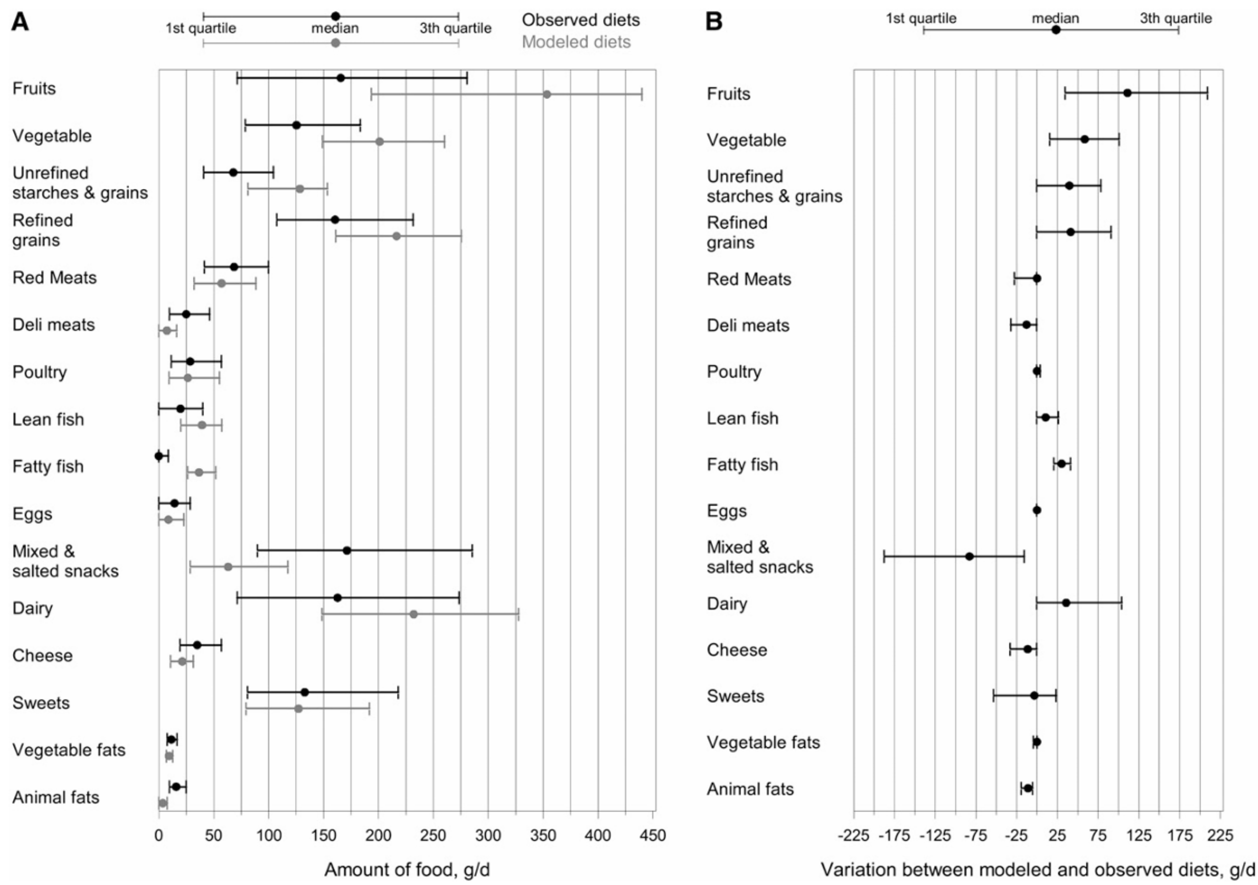


Figure 1.2: The A part compares observed with modified diets; the B part shows the variation between observed and modeled diets

First of all, as shown in the figure 1.2 an increase in fruit, vegetables, grains, fresh dairy products, fatty fish, walnuts was observed, while, on the other hand, a strong diminishing

of sweets, cheese, red meat and deli meat was pointed out.
Then, what makes this research unique and distinct from others dealing with the same topic, is the ability to convert nutritional recommendations into foods and, especially, to provide dietary plans adjusted to each individual needs - and for such a reason the data were taken based on individual and population preferences. Indeed, the objective function was modeled with the goal to minimize variations from the current individual diet plan. For this, only 5 foods not included in the already consumed ones were added ($< 5\%$ contribution to the diet weight). This aspect implicitly highlights the importance of some foods that were inserted in a lot of nutritional plans. That is, the only way to naturally assume fundamental nutrients is by consuming these added foods - as fatty fish and walnuts.
Another interesting conclusion is related to the amount of different foods, about 35-40, that are necessary and sufficient to constitute a nutritionally satisfactory diet plan. Comparing this conclusion with the Stigler's one, is evident how a price constraint may affect not only the food choice, but also its quantity.

The negative effects that cost constraints can have on the choice of foods have been evaluated by a study conducted at the beginning of 2000s by Darmon, Ferguson and Briend. [6]
The data for the study are taken from a survey made in the district of Val de Marne, Paris. Those data were used to define the characteristic diet pattern of French people living in such area, and for implementing the LP model (only with respect to those data related to adults - both male and female). The survey refers to 73 foods and 28 nutrients.
Moreover, as Maillot *et al.* did, the problem was modeled with the goal to minimize departure from the current diet plan, and assess its changes as cost constraints were gradually strengthened.
The objective function was at first defined in a non-linear way - expressing the total departure from the mean food intake:

$$minimize \quad \sum_{i=1}^{n} \frac{|m_i \ - X_i|}{m_i} \qquad \forall i = 1, ...n$$

being m_i the mean portion size observed in population for food i; X_i the food variable portion size for food i.
The objective function was so linearized by rewriting the absolute value as:

$$Z_i \geq \frac{(m_i \ - X_i)}{m_i}, \ Z_i \geq -\frac{(m_i - X_i)}{m_i} \quad \forall i$$

So that the objective function becomes:

$$minimize \quad \sum_{i=1}^{n} Z_i = \frac{m_1 - X_1}{m_1} \ + \ \frac{m_2 - X_2}{m_2} \ + \ ... \ + \frac{m_n - X_n}{m_n}$$

That is equivalent to:

$$minimize \quad \sum_{i=1}^{n} Z_i = n - \frac{X_1}{m_1} + ... + \frac{X_n}{m_n}$$

The constraints were introduced at multiple levels: energetic, nutritional and, the core of the study, economic. The energy of the proposed diet plans was chosen based on the average daily energy intakes of the population, distinguished by gender.
The food constraints were introduced with the purpose of proposing adjusted diets aligned with the typically eaten foods, that is, by lowering the variance. To achieve this goal, an upper bound correspondent to the 75^{th} percentile of the population's daily portions (in g/d) for each considered food is set. Moreover, the typical French diet was preserved by setting constraints on the energy contributed by each of the six food group and of the 21 subgroup. The energy intake was so limited to the range between $10^{th} - 90^{th}$ percentile and to a range $25^{th} - 75^{th}$ for each subgroup, distinguish for gender. For the same reason, those foods consumed by less than $< 10\%$ were excluded, as well as beverages of any kind.
Lastly, cost constraints were gradually improved of 50 cents. at a time.
The results show how cost constraints impact diet cost, energy intake and nutritional values. First of all, the most expensive foods in a diet are meat, fish, and eggs, corresponding to 44% and 41% of the diet cost for males and females, respectively. Then, a great contribution is provided by fruit and vegetables, followed by dairy and cereals products, concluding with added fats and sweets ($< 3\%$ of total diet cost). So, as the cost constraint is applied, an evident decrease in the meat, fish, eggs, fruits and vegetables is noticed. Looking at the minimum daily diet cost (2.52\$ for men, 1.78\$ for women), the contribution of the most expensive food groups is highly reduced, with a slight increase in cereals.
The cost constraint has consequences also on the energy level. Indeed, based on what discussed above, seems quite intuitive to get that as cost constraint is strengthened, the main energy sources will be cereals, sweets and added fats foods - the cheapest ones. Further going into details, also the subgroups choice changes as a consequence of the limitations imposed - as example, the fresh fruit intake decreases in favor of dried fruit and nuts.
Related to this, from a nutritional point of view, the decrease in protein supply, associated to meat and fish, is compensated by carbohydrates and fats, that become the main energy sources. However, differently than proteins that, even if reduced, remain above the Population Reference Intake index, some micronutrients (as iron, selenium, vitamin C, calcium, magnesium, zinc) go below that value as the diet cost gets $\leq 3.0\$/d$ for men and $\leq 2.5\$/d$ for women. The micronutrients deficiency can have negative effects on health. ([7], [8], [9])
Thus, the negative effects that a diet cost restriction could have on health, especially if kept for a long time, are evident from Darmon's study. For this reason, an economic sustainment to those belonging to low socioeconomic groups, accompanied with proper nutritional education, may have positive effects on people's health.

The question, at this point, is: what if cost constraints are combined with other factors?
The answer was provided by van Dooren *et al.* [10] who tried to combine a cost constraint with the environmental impact of the diet plans for Dutch people (31 to 50 years old, both male and female). This imply a definition of a sustainable diet: "Sustainable diets are those diets with low environmental impacts which contribute to food and nutrition security and to healthy life for present and future generations. Sustainable diets are protective and respectful of biodiversity and ecosystems, culturally acceptable, accessible, economically fair and affordable; nutritionally adequate, safe and healthy; while optimizing natural and human resources" (Food and Agriculture Organization of the United Nations, FAO). [11] To introduce and assess such type of constraint, the parameters included in the study are: greenhouse gas emissions (GHGE), land use (LU), fossil energy use, chosen based on the Life Cycle Assessment (LCA) - a methodology used to quantify the products impact on environmental aspects, as water use and land use. [12]
The nutritional data are taken from Dutch Nutrient Database and Nutrition Centre, while data related to diet costs come from Dutch National Institute for Family Finance Information.
The objective function was designed with the intent of giving more credit to those foods usually eaten by the population, penalizing the introduction of non popular ones, like what was made by Maillot [5]. The Food Consumption Survey is used as reference to get which are popular foods.
The constraints were arranged according to the following criteria:

- The nutrient constraints are aligned with the recommendations used for 33 macro and micronutrients from the Netherlands Nutrition Center and the Health Council of the Netherlands.

- The environmental restrictions were introduced to half the GHGE average daily impact, expressed in CO_2eq/day (so, is set equal to 1.6 kg CO_2eq/d, which would be needed to achieve the European goal of keeping global warming below 2° within 2050 - information referred to 2015, when the paper was written). The GHGE evaluation for each product is made based on the import vs export and, for vegetables mainly, on open field vs heated greenhouse production.

- The economic limitation, that is the maximum diet cost, is set equal to 2.50€ (50% reduction of reference costs and close enough to the 2€/day poverty line, according to World Bank 2015).

To underline the different impacts that the different constraints have on the diet plan definition, they are applied one after the other.
At a first stage, nutritional limitations are imposed, resulting in a model that corrects for

possible energy, fat, salt and carbohydrate overconsumption, with a diet plan characterized by fruit, vegetables and legumes rather than on meat or dairy products.
The second run introduces the GHGE restriction, which led to 55% and 43% reduction for males and females, respectively.
The economic constraint move the diet cost to the minimum of 2.59 €$/d$, a significant reduction if compared with the results obtained after both nutritional and environmental constraints were introduced (3.23 €$/d$ for men and 3.18 €$/d$ for women), but not sufficient to reach the imposed requirement (2.50 €).
The results point out how sustainability limitations almost automatically generate a low-price diet plan. Indeed, as proven in the previous study [6], the more expensive foods are meat and dairy products, that are also responsible for negative environmental effects. [13] Indeed, even if the objective function was designed to minimize substitution of typically eaten foods, the meat and dairy products are replaced by soy products.

Similar analysis has been proposed by Liu *et al.* in their recent article (published in 2024), which encompasses 181 countries around the world.[14] Starting by assessing the nutritional, economical and environmental affordability, so to optimize at different levels.
The food data were taken from FAO 2020 database, with a total of 74 food items classified into 10 main categories (such as eggs, meat, legumes, fruit etc.). The associated price of those foods is taken from the U.S. Bureau of Labor Statistics.
Since on a first stage the purpose was to assess nutritional level of the countries, two parameters have been introduced:

$$N_{i,k} = \sum_{j=1}^{10} Q_{j,k} \times n_{i,j}$$

is the average daily intake of nutrient i in $gr/capita * day$ for the population of the k-th country, $Q_{j,k}$ is the daily consumed amount of food j by people of country k, $n_{i,j}$ is the amount of nutrient i in food j.

$$NAR_k = \frac{1}{24} \times \sum_{i=1}^{24} \frac{N_{i,k}}{RDA_i} \times 100\%$$

is the nutrient adequacy ratio for country k, and RDA_i is the recommended daily allowance for food i. The NAR was evaluated for the considered countries.
The resulting measures show interesting results. First of all, only 14.4% of the countries have a NAR above 80%. Interesting to notice, that among these the majority are European countries (as Greece, Spain, Portugal, Italy etc.). Indeed, compared to other continents, Europe has a NAR median much higher than the others (confirming the numerous articles that suggest Mediterranean diet as the optimal under wide types of health disease [15], [16]). Instead, Madagascar, Yemen, Afghanistan, Haiti, Lesotho, Guinea-Bissau have the lowest registered NAR value.

These differences in nutritional adequacy are influenced by dietary habits, food availability and affordability, consequently related to the country income level (greater income level, greater affordability of food): the low-income countries show a median NAR of 57.6%, the high-income ones of 77.8%.

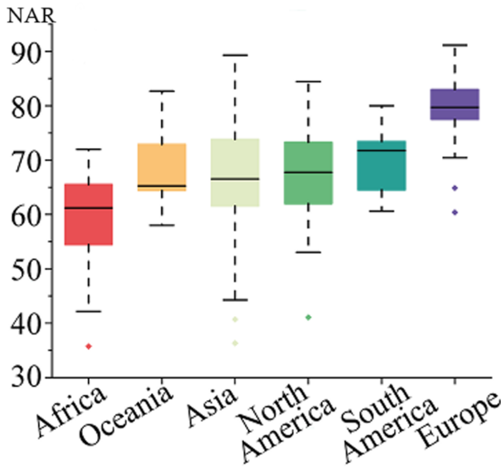


Figure 1.3: Box plot representing the NAR value in the different continents.

Regarding the RDA, only six over the 24 essential nutrients considered satisfy the values. Also in this case, not all countries can satisfy the same nutrients - for example, carbohydrates RDA is met by 175 countries, while only 58 met the protein recommended intake.
Introducing the optimization model formulation, the objective function is defined as:

$$minimize\ f = \frac{1}{n}\sum_{p=1}^{n}\left|\frac{N_{opt_p} - RDA_p}{RDA_p}\right|$$

$$N_{opt_p} = \sum_{j=1}^{10} Q_{opt_j} \times n_{p,j}$$

where f is the function expressing the distance between the nutrient levels and the recommended dietary allowances in the current diet, n is the number of foods not meeting the adequacy standards, N_{opt_p} is the optimal quantity of nutrient p that does not meet the recommended dietary allowance under the current diet structure.
Then, the nutritional, economical, environmental and other constraints are introduced:

- **environmental parameters**:

$$LU_k = \sum_{j=1}^{10} Q_{j,k} \times l_j$$

$$FW_k = \sum_{j=1}^{10} Q_{j,k} \times w_j$$

$$GHG_k = \sum_{j=1}^{10} Q_{j,k} \times g_j$$

with LU_k, FW_k, GHG_k representing, respectively, the land use m^2, freshwater withdrawal L and greenhouse gas emissions $kg\ CO_2eq$;

- **economic constraint**:

$$\sum_{j=1}^{10}(Q_{opt_j} \times Price_j) = \sum_{j=1}^{10}(Q_j \times Price_j)$$

to ensure that the optimized diet plan has the same cost as the current one;

- **acceptability constraint**:

$$\left|\frac{(Q_{opt_j} - Q_j)}{Q_j}\right| \leq 0.3$$

limiting the diet structure deviation from current to 30%;

- N_{opt_p} **not decreasing**

$$N_{opt_q} \geq RDA_q$$

to avoid that the existing nutrient levels did not decrease, the nutrients that met the RDA in the current diet structure have to be greater than that RDA;

- **non negativity constraint**:

$$Q_{opt_j} \geq 0$$

- **nutritional upper limit**:

$$N_{opt_i} \leq UL_i$$

So, the research proceeds by optimizing the diet structure under acceptability and affordability constraints, with a NAR increase ranging between 1.8% to 11.9%. Then, as environmental constraints are applied, the optimization solution suggested a greater mitigation in middle- and high-income countries.
In the end, optimized cases were varied, proposing three different alternatives: environmental impact and nutrition are optimized together; relax the acceptability constraint; relax the affordability constraint. The first new scenario suggests a decrease in meat consumption, and an associated slight increase in vegetables; moreover it increased the NAR of $0.4\% - 6.8\%$ from current diet structure, so less the increased observed when only nutritional constraints are applied. The second case led to a NAR median increase to 90%; however, it is worth to notice that, differently than the first scenario that generates diet structures potentially adoptable, this second scenario is not so acceptable due to factors as habits, culture and food availability. In the third case, the NAR median moves to 77%, with greater amount of meat and vegetables in the diet.

1.3 Children Health and Diet

Frequent applications of the diet problem are referred to those world regions that are sadly known for their difficulties in satisfying the primary needs, with significant consequences on children. Focusing on this, Darmon *et al.*, took as reference the Malawian children (3 to 6 years old), to design a nutritional satisfying diet by taking into account the differences in food availability between harvest and non-harvest season. [17]
The LP method was used to model the problem under both nutritional and food consumption constraints - so to ensure, respectively, the achievement of nutritional recommendations (considering the Reference Nutrient Intakes defined by United Kingdom) and palatability of foods group combinations. All the considered foods were taken from a 2-season survey performed in Malawi, and thus including only locally available foods.
The linear programming model was defined, by performing also mathematical transformations whenever the nutritional requirements were imposed as ratios, so to obtain linear functions to be optimized.
Specifically, the objective function was written as:

$$minimize\ a_1W_1\ + a_2W_2\ +\ ...\ + a_nW_n$$

where the $a_1, a_2,\ ...,\ a_n$ coefficients express the energy content per unit weight for the correspondent foods $F_1,\ F_2,\ ...,\ F_n$, and $W_1,\ W_2,\ ..., W_n$ are the weights of each food.
The requirements have been applied through multiple steps, starting by introducing only the nutritional constraints, inserting then an energy proportion from multiple foods to 25th and 75th percentiles, and taking into account that constraints related to vegetables were stronger during the non-harvest season. Lastly, also limitations on the gr/day amount of eaten food were imposed.
The constraints for minimal nutrient intakes are $N_i\ \geq RNI_i$ for i^{th}nutrient, where N_i is equivalent to the total amount of each nutrient provided by the diet, and RNI_i equals the selected nutrient recommendation for that nutrient.

Modifying the constraints, strengthening them, led to two main general conclusions. First of all, changing the range of percentile values for food group, eligible foods and maximum weight of food, is actually improving the minimum energy required to satisfy the nutritional constraints. The second relevant issue addressed is quite intuitive: the minimum energy required to satisfy the requirements, considering the same foods group, is higher during non-harvest season. This shows the difference in nutrient-rich foods' availability between harvest and non-harvest season, together with the different percentage of energy extracted by same foods in different periods.
In conclusion, such an analysis, point out the importance of vegetables and foods of animal origin to increase the nutrient density of diet. Moreover, the results suggested that the in-

adequacy of diet persists during the non-harvest season due to the low availability of zinc and riboflavin rich foods, requiring an intervention on available foods during that period.

1.4 Computer Implementation

An unquestionable advantage of Linear Programming is its easy implementation on computer support.
Among various examples, let me illustrate a brand new one, published at the beginning of this year, that implements with Gurobi's support, multiple diet plans according to the pursue that has to be reached. Specifically, it addresses diet plans each characterized by vitamin D deficit, iron deficit, calorie deficit, high protein diet, maxed-out diet (that is, the maximum possible intake of all nutrients is reached, maintaining the cost constraint as a secondary limit), with the aim to minimize the total cost under specific constraints - setting both lower and upper bounds. Differently than the studies previously shown, in this case more realistic-oriented decisions have been performed: it is assumed that no ingredient can exceed the 20% of the total meal; the weights are considered as continuous positive variables. The linear model is so presented in the form:

$$\begin{aligned}
&minimize \sum_{i \,\in\, ingredients} W_i P_i \;\; \forall i \in ingredients \\
&such\ that \sum_{i \,\in\, ingredients} W_i N_{i,k} \geq min_k \;\; if\ k \in mini \\
&\sum_{i \,\in\, ingredients} W_i N_{i,k} \leq max_k \;\; if\ k \in maxi \\
&W_i \leq \frac{1}{5} \sum_i W_j \;\; \forall i \\
&W_i \geq 0 \; \forall i
\end{aligned}$$

were W_i are the weights of ingredient i, P_i the correspondent price per unit weight, $N_{i,k}$ the nutrient k in ingredient i. Comparing the results, is interesting to notice that the diet modeled with the purpose to recover vitamin D deficit is the one with the highest cost, compared to the others. Specifically, the daily expense associated to such diet is 12.3$ while the others have an average of 4.23$.

Chapter 2

Iron Deficiency and Sensitivity Analysis

The Linear Programming models offer the possibility to implement variations to assess the strength of the solution and of the constraints - that is, performing a sensitivity analysis.
It will be implemented in the following case study, that focuses on iron daily intake, differentiating by food sources.

Among the various micro-nutrients and minerals, iron is by far one of those covering an essential role in human body. Indeed, it binds with hemoglobin (in red blood cells) and myoglobin (in muscles) ensuring oxygen transportation through all tissues and organs. There are other metabolic processes involving iron, as DNA synthesis and electron transport.
Despite is high bio-availability, iron deficiency is the most common nutritional deficiency in the world, causing anemia. A low intake of iron is of great significance specifically for women (from adolescence to menopause). However, due to a particular behavior of iron, it is important to highlight that iron overload may have negative consequences on tissues. ([18], [19])
There are two main forms of iron: heme and non-heme. The main sources of the first are animal proteins, from meat to fish, with an absorption ranging between $10-35\%$; differently, the main sources of the second type are vegetables, legumes, cereals with an absorption rate between $2-10\%$. [20]

The proposed case study wants to minimize the costs associated to the combination of three main heme iron sources (specifically, beef, mollusks and eggs), and of two non-heme iron sources (cocoa powder and lentils).
The foods cost, iron content, energy supply and other information that will later be used are shown in Table 2.1 and 2.2. At a first stage nutritional and energetic intake constraints will be applied for different case studies - distinguishing by gender and age (18-29 years old,

50-59 years old), assuming for all the same physical activity level index, that is $PAL = 1.6$ (moderately active). In all the addressed cases, the involved variables are assumed to be ≥ 0. Then, sensitivity analysis is performed, checking how the optimal solution changes varying the constraints.

Table 2.1: Heme Iron Sources considered

Food	Energy $kcal/100gr$	Iron $mg/100gr$	Cost $\$/100gr$	Environmental Impact $kg\,CO_2eq/100gr$
Beef, cooked, braised (x)	145	39.4	1.7	25
Mollusks, clam, mixed species, cooked (y)	202	13.9	2.3	3.5
Eggs whole cooked (z)	143	1.75	0.9	3.8

Table 2.2: Non-Heme Iron Sources considered

Food	Energy $kcal/100gr$	Iron $mg/100gr$	Cost $\$/100gr$
Cocoa powder, unsweetened (x)	228	13.9	2.5
Lentils, raw (y)	360	7.2	0.4

The nutritional data are taken from FoodData Central, USDA; the costs data from Bureau of Labor Statistics, US; the environmental impact data from Our World Data.

Before delving into the various cases, some aspects should be specified, starting by the food choice. The selection of beef, mollusks and eggs lies on few reasons. First of all, they belong to three different food groups (meat, fish, dairy products) and are available, even if in different quantities, through the entire world; thus, the analysis performed here can be of interest for multiple regions. Similar consideration has been made for non-heme iron sources. Moreover, the choice of specific foods is due to the consulted databases from which information about nutrient content, energy supply and cost have been taken. Also, the selection occurred between those foods that have an higher iron content (for example, among meat the choice was beef rather than poultry for such reason).
Regarding the values of recommended iron intake, the choice of inserting an upper bound is motivated by the diseases that iron overload may have on organs and tissues. [21]

2.1 Women Iron Intake

Women requirement of iron varies through their different stages of life. Indeed, starting from pre-adolescence (when the first menstruation occur) the body requests of iron increases, reaching the peak during pregnancy, and then moving down as menopause starts.
The focus of this analysis is on 18-29 years old and 50-59 years old, to compare the different requirements and solutions between pre and post menopausal women.

Table 2.3: Women Iron and Energy Daily Intake Recommendation

Age (years)	Iron (PRI, mg/day)	Energy (AR, $kcal/day$)
18-29	16-45	2150
50-59	11-45	2030

Data from https://multimedia.efsa.europa.eu/drvs/index.htm; PRI is the population reference intake; AR is the average requirement

2.1.1 Heme Iron Sources, 18-29y

As first case study, let's model the minimization LP problem, subjected to nutritional and energy intake constraints.

$$minimize \;\; f : 1.7x \; + \; 2.3y \; + \; 0.9z$$

$$s.t. \;\; 39.4x \; + \; 13.9y \; + \; 1.75z \geq 16$$

$$39.4x \; + \; 13.9y \; + \; 1.75z \leq 45$$

$$145x \; + \; 202y \; + \; 143z \geq 2042.5$$

$$145x \; + \; 202y \; + \; 143z \leq 2257.5$$

The iron supply is constrained to avoid anemia on one side and overload on the other side. The energy intake ranges around the average requirement ($AR = 2150 \;\; kcal/day$) of a 5% variation, to provide a more realistic situation.

The problem optimal solution is:

$$f^* = 12.85\$, \quad (x, y, z) = (0, 0, 14.28)$$

So, it suggests a daily intake of about $1.4kg$ of eggs, which is quite unrealistic. To fix this, an upper bound can be introduced. For example, it is possible to modify the model to get a more realistic solution by setting upper bounds in the consumption of the considered foods. So, limiting the eggs content in the diet plan will consequently require the introduction of

the other foods, to satisfy the constraints. Indeed, for example, introducing as upperbound $z \leq 12$ the solution would include also mollusks ($(x, y, z) = (0, 1.616, 12)$). Thus, the diet variety is increased a little, even if the solution still appears unrealistic (suggesting to consume $1.2kg$ of eggs per day). Further implementation of this is left as suggestion for future insights.
In addition, the diet cost is pretty high if compared with those used as benchmarks by Darmon. [6]
It is interesting to notice that even by removing the upper bound constraint referred to iron, nothing changes.
A further analysis may be performed restricting the diet with respect to sustainability (expressed as the amount of $kg\ CO_2eq/day$), imposing the total amount being $\leq 1.6\ \ kg\ CO_2eq/day$ (value chosen to guarantee that the global temperature increase remains below $2°$ within 2050, [10]). However, due to the strength of the constraint, the final solution is infeasible.

Let's now start performing sensitivity analysis, to check how much the cost of beef and mollusks has to decrease in order to insert them into the diet plan.

Change in Beef Cost

Modify the $\mathbf{c}^T$ vector, $c_x \leftarrow c_x\ +\ \Delta$; since the x variable is non-basic, the mathematical procedure is shown in section A.2.2.
Look at the final optimal simplex tableau of the previous iteration:

Basic	x	y	z	S_1	S_2	S_3	S_4	Solution
Z	$-\frac{563}{715}$	$-\frac{1471}{1430}$	0	0	0	0	$-\frac{9}{1430}$	$\frac{7353}{572}$
z	$\frac{145}{143}$	$\frac{202}{143}$	1	0	0	0	$-\frac{1}{143}$	$\frac{4085}{286}$
S_2	$\frac{107609}{2860}$	$\frac{8171}{715}$	0	0	1	0	$\frac{7}{572}$	$\frac{22885}{1144}$
S_3	0	0	0	0	0	1	1	215
S_1	$-\frac{107609}{2860}$	$-\frac{8171}{715}$	0	1	0	0	$-\frac{7}{572}$	$\frac{10291}{1144}$

Figure 2.1: Final Tableau of Simplex Algorithm, only nutritional and energetic constraints applied (women, 18-29y)

In this specific case, the objective is to lower c_x so to include beef, that is, the objective function optimal has to be modified. Thus, looking at Figure 2.1, $\bar{c}_x = -\frac{563}{715}$, from which follows that to maintain valid the optimality condition and get a new optimal for f:

$$\Delta \leq -\frac{563}{715}$$

$$-\infty \leq \bar{c}_x^{new} \leq 1.7 + \Delta$$

The objective function is reformulated with this new value:

$$minimize \;\; f: 0.91x \; + \; 2.3y \; + \; 0.9z$$

$$s.t. \;\; 39.4x \; + \; 13.9y \; + \; 1.75z \geq 16$$

$$39.4x \; + \; 13.9y \; + \; 1.75z \leq 45$$

$$145x \; + \; 202y \; + \; 143z \geq 2042.5$$

$$145x \; + \; 202y \; + \; 143z \leq 2257.5$$

Getting:

$$f^* = 12.85\$, \quad (x, \; y, \; z) = (0.53, \; 0, \; 13.74)$$

The objective function optimal value is the same, but obtained combining two variables (that is, two heme iron sources).

Change in Mollusks Cost

Following the same procedure for the y variable (referred to mollusks), with a $\overline{c}_y = -\dfrac{1471}{1430}$ (Figure 2.1). The Δ will be:

$$\Delta \leq -\frac{1471}{1430}$$

and so

$$-\infty \leq \overline{c}_y^{new} \leq 2.3 \; + \; \Delta$$

Rewriting the problem for $\overline{c}_y^{new} = 2.3 \; - \; \dfrac{1471}{1430}$:

$$minimize \;\; f: 1.7x \; + \; 1.27y \; + \; 0.9z$$

$$s.t. \;\; 39.4x \; + \; 13.9y \; + \; 1.75z \geq 16$$

$$39.4x \; + \; 13.9y \; + \; 1.75z \leq 45$$

$$145x \; + \; 202y \; + \; 143z \geq 2042.5$$

$$145x \; + \; 202y \; + \; 143z \leq 2257.5$$

That provides:

$$f^* = 12.85\$, \quad (x, \; y, \; z) = (0, \; 1.75, \; 11.81)$$

Thus, a change in the cost both for beef and mollusks, separately, lead to the introduction of these foods into the optimal combination for the diet plan, maintaining fixed the optimal function value.

Change in Absorption Level

However, in this first problem formulation an important data has not been included: the absorption level of iron. In fact, as specified above, the heme iron sources have an absorption rate between $10-35\%$, which means that not the entire amount of iron in a portion ($100gr$) can be assimilated by the organism.
For example, let's suppose that the absorption level is of 30%, 25%, 20% respectively for (x, y, z), that is beef, mollusks and eggs. Then, the final solution will be

$$f^* = 13.61, \;\; x = 0.95, \;\; y = 0, \;\; z = 13.31$$

Beef is included in the plan and the total diet cost rise up.
This case can be assumed as an example of sensitivity analysis, where the a_{ij} parameters are modified.

2.1.2 Heme Iron Sources, 50-59y

Moving the attention to another age interval, specifically for women referred to the menopausal period, the constraints of the model are modified as

$$minimize \;\; f: 1.7x \; + \; 2.3y \; + \; 0.9z$$

$$s.t. \;\;\; 39.4x \; + \; 13.9y \; + \; 1.75z \geq 11$$

$$39.4x \; + \; 13.9y \; + \; 1.75z \leq 45$$

$$145x \; + \; 202y \; + \; 143z \geq 1928.5$$

$$145x \; + \; 202y \; + \; 143z \leq 2131.5$$

With the final simplex tableau being:

Basic	x	y	z	S_1	S_2	S_3	S_4	Solutior
Z	$-\frac{563}{715}$	$-\frac{1471}{1430}$	0	0	0	0	$-\frac{9}{1430}$	$\frac{34713}{2860}$
z	$\frac{145}{143}$	$\frac{202}{143}$	1	0	0	0	$-\frac{1}{143}$	$\frac{3857}{286}$
S_2	$\frac{107609}{2860}$	$\frac{8171}{715}$	0	0	1	0	$\frac{7}{572}$	$\frac{24481}{1144}$
S_3	0	0	0	0	0	1	1	203
S_1	$-\frac{107609}{2860}$	$-\frac{8171}{715}$	0	1	0	0	$-\frac{7}{572}$	$\frac{14415}{1144}$

Figure 2.2: Final Tableau of Simplex Algorithm, only nutritional and energetic constraints applied (women, 50-59y)

The solution is

$$f^* = 12.13\$, \quad (x, \; y, \; z) = (0, \; 0, \; 13.48)$$

The total cost of the objective function is lower than the equivalent iteration referred to the other age interval ($f^* = 12.85\$$), while the food combination is similar to the one of the previous case, with again an unrealistic daily consumption of $1.3kg$ of eggs.

Change in Beef Cost

Through a technique equivalent to the one previously adopted (explained in section A.2.2), analyze what happens as the cost of beef changes, with the purpose to include it into the plan.
The reduced cost associated to beef is $-\frac{563}{715}$, thus:

$$\Delta \leq -\frac{563}{715}$$

$$-\infty \leq \overline{c}_x^{new} \leq 1.7 \ + \ \Delta$$

As the problem is rewritten for the boundary value, $\Delta = -\frac{563}{715}$:

$$minimize \ \ f : 0.9x \ + \ 2.3y \ + \ 0.9z$$

$$s.t. \ \ 39.4x \ + \ 13.9y \ + \ 1.75z \geq 11$$

$$39.4x \ + \ 13.9y \ + \ 1.75z \leq 45$$

$$145x \ + \ 202y \ + \ 143z \geq 1928.5$$

$$145x \ + \ 202y \ + \ 143z \leq 2131.5$$

The optimal solution becomes:

$$f^* = 12.13\$, \ \ (x, \ y, \ z) = (0.56, \ 0, \ 12.9)$$

As before, the optimal value of the objective function is not modified, while the final solution set of variables is obtained as a combination of meat and eggs.

Change in Mollusks Cost

Iterating again the procedure, now to let mollusks enter in the set of variables, setting $\Delta = -\frac{1471}{1430}$ the new optimal value becomes:

$$f^* = 12.13\$, \ \ (x, \ y, \ z) = (0, \ 1.87, \ 10.84)$$

Change in Absorption Level

Assess the optimal function and variables value change if a proportion of $30\%,\ 25\%,\ 20\%$ is considered for (x, y, z), that is beef, mollusks and eggs. The new solution is

$$f^* = 12.56, \quad (x,\ y,\ z) = (0.54,\ 0,\ 12.93)$$

Pay attention to the (minimal) increase incurred by the objective function with respect to the value obtained if considering that the total iron content of a given food is absorbed by human body ($f^* = 12.13$).

2.1.3 Non-Heme Iron Sources

Iron can be introduced into the human organism through vegetal sources (as cereals, beams and broad leaf vegetables). In this case study, the focus is on two sources (cocoa and lentils) that are commonly eaten worldwide - or, at least, easy to be introduced.
The linear programming problem, now in two variables, can be formulated as ($18 - 29$ years old women):

$$minimize \quad f: \ 2.5x \ + \ 0.4y$$

$$s.t. \quad 13.9x + 7.2y \geq 16$$

$$13.9x + 7.2y \leq 45$$

$$228x + 360y \geq 2042.5$$

$$228x + 360y \leq 2257.5$$

That can be represented graphically:

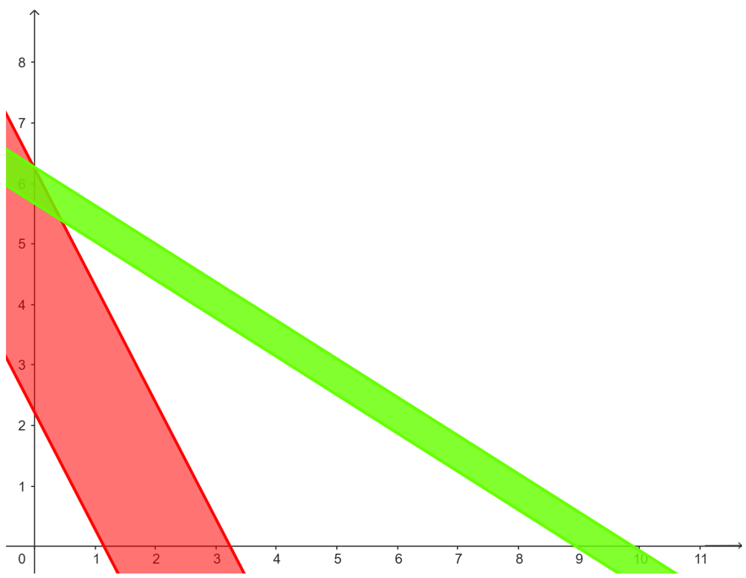


Figure 2.3: Graphical Representation Non-Heme Iron sources (women, 18-29y).

From this, the resolution is pretty much immediate, by identifying the intersection area (a small triangle) and its vertices. Thus, applying Theorem 2, can reduce the set of feasible optimal solutions to those corresponding to the vertices (Figure 2.4).

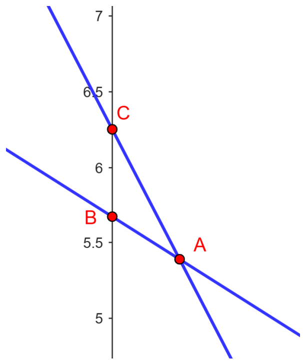


Figure 2.4: Vertices of the Polytope (women, 18-29y).

Where $A = (0.44,\ 5.39)$, $B = (0,\ 5.67)$, $C = (0,\ 6.25)$. Substituted into the objective function, they give: $f_A^* = 2.2692$, $f_B^* = 2.2699$, $f_C^* = 2.5$. Even if the result of f_A^* and f_B^* are pretty much equivalent, especially if considering the physical meaning of it, that is the cost of a daily diet, for a small difference the optimal is f_A^*, and thus A.
Interesting to point out is that applying the simplex method, the reduced cost associated to the variable x (cocoa) is

$$\Delta = -\frac{337}{150}$$

and so, replacing $c_x = 2.5$ with $c_x^{new} = 2.5\ -\ \Delta$, the new optimal value will be

$$f^* = 2.269,\ (x,\ y) = (0,\ 5.67)$$

That is exactly what would get if picking vertex B rather than A. Notice that the change from solution A to B requires cocoa's cost to exceed the c_x of a Δ amount, for this reason the c_x^{new} is obtained as the *difference* between the two. Practically, this means that the cost of cocoa is above the maximum value it can take to still be included in the final solution.
To stress out the concept, the choice between A and B is purely related to personal taste, since the optimal value function is unaffected. Moreover, this diet plan is for sure non-adoptable, lacking in palatability and not only.

Considering the other age range, that is 50-59 years old, the resolution procedure is similar. The difference in upper and lower bounds for energy and in lower bound for iron, though, led to a variation in the polytope, shown below.

$$minimize\ \ f:\ 2.5x\ +\ 0.4y$$

$$s.t.\ \ 13.9x + 7.2y \geq 11$$

$$13.9x + 7.2y \leq 45$$

$$228x + 360y \leq 2131.5$$

$$228x + 360y \geq 1928.5$$

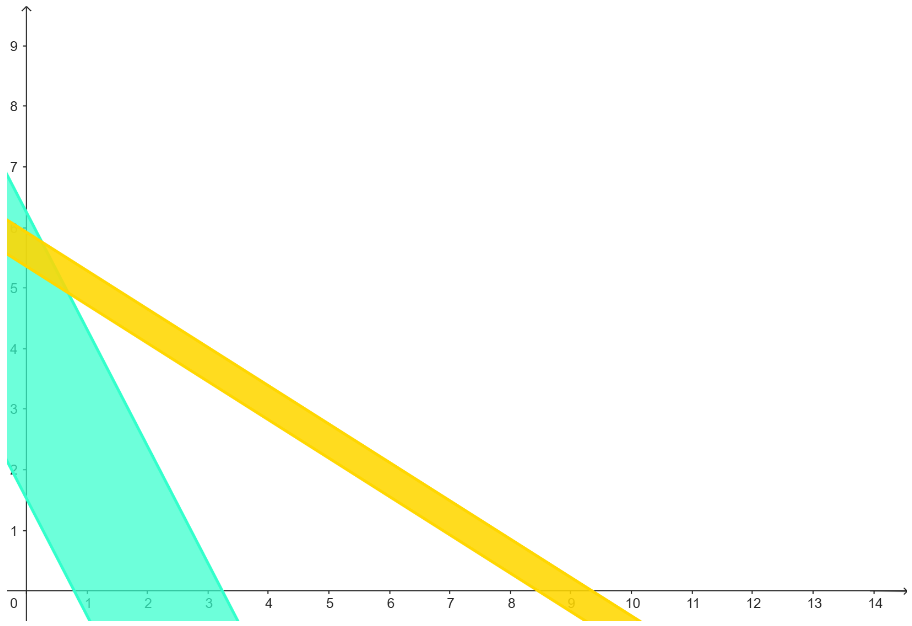


Figure 2.5: Graphical Representation Non-Heme Iron sources (women, 50-59y).

Indeed, the new polygon is a quadrilateral, which means that there are four possible optimal solutions.

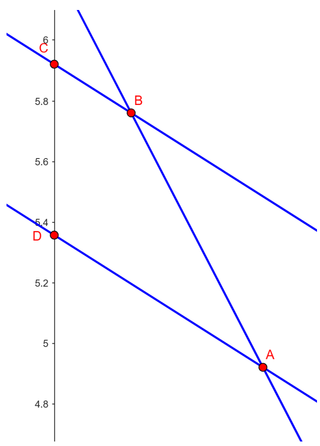


Figure 2.6: Vertices of the Polytope (women, 50-59y).

The vertices are $A = (0.68,\ 4.92),\ B = (0,\ 5.35),\ C = (0.25,\ 5.76), D = (0,\ 5.92)$ with

corresponding optimal: $f^*_A = 3.66$, $f^*_B = 2.14$, $f^*_C = 2.92$, $f^*_D = 2.36$. The lowest cost is given by option B.
As the cocoa cost is lowered of the associated $\Delta = -\frac{337}{150}$ (same as before), the optimal moves to point A.

Further consideration can be made as a change in the absorption coefficient occurs. Indeed, as for the heme-sources, not the entire iron contained into the food considered is actually assimilated and metabolized by human body. Specifically, for vegetable sources, the absorption rate is quite low ($2 - 10\%$). Some different values have been used to check if considering the actual absorption rate the solution still exists, but it emerges to be infeasible in all addressed cases, for both the age ranges.

2.2 Men Iron Intake

Differently than for women, men do not have physiological variation in iron requirement during their life, which is pretty much fixed at a level of 11 mg/day. The reference data that will be used to model the problem are shown the the table 2.4

Table 2.4: Men Iron and Energy Daily Intake Recommendation

Age (years)	Iron (PRI, mg/day)	Energy (AR, $kcal/day$)
18-29	11-45	2670
50-59	11-45	2520

Data from https://multimedia.efsa.europa.eu/drvs/index.htm; PRI is the population reference intake; AR is the average requirement

2.2.1 Heme Iron Sources, 18-29y

Model the LP problem, assuming also in this case a range of 5% for energy supply:

$$minimize \quad f : 1.7x \ + \ 2.3y \ + \ 0.9z$$

$$s.t. \quad 39.4x \ + \ 13.9y \ + \ 1.75z \geq 11$$

$$39.4x \ + \ 13.9y \ + \ 1.75z \leq 45$$

$$145x \ + \ 202y \ + \ 143z \geq 2536.5$$

$$145x \ + \ 202y \ + \ 143z \leq 2803.5$$

The associated final tableau is

Basic	x	y	z	S_1	S_2	S_3	S_4	Solution
Z	$-\frac{563}{715}$	$-\frac{1471}{1430}$	0	0	0	0	$-\frac{9}{1430}$	$\frac{45657}{2860}$
z	$\frac{145}{143}$	$\frac{202}{143}$	1	0	0	0	$-\frac{1}{143}$	$\frac{5073}{286}$
S_2	$\frac{107609}{2860}$	$\frac{8171}{715}$	0	0	1	0	$\frac{7}{572}$	$\frac{15969}{1144}$
S_3	0	0	0	0	0	1	1	267
S_1	$-\frac{107609}{2860}$	$-\frac{8171}{715}$	0	1	0	0	$-\frac{7}{572}$	$\frac{22927}{1144}$

Figure 2.7: Final Tableau of Simplex Algorithm, only nutritional and energetic constraints applied (men, 18-29y)

Which solved gives

$$f^* = 15.96, \quad (x, \ y, \ z) = (0, \ 0, \ 17.73)$$

Change in Beef Cost

The required variation of the reduced cost is the same as in 2.1.1 and 2.1.2. Thus, introducing a $\Delta = -\frac{563}{715}$ the new optimal will be:

$$f^* = 15.96, \quad (x,\ y,\ z) = (0.37,\ 0,\ 17.36)$$

Change in Mollusks Cost

Also in this case, the Δ threshold is the same as before. The implemented solution gives:

$$f^* = 15.96\$, \quad (x,\ y,\ z) = (0,\ 1.22,\ 16.01)$$

Change in Absorption Level

As performed previously, to better capture the real amount of iron that the body can assimilate, consider three different levels of absorption (as before, 30%, 25%, 20% respectively for beef, mollusks and eggs). In this case, the optimal result reached is:

$$f^* = 16.29\$, \quad (x,\ y,\ z) = (0.417,\ 0,\ 17.31)$$

To highlight are the high in minimum diet cost, if compared with the one shown later (2.2.2) and the high amount of eggs suggested, whose presence is almost a certainty in the models seen so far. Thus, can be deduced that even if contributing with lower iron per $100gr$ portion, the eggs' low cost maintains them more convenient than mollusks (or beef).

2.2.2 Heme Iron Sources, 50-59y

Change the linear model as:

$$minimize \quad f : 1.7x \ + \ 2.3y \ + \ 0.9z$$

$$s.t. \ \ 39.4x \ + \ 13.9y \ + \ 1.75z \geq 11$$

$$39.4x \ + \ 13.9y \ + \ 1.75z \leq 45$$

$$145x \ + \ 202y \ + \ 143z \geq 2394$$

$$145x \ + \ 202y \ + \ 143z \leq 2646$$

The iron consumption boundaries are not modified, while the recommended energy daily intake decreases with respect to the previous case, due to a lower energy supply required by older men. The energy boundaries are evaluated as a 5% from the actual value, to allow more flexibility related to different lifestyles.

The final tableau is:

Basic	x	y	z	S_1	S_2	S_3	S_4	Solution
Z	$-\frac{563}{715}$	$-\frac{1471}{1430}$	0	0	0	0	$-\frac{9}{1430}$	$\frac{10773}{715}$
z	$\frac{145}{143}$	$\frac{202}{143}$	1	0	0	0	$-\frac{1}{143}$	$\frac{2394}{143}$
S_2	$\frac{107609}{2860}$	$\frac{8171}{715}$	0	0	1	0	$\frac{7}{572}$	$\frac{4491}{286}$
S_3	0	0	0	0	0	1	1	252
S_1	$-\frac{107609}{2860}$	$-\frac{8171}{715}$	0	1	0	0	$-\frac{7}{572}$	$\frac{5233}{286}$

Figure 2.8: Final Tableau of Simplex Algorithm, only nutritional and energetic constraints applied (men, 50-59y)

With solution:

$$f^* = 15.06, \quad (x,\ y,\ z) = (0,\ 0,\ 16.74)$$

This case, similarly to the others, suggests a very high and unsustainable, from a taste perspective, intake of eggs. Moreover, the increase in function value, that is the cost of the diet plan, is due to the greater energy supply required by men - for both the considered age ranges.

Change Beef Cost

Setting $\Delta = -\dfrac{563}{715}$ gives:

$$f^* = 15.06\$, \quad (x,\ y,\ z) = (0.417,\ 0,\ 16.31)$$

Change Mollusks Cost

For mollusks, set $\Delta = -\dfrac{1471}{1430}$ and obtain:

$$f^* = 15.06\$, \quad (x,\ y,\ z) = (0,\ 1.37,\ 14.8)$$

To point out is the difference between the two cases (Figure 2.1): applying the same Δ variation is not providing the same amount of beef nor mollusks - and this is due to the different strength of the constraints.

Change in Absorption Level

Considering the already used $30\%,\ 25\%,\ 20\%$ absorption levels for beef, mollusks and eggs, respectively. The solution becomes:

$$f^* = 15.42\$, \quad (x,\ y,\ z) = (0.448,\ 0,\ 16.28)$$

which is slightly higher than what would obtain if the beef cost reduces of a $\Delta = -\dfrac{563}{715}$.

2.2.3 Non Heme Iron Sources

Modeling the problem for both the age ranges considered, it appears unfeasible. Indeed, as shown in Figure 2.9 and 2.10, the intersection area (that is, the *polytope*) is on the left side of the cartesian plane, not satisfying $x \geq 0$.

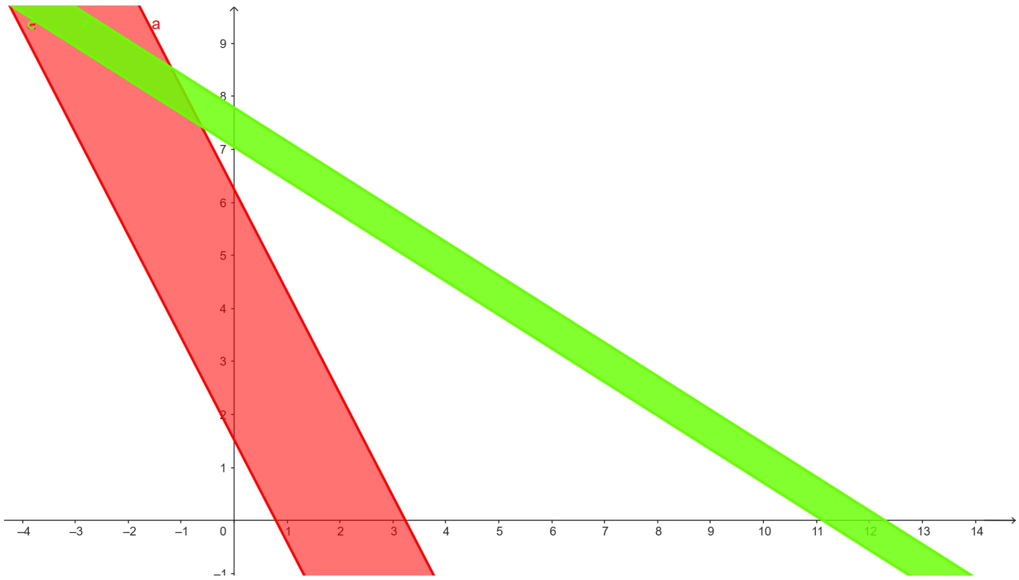


Figure 2.9:
Infeasible Region for Non-Heme Iron Sources (men, 50-59 years old)

What remains to analyze is if a solution exists, and so how the constraints are affected, when a change in the matrix A entrees occur. In practical terms, investigate what happens as realistic absorption rate of iron are introduced in the model. Under this case, the solution becomes feasible for a given range of absorption rate.
At first, consider the year range 18-29. Suppose that cocoa and lentils have the same absorption coefficient. Then, the solution exists as long as it is between $10\% - 6.44\%$. In the first extrema, the objective function value is $f^* = 16.08$ and $(x,\ y) = (5.77,\ 4.13)$. In the lowest value of the interval, $f^* = 30.71$ and $(x,\ y) = (12.28,\ 0.006)$. The difference between the two is evident and emerges both by looking at the objective function and at the contribution of the two variables. In the first case, a lower cost is given thanks to the combination in the final solution of both foods in comparable amounts, with the lentils that, given their low price, help in lowering the total cost. In the other case, instead, the high cost is a direct consequence of the high cocoa contribution to the solution. Moreover, recalling that the function to be minimized is of the total costs, the change in absorption is not favouring the insertion of lentils in the final solution, but cocoa is preferred due to its high content of iron, to satisfy the constraints.

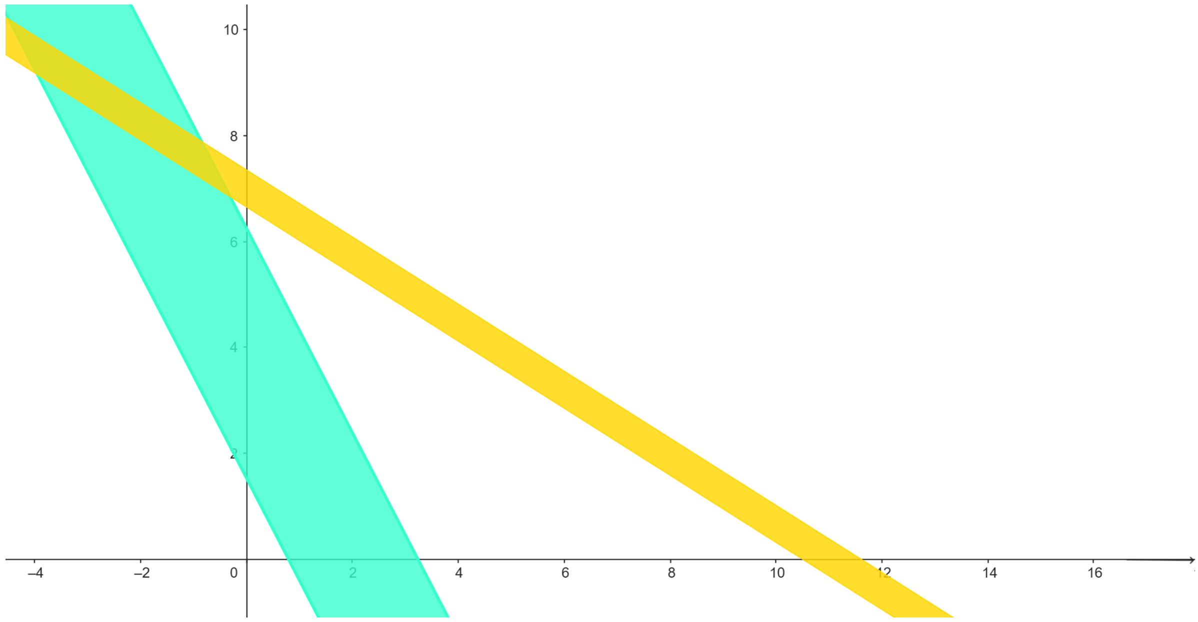


Figure 2.10:
Infeasible Region for Non-Heme Iron Sources (men, 18-29 years old)

Another case addressed, is to consider the two extrema of absorption rate in the case of non-heme iron sources, $10\% - 2\%$, and assuming that they are associated to the foods alternatively. That is, at first 10% for cocoa and 2% for lentils, then the opposite. In the former, the optimal solution is: $f^* = 20.09$, $(x,\ y) = (7.68,\ 2.17)$, with an amount of lentils greater than in previous case. In the latter, the solution is infeasible, not satisfying the non-negativity of the variables.

Regarding the second year range, 50-59, a feasible solution is obtained under the same cases. Assuming that both the foods have the same absorption coefficient, if it varies between $10\% - 6.9\%$, the solution exists. Specifically, having both 10% rate of absorption, $f^* = 16.67$, with $(x,\ y) = (6.11,\ 3.48)$. At the other extreme, the optimal value is way greater: $f^* = 28.52$ and $(x,\ y) = (11.38,\ 0.138)$. In both cases, the amount of cocoa is greater than lentils, meaning that the high iron content of cocoa is "compensating" the higher cost. In other words, even if lentils are significantly less expensive, their low absorption is relevant in choosing the optimal combination.
Alternatively, one may wonder what happens as the absorption of the foods is distinct. Specifically, consider two cases: 10% for cocoa and 2% for lentils, and viceversa. In the first case, the solution exists and shows again that the high cocoa cost is not a good reason to lower its amount, since the associated absorption is great. In numbers: $f^* = 20.03$, $(x,\ y) = (7.73,\ 1.75)$. In the second case, the solution is infeasible.
As with the majority of solutions seen, also these ones are not feasible from the human perspective. Indeed, it is absurd, unsustainable and low in palatability, to even think about

the possibility to eat daily $773gr$ of cocoa and $175gr$ of lentils - just to make an example from the seen solutions.

2.3 Final Considerations

The sensitivity analysis shown may represent a very important starting point for further, deeper and more specific analysis - for example, considering a particular country, population or year range, or again, the actual food sources of iron, or some health problems to be fixed with diet change. Thus, as a starting point, the analysis made had the purpose to be the more general as possible. For this reason, as already explained, the selected foods were those seen above. However, may consider also iron fortified foods, those enriched with iron during the processing.

Regarding future implementations, a more detailed study over the variation of absorption coefficients can be performed, as well as new constraints could be introduced to have a more realistic plan - avoiding extrema as those of $1.4kg/day$ of eggs or $1.2kg/day$ of cocoa powder. Moreover, may be interesting to check how the final solution is affected as this evaluation is inserted into a greater diet plan, so with more foods and, especially, more constraints - for daily intake of carbohydrates, proteins, fat etc. Similarly, can also consider to favour the iron absorption by combining those foods, particularly those of vegetable source, with others. Indeed, has been proven that combining foods rich in iron with other specific nutrients, the iron absorption by the human organism is favoured; similarly, there exist other foods that act on the opposite sense, hindering the assimilation. [22]

In addition, while in this case the environmental impact was not deepened due to infeasibility and lacking of data, an interesting future analysis can also regard this aspect, becoming very important nowadays.

Chapter 3

Quadratic Programming

In the last years, the resolution of the diet problem has also been implemented with a new mathematical methodology: Quadratic Programming. It represents a good alternative to Linear Programming. Indeed, if on one side LP is very flexible, easy to understand and to fit for a computer implementation, some shades are not clearly get. For this and more, the implementation of the diet problem with Quadratic Programming leads to more realistic results.

3.1 Diet effects on Water Consumption

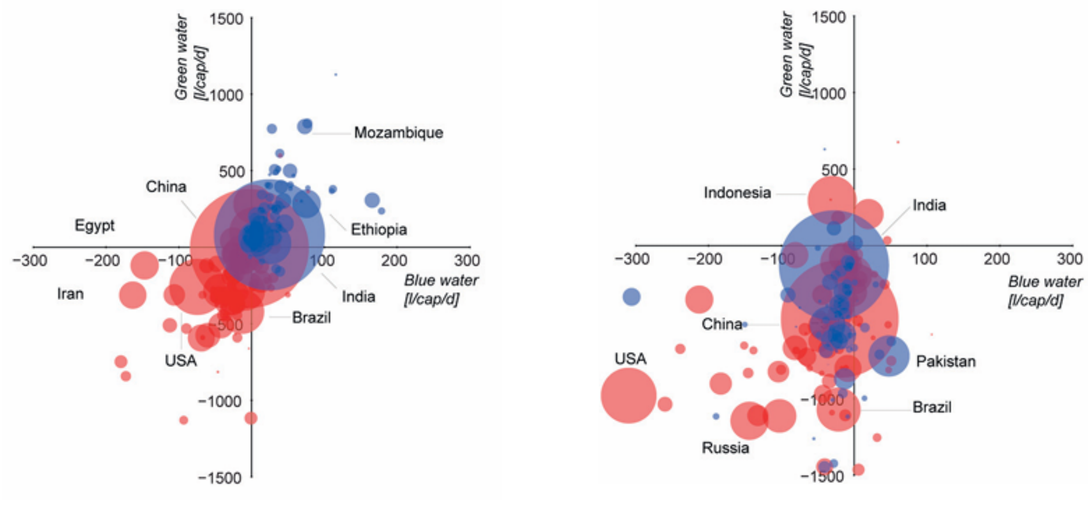
Water is essential for life. Moreover, it is at the basis of wide industrial sectors production processes. It is obvious that cannot be wasted. However, when evaluating the diet environmental impact, the water consume factor is rarely considered as a constraint.
Jalava *et al.*, as Chaudhary and Krishna ([23], [24]), decided to focus their attention toward diet optimization constrained to water consumption, together with other nutritional and environmental impact constraints. The mathematical method used is based on quadratic optimization.

Jalava, Kummu *et al.* considered that as well as diet structure is affected by food availability, culture and traditions of the considered country, also for water availability and consumption a similar evaluation has to be made. Specifically, different countries around the world show different freshwater accessibility and use. In the study, they take the reference nutritional values from the United Nations Food and Agriculture Organization (FAO) data. Also in this research, as those previously seen, the deviations from current diet are limited – for doing so, through quadratic programming is possible to add a "cost" to any deviation occurrence.

At first only nutritional constraints are applied, then animal protein sources are limited to gradually decrease the animal content in diet and asses the effect on water consume –

distinguishing between *blue* water (freshwater in rivers and aquifers) and *green* water (infiltrated rain into soil and accessible by plants' roots). Specifically, the main constraint is applied on meat (whose amount decreases from $\leq 16.7\%$ of protein to 0%). This requires a replacement with vegetable sources, specifically oilseeds and root crops. After a first dietary adjustment related to nutritional constraints, not only protein and fat contribution to the total energy supply vary, but also the global blue water footprint decreased of 1%, while the green water footprint of 2%. Then, as animal-based protein supply is decreased (from 50% of total protein, gradually set as 25%, 12.5%, 0%), a slight reduction in blue water consumption (reaching a maximum reduction of 14% in the case of 0% animal-based protein supply) and a more significant green water decrease (with a maximum of 21% in the case of 0% animal-based protein supply). Those data are referred to a global evaluation.
Moving deeper to each country, some of them show even an increase in blue water footprints as animal protein is increased. This is due to the fact that, as animal sources decreased, the protein contribution has to be compensated by vegetal products, that for cultivation require high irrigation (indeed, the irrigated land due to agriculture is way greater than water use for pasture, 61% and 7.4% respectively).

Thus, a diet change can have global impact on the green and blue water waste, even if locally it may lead to an opposite trend.



(a) From Original Diet to Recommended Diet

(b) From Recommended Diet to Animal Zero

Figure 3.1: The size of the bubbles is proportional to the population of the country.

So, to significantly modify water waste is necessary to strongly change the diet composition, with a drastic decrease in animal protein sources. However, such type of substitution may generate other types of nutritional deficiencies (especially if considering micronutrients). In fact, a limitation of this research is the lack in nutritional constraints, referring only to macronutrients and energy supply. Moreover, as recognized by Jalava *et al*. themselves, the paper focuses too much on diet and water consumption, without assessing or somehow limiting the potential changes that the observed results require (e.g., an increase in agricultural activities may not be sustainable by all countries) and neither considering other possible constraints, as diet cost.

Taking as a starting point the global analysis made by Jalava *et al*., the Chaudhary and Krishna research [24] introduced a more detailed diet structure, and with more factors to assess its environmental footprint.
Specifically, they took nutritional recommendation data from the World Health Organization (WHO) for 24 essential nutrients and 4 constraints for preserving human health (total fats, cholesterol, saturated fats, sugar).
The environmental footprint constraints refer to: greenhouse gas emission, freshwater use, cropland use, nitrogen application, phosphorous application – these last two limits are introduced since such fertilizers have negative impacts on aquatic biodiversity.
The objective function is defined with the goal to minimize the possible deviation from the current diet:

$$minimize \ \ \Delta = \sum_{i=1}^{221} \left[\frac{Q_{opt,i} - Q_{obs,i}}{Q_{obs,i}} \right]^2$$

where $Q_{opt,i}$ and $Q_{obs,i}$, are the optimum and current intake amounts ($g \ capita^{-1} \ day^{-1}$ of food item i).
The final results show that 8 of the 24 micronutrients did not reach the recommended level. Moreover, the countries with higher income are those to which is required a greater reduction in environmental footprints, while in some other nations increase due to an addition of fruit, vegetables and legumes, in order to reach nutritional recommendations. On a global level, the GHG emissions reduced of 32%, blue water consumption reduced of 8%, 12% and 13% decrease for nitrogen and phosphorous.
Interesting to notice, but not different from expectation, is that among all those countries showing a departure from current diet, those with highest variation are spread through all continents (USA, Mexico, Honduras, Philippines etc.) due to different reasons. Specifically, South and North American countries' displacement is due to the typical high meat and eggs consumption (replaced with cereals), while African countries' variation is motivated simply by the requirement of increasing the energy level and more in general for the nutritional values. Differently, the Mediterranean regions do not show an high displacement from current diet since the diet already contains high fruit, vegetables and legumes, remaining within the optimal boundaries. [15]

The limitation of this study, even if it implements in a more complete way the optimization with respect to Jalava *et al*., is to not consider the economic aspects. Moreover, it could be further implemented if cultural aspects are taken into account, as well as cooked foods.

3.2 Diet and GHGE: UK and Danish comparison

The relationship of the diet plan with the greenhouse gas emissions is well known, as addressed in the previous paragraphs, and can be easily summarized by the FAO's definition of sustainable diet [11]. Indeed, food plays a crucial role in GHG emissions, from processes of agricultural production, processing, transport, storage, cooking, waste, and more developed countries tend to consume diets with higher GHG emissions. Moreover, the more foods are processed foods, the more the GHG emission level associated to them. It's also relevant to consider that those foods, in general, are those contributing less (and even negatively) to the construction of a health dietary plan. Thus, countries around the world have committed a goal to decrease their GHGE of a given amount to avoid negative consequences for humanity. The United Kingdom has to reduce such emissions of 80% relative to 1990 levels by 2050, and has been suggested a 70% reduction in emissions from food to achieve this goal. Specifically, in the case of the UK, current diets fail to satisfy basic recommendations such as daily portions of fruit and vegetables.

In their research, Green *et al.* [25] considered data for 1571 adults to estimate the average composition of diet distinguished by gender. The food used for the analysis were aggregated into 42 groups. The objective function required to minimize the loss of welfare for all the 42 food groups, defined as:

$$\Delta W_i \propto \frac{s_i}{\epsilon_i}\left(\frac{\Delta X_i}{X_i}\right)^2$$

s_i is the share of expenditure for that food group, and ϵ_i is the price elasticity of demand, the ratio of these terms acts as a proxy measure of how likely people would be to modify their consumption of that food group. X_i is the current consumption for group i and ΔX_i is the difference between current and ideal consumption for food group i. So, the objective function is:

$$minimize \quad \sum_{i=1}^{42} \Delta W_i \propto \frac{s_i}{\epsilon_i}\left(\frac{\Delta X_i}{X_i}\right)^2$$

The applied constraints can be classified as:

- **nutritional**: taking as reference the WHO recommendations for daily consumption of macronutrients. Moreover, the calories and liquids consumed are maintained, to ensure similarity with the actual diet plan;
- **environmental**: gradual reduction of the GHGE emission is imposed, moving with 10% steps, till 60% cut.

At a first stage, as only nutritional constraints are applied, a decrease of emission of 17% is registered. This confirms what was previously addressed, that is an healthy diet is good also

for the environment. Specifically, those foods with greater greenhouse-gas emissions are red meat and butter, with eggs and dairy products. Thus, they are reduced in favor of cereals, fruit and vegetables.
Then, the allowed emissions are decreased gradually. As a 40% reduction is applied, fruits and vegetables are substituted by sugary snacks, more processed but nutrient and energy dense. Such introduction is highly debatable, but the nutritional constraints are satisfied. Furthermore, the extreme reduction of GHGE suggests a vegetarian diet plan, that is not substainable for the entire population.

A similar study, considering Danish adults diets, was performed by Nordman *et al.* [26] taking data from the Danish National Survey of Diet and Physical Activity.
The 434 foods were aggregated into 50 sub-groups based on nutritional, culinary and environmental impact characteristics; the foods are considered both in raw/uncooked and cooked/processed states. The advantage that lies behind the food clustering is that the final result will be expressed as function of a subgroup, which actually make the final plan more realistic, ensuring flexibility within a subgroup. The content of nutrients in all subgroups is evaluated through the following function

$$A_{ij} = \sum_{k=1}^{n_j} \frac{x_{kj}}{\sum_{k=1}^{n_j} x_{kj}} \times a_{ik}$$

where A_{ij} is the content of nutrient i per gram of food sub-group j; n_j is the number of food items belonging to food sub-group j; x_{kj} is the quantity of food item k of food sub-group j consumed in the population and a_{ik} is the content of nutrient i in food item k.
The objective function was defined as a quadratic function expressing the deviation from the current plan. The quadratic form ensures to have a greater penalization of variations from the observed diet plan.

$$minimize \quad \sum_{j=1}^{50} \left(\frac{x_j - x_{obs,j}}{x_{obs,j}} \right)^2$$

with x_j and $x_{obs,j}$ being the population average amount of food sub-group j in grams in the optimized and observed diet, respectively. With the purpose to monitor variations from the present plan, a coefficient (called *diet departure score*) is used:

$$\Delta_{diet}(\%) = \frac{1}{50} \times \sum_{j=1}^{50} \left| \frac{x_j - x_{obs,j}}{x_{obs,j}} \right| \times 100$$

The constraints are applied at nutritional level, health target level and GHGE, applied alone and combined.
The results confirm the result of the UK study, that is decreasing meat and dairy products, raising vegetables, fruits and cereals. Thus, another evaluation was made, checking the similarity between the plan resulting from nutritional, healthy and GHGE constraints applied,

and the plant-based plan, being both optimal under the environmental impact constraint. In this case, the difference emerges in the presence of (low) meat amount in the optimized model, absent in the plant-based diet. So, the two can be considered alternatives leading both to low greenhouse gas emissions, offering the possibility to choose according to personal taste.

Conclusion

The diet problem, born in a war context, is nowadays applied and applicable in all the world areas and for multiple optimization reasons - above the economical one.
Indeed, it can be tailored to address specific local cases, so to consider local habits, food availability and culture, factors that affect the diet plan. In addition, a global analysis can even be performed, considering the different nutritional behaviours around the world, and comparing them through a given criterion, like environmental impact or economic affordability.
The different methodologies seen have advantages and disadvantages, requiring to look for a trade-off when dealing with this problem. In fact, Liner Programming, the first methodology used in this field, provides great results and flexibility, even for computer implementation (as an example, is left for curiosity the link for the code used to solve the diet problem with the data used by Stigler: `https://developers.google.com/optimization/lp/stigler_diet`, whose code is reported in Section A.4). In addition, Linear Programming is also easy to utilize for performing sensitivity analysis, as seen in Chapter 2.
On the other side, Quadratic Programming represents just a specific application of what is actually related to Non-Linear Programming. This method offers the possibility to give strong importance (that is, emphasized by quadratic function) to the deviation from the current diet, discouraging it as is possible, with the purpose to have a new plan that is both acceptable from a cultural and habit perspective, but also in terms of economical, nutritional and/or environmental aspects. So, future analysis may focus on the application of this method not widely used yet.
Moreover, a problem that has not be treated by (almost) any of the examples seen in this work, is related to the combination of foods. Specifically, can consider this under two perspectives. First, food combination is essential to avoid mixing those foods that would never eat together - like meat and fish. Second, some foods, if eaten together, can raise the absorption of some micro-nutrients, thus their mixing should be encouraged ([27]). The implementation of a diet plan considering these factors was proposed by Donati *et al.*, [28].
To conclude, the deeper aspects of such problem have yet to be discovered and analyzed, offering wide choice both in terms of case under study and of mathematical method to use.

Appendix

A.1 Simplex Method

It is essential to look at the mathematical basis and the subsequent methodologies that lie behind the Simplex Algorithm. [29]
Linear programming problems are presented in the canonical form:

$$minimize \;\; z = \mathbf{c}^T\mathbf{x} \;\; s.t. \; A\mathbf{x} \leq \mathbf{b} \; and \; \mathbf{x} \; \geq \mathbf{0}$$

where $\mathbf{x}$ is a vector with components the unknown variables, $\mathbf{b}$ is the column vector associated to the constant terms, $\mathbf{c}$ is a row vector of constant terms, A is a matrix of size m by n.
The feasible region, basically a polyhedron (more generally a *polytope*), is a convex set of the space delimited by the intersection of hyperplanes ($\{x \mid a'x \geq b\}$) and half-spaces ($\{x \mid a'x \; = b\}$).

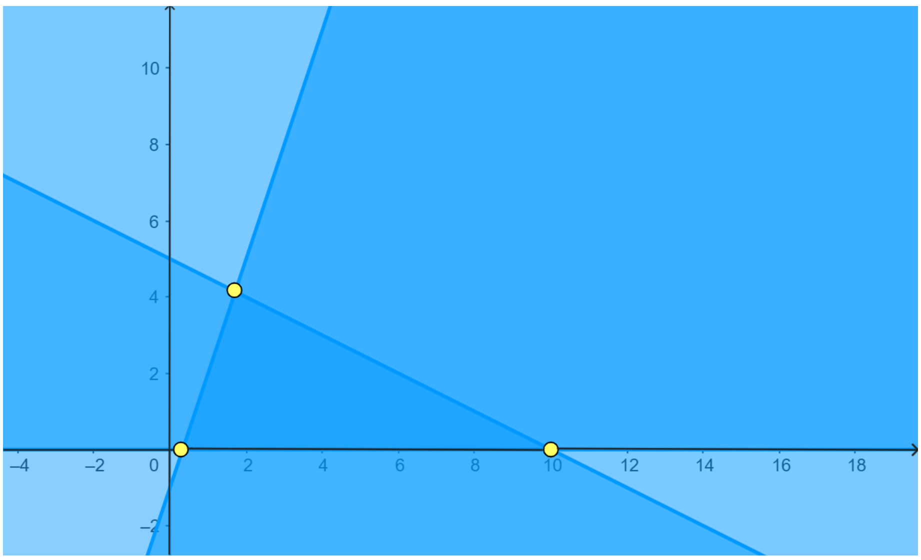


Figure A.2: The dark area with yellow vertices corresponds to the polytope identified as the intersection of the various constraints

Theorem 1. *Being $P \subset \mathbb{R}^n$, with vertices $v^1,\ v^2, ...,\ v^k, \forall\ v^i \in \mathbb{R}^n$, then x is a convex combination of the vertices of P, so holds that $x \in P,\ x\ = \sum_i \lambda_i v^i,\ \lambda_i \geq 0,\ \sum_i \lambda_i = 1$. This means that is possible to write any point of the polyhedron as a convex combination of its vertices. For example, having only two points $x,\ y$ can write the convex combination as $\lambda x\ + (1-\lambda)y,\ \lambda \in [0,1]$ to identify all the points, along the segment, between x and y.*

Theorem 2 (Optimal Vertex)**.** *Given a linear programming problem in the form $min(max)\ \{\boldsymbol{c}^T\boldsymbol{x}\ : \boldsymbol{x} \in P\}$, and being $P = \{\boldsymbol{x} \in \mathbb{R}^n\ :\ A\boldsymbol{x} = \boldsymbol{b}, x \geq 0\}$ a bounded non-empty region, x is a feasible basic solution if and only if it correspond to a vertex of P and among those is possible to identify the optimal.*

Theorem 2 limits the search for optimal among the vertices of the feasible region.

Consider now the standard linear programming problem formulation. The matrix $A \in \mathbb{R}^{m\ \times\ n},\ n > m$, has at maximum a rank of size m. Consequently, is possible to identify a basis B for A, $B \in \mathbb{R}^{m \times m}$ by selecting m independent columns of the matrix A - which also implies $det(B) \neq 0$, so B is invertible. In this way, is possible to rewrite the matrix A as $[B \mid F]$, with $F \in \mathbb{R}^{(n-m) \times n}$.
Also $\mathbf{x}$ can be rewritten as a combination of these two matrices:

$$\mathbf{x} = [\mathbf{x}_B\ \mathbf{x}_F]^T$$

where the vector $\mathbf{x}_B \in \mathbb{R}^m$ refers to the *basic* variables, while $\mathbf{x}_F \in \mathbb{R}^{n-m}$ is the vector of the *non-basic* variables.
So, from $A\mathbf{x} = \mathbf{b}$ can write $\mathbf{b} = B\mathbf{x}_B\ +\ F\mathbf{x}_F$, from which is immediate:

$$\mathbf{x}_B = B^{-1}\mathbf{b}\ -\ B^{-1}F\mathbf{x}_F$$

The *basic solution* is the one obtained by setting $\mathbf{x}_B\ =\ B^{-1}\mathbf{b}$ and $\mathbf{x}_F\ =\ 0$. If the *non-negativity* condition is satisfied ($\mathbf{x}_B = B^{-1}\mathbf{b} \geq 0$), the basic solution is *feasible*, so inside the polyhedron identified by the constraints.

Based on these fundamental theorems and results, is possible to rewrite an LP problem in the canonical form with respect to the basis B: all the basic variables and the objective function will be expressed as a function of the *non*-basic variables as follows:

$$z = z_B\ + c_{F_1}x_{F_1}\ +\ ...\ +\ c_{F_{n-m}}x_{F_{n-m}}$$

$$x_{B_i} = b_i\ -\ a_{iF_1}x_{F_i}\ -\ ...\ - a_{iF_{n-m}}x_{F_{n-m}}$$

where z_B and b_i are scalars; c_{F_j} is the coefficient of the j-th non-basic variable in the objective function (*reduced cost*); $-a_{iF_j}$ is the coefficient of the j-th non basic variable in the constraint that refers to the i-th basic variable, expressed with negative sign for pure convenience. In matrix form:

$$\mathbf{x}_B = B^{-1}\mathbf{b} \ - \ B^{-1}F\mathbf{x}_F$$

Being the coefficients of the non-basic variables

$$A_{F_j} = F_j = B^{-1}F_j$$

The objective function is

$$z = \mathbf{c}^T\mathbf{x} = \mathbf{c}_B^T\mathbf{x}_B \ + \ \mathbf{c}_F^T\mathbf{x}_F = \mathbf{c}_B^T(B^{-1}\mathbf{b} \ - \ B^{-1}F\mathbf{x}_F) \ + \ \mathbf{c}_F^T\mathbf{x}_F$$

That is, grouping $\mathbf{x}_F$:

$$z = \mathbf{c}_B^T B^{-1}\mathbf{b} \ + \ (\mathbf{c}_F^T \ - \ \mathbf{c}_B^T B^{-1}F)\mathbf{x}_F$$

The reduced costs, that is the coefficients of the non-basic variables in the objective function are:

$$\bar{c}_j = c_j \ - \ c_B^T B^{-1}A_j$$

with A_j referring to the j-th column in matrix A in the initial form of the matrix - before matrix manipulation starts.

Let's now introduce two fundamental conditions to check optimality and feasibility of a solution.

- **Optimality Condition:** Considering a linear programming problem in the canonical form with respect to the basis B, if all the reduced costs with respect to this basis are *non-negative*, that solution is optimal.
  Notice that the reduced costs appear in the objective function only as related to the non-basic variables, so can assume that those associated to the basic variables are null. In other words, that condition can be reformulated as referred only to the reduced costs of the non-basic variables, imposing that have to be non-negative. To sum-up, the optimality condition is satisfied whenever $c_j \geq 0 \ \forall j$. Equivalently:

$$\mathbf{c}^T \ - \ \mathbf{c}_B^T B^{-1}A \geq \mathbf{0}^T$$

- **Unlimited Condition:** If there is *any* non-basic variable x_h such that

$$c_h < 0 \ \wedge \ a_{ih} \leq 0 \ \forall i = 1, \ ..., m$$

  the problem is unbounded. Indeed, it would mean that is possible to increase the basic variables associated to a negative coefficient without any limit, consequently reducing the objective function optimal value (remember that the goal is to minimize the z function). In other words, there's no bound in the feasible region, so the objective function can decrease infinitely, till $-\infty$.

Both those conditions have to be verified when iterating the Simplex Algorithm, to assess when the optimum is reached.
Indeed, the steps of the algorithm can be reformulated as follows:

1. Identify an initial basis $B = [A_{B_1}|...|A_{B_m}]$ and compute B^{-1} and $c_B^T B^{-1}$;

2. Compute the reduced costs $\overline{c}_j = c_j \; - \; c_B^T B^{-1} A_j$ referred to the non-basic variables;

3. Perform the Optimality Condition test: if $\overline{c}_j \geq 0$ for every non-basic variable, the solution is optimal;

4. If not, keep going by selecting the non-basic variable x_j with $c_j \; < \; 0$ (picking the smallest if there are more than one negative reduced costs);

5. Calculate $\overline{\mathbf{b}} = B^{-1}\mathbf{b}$ and $\overline{A}_j = B^{-1} A_j$; if one of $a_{ij} \leq 0 \; \forall i$ the problem is unbounded and there's nothing more to do (Unlimited Test);

6. Otherwise, compute the *ratio test*: select the minimum among the ratios $\dfrac{\overline{b}_i}{\overline{a}_{ij}}$ and use it as the pivot for Gauss-Jordan matrix computations;

7. The basis is now updated with a new entry variable (the j-th one); the algorithm can restart by computing B^{-1} and $\mathbf{c}_B^T B^{-1}$ (the basis B is the same chosen previously).

This definition of the simplex method based on matrices can be moved to a different perspective of computations through the following steps:

1. Set the problem in the maximization form, that is, from the canonical linear programming form, multiply by -1 the objective function, maintaining the constraints in the "less than or equal to" form;

2. Introduce slack variables, needed to move from inequalities to equalities constraints. There are added to the constraints with a positive coefficient and their non negativity has to be imposed;

3. Set the tableau, representing in each row the coefficients of the variables (slack included) and on the last column the b constants. The last row is for the objective function;

4. If $c_j \geq 0 \; \forall j \rightarrow$ optimal solution is reached, otherwise means that can improve the solution;

5. To improve the solution, everything starts by identifying the pivot variable, that will be increased to minimize the objective function. To address the pivot:

   (a) Pick smallest value in last row (considering negative sign!) and select that column;

(b) Compute the ratio test: the ratio between the **b** coefficients and the corresponding coefficients in the selected column;

(c) Consider the higher ratio value (in absolute value);

(d) The pivot comes from the intersection between the row associated to the higher ratio and the previously selected column;

(e) Start using that pivot as a pivot: maintain it as the unique 1 in the column, setting 0 the other coefficients through specific operations, and recompute the **b** values and the objective function value;

(f) Two possibilities:

   i. All coefficients in last row are positive or null $\rightarrow$ the optimum is reached;

   ii. one or more coefficients in last row are negative $\rightarrow$ re-iterate the process until the final row is composed of non-negative values, at that point the maximum for the objective function will be reached.

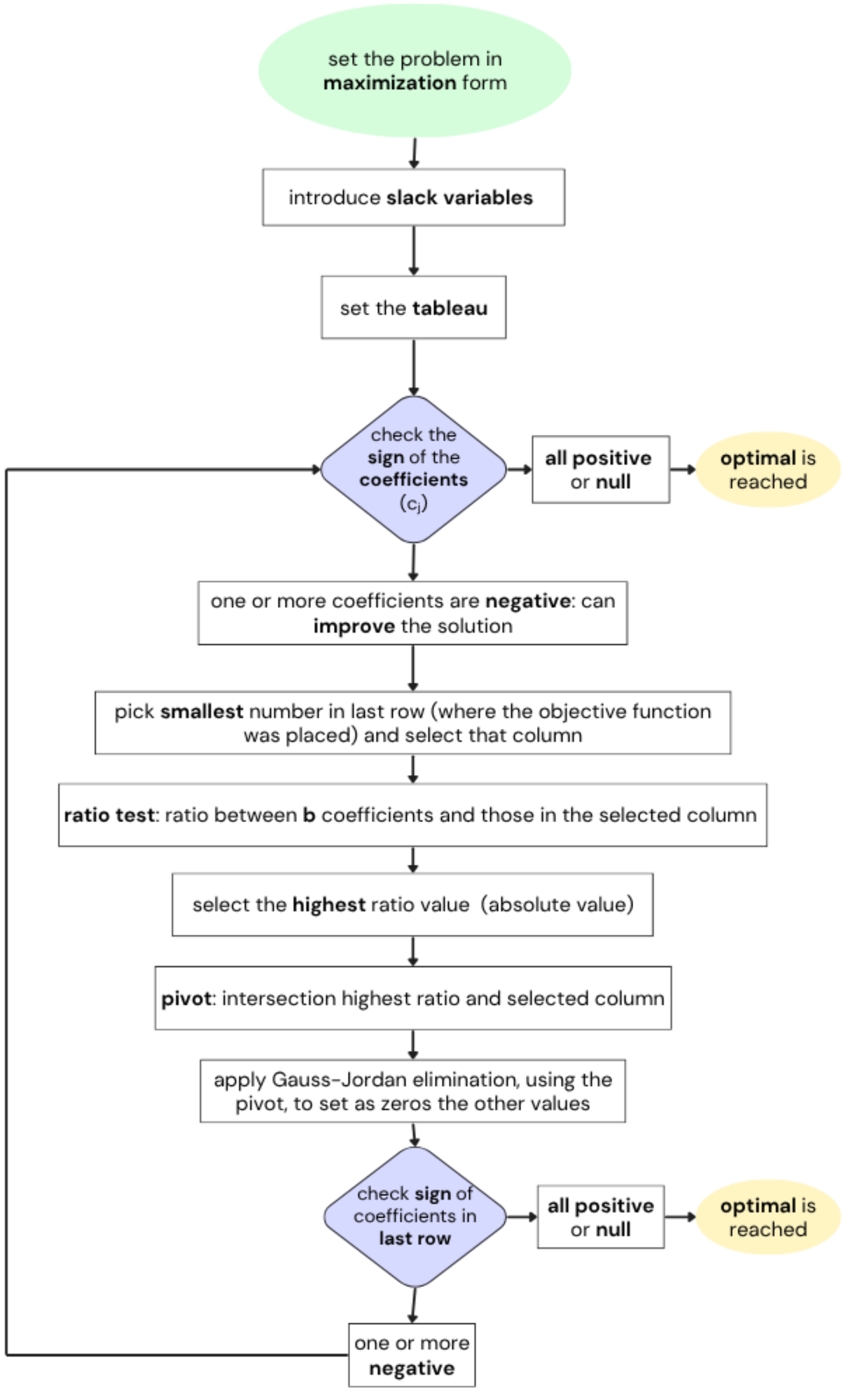


Figure A.3: Flowchart describing the Simplex's steps.

A.2 Sensitivity Analysis

The simplex method is the algorithmic approach to linear programming problems resolution. Looking at this problems from a more "practical" perspective, that is, for example, as maximization of costs under production and capacity constraints, what is relevant for managers is to look at all the various scenarios - as a decrease in item's cost, an increase in production rate and similar. That is, performing *sensitivity analysis*.
As changes to the elements involved in the LP problem occur, is necessary to verify that the two fundamental conditions of optimality and feasibility hold (A.1):

- Feasibility Condition: $\mathbf{x}_B = B^{-1}\mathbf{b} \geq 0$
- Optimality Condition: $\mathbf{c}^T \;-\; \mathbf{c}_B^T B^{-1} A \geq \mathbf{0}^T$

A.2.1 Change in b

Whenever a change in the **b** vector is performed, that is the values of the right-hand side are modified (*e.g.* there is an higher availability of one type of resource), it is important to check that the problem is still feasible and if something has changed in the optimal solution. Actually, by looking at the optimality and feasibility conditions is evident that a change in **b** cannot affect the optimal (to be more precise, if the modified right hand side is the one of a binding constraint, a great variation of it may affect the solution).
The linear programming problem with **b** variability can be written as:

$$minimize \;\; \mathbf{c}^T\mathbf{x}$$

$$s.t. \; A\mathbf{x} = \mathbf{b} \;+\; \Delta\mathbf{e}_i, \;\; \mathbf{x} \geq 0$$

So, being $\beta_{ij} = [B^{-1}]_{ij}$ and $\bar{b}_j = [B^{-1}\mathbf{b}]_j$, the feasibility condition can be rewritten as:

$$(B^{-1}\mathbf{b})_j \;+\; \Delta(B^{-1}\mathbf{e}_i)_j \geq 0$$

$$\bar{b}_j \;+\; \Delta\beta_{ij} \geq 0$$

From which follows:

$$\overline{\Delta} \leq \Delta \leq \underline{\Delta}$$

with $\overline{\Delta} = max(-\frac{\bar{b}_j}{\beta_{ij}}), \;\; \beta_{ij} > 0$ and $\underline{\Delta} = min(-\frac{\bar{b}_j}{\beta_{ij}}), \;\; \beta_{ij} < 0$. Until the range of variation of a component of **b** remains within this range, the feasibility condition is not affected.

A.2.2 Change in c

x_j **basic**

Differently than previously, now only the optimality condition may fail. In the specific case of x_j basic variable, assume that $\mathbf{c}_B^T$ is replaced by $\mathbf{c}_B^T + \Delta\mathbf{e}_j$. The condition transforms into:

$$\mathbf{c}^T - (\mathbf{c}_B^T + \Delta\mathbf{e}_j)B^{-1}A \geq 0$$

$$(\mathbf{c}^T - \mathbf{c}_B^T B^{-1}A) - \Delta\mathbf{e}_j B^{-1}A \geq 0$$

Taking into account that $\mathbf{x}_j$ is basic and so has reduction costs equal to zero ($\mathbf{c}_B^T B^{-1}A = 0$), the condition becomes:

$$\mathbf{c}^T - \Delta\mathbf{e}_j B^{-1}A \geq 0$$

Selecting the i-th component:

$$\overline{c}_i \geq \Delta\overline{a}_{ji}$$

To conclude

$$\overline{\Lambda} \leq \Delta \leq \underline{\Lambda}$$

with $\overline{\Lambda} = max(\frac{\overline{c}_j}{\overline{a}_{ji}}), \;\; \overline{a}_{ji} < 0$ and $\underline{\Lambda} = min(\frac{\overline{c}_j}{\overline{a}_{ji}}), \;\; \overline{a}_{ji} > 0$.

x_j **non-basic**

The $\mathbf{c}_B$ is unaffected. The optimality condition to be verified is:

$$(\mathbf{c}^T + \Delta\mathbf{e}_j) - \mathbf{c}_B^T B^{-1}A \geq 0$$

$$(\mathbf{c}^T - \mathbf{c}_B^T B^{-1}A) - \Delta\mathbf{e}_j \geq 0$$

So it gets:

$$\overline{\mathbf{c}} + \Delta\mathbf{e}_j \geq 0$$

$$\Delta \geq -\overline{c}_j$$

A.2.3 Adding a constraint

As a new constraint is introduced in the problem, the previous solution can or cannot be the optimal. If it still is, means the constraint is no more restrictive than those already present. Otherwise, the problem may be unfeasible (which is the most common outcome) or a new optimal solution exists that would further minimize your objective function value.

A.2.4 Change in A

Suppose that a_{ij} is implemented as $a_{ij} + \Delta$, assuming that A_j is not in the basis B. Then, the feasibility condition is not affected, differently than the optimality one:

$$c_j - \mathbf{c}_B^T B^{-1}(A_j + \Delta \mathbf{e}_i) \geq 0$$

$$\bar{c}_j - \Delta \rho_i \geq 0, \quad \rho = \mathbf{c}_B^T B^{-1}$$

$$\Delta \leq \frac{\bar{c}_j}{\rho_i}$$

For Δ below this ratio, the optimal solution is not affected.
If A_j is in the basis, other considerations regarding also feasibility should be addressed.

A.3 Quadratic Programming

Quadratic Programming (QP) is a mathematical approach to solving optimization problems in non-linear form, being a subclass of Non-Linear Programming (NLP).
Quadratic Programming problems, as the name says, are characterized by an objective function in quadratic form, that is:

$$minimize \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q \mathbf{x} + \mathbf{c}^T \mathbf{x} + cost$$

with Q the coefficient *symmetric* matrix, $\mathbf{x}$ and $\mathbf{c}$ being the coefficient vector of linear part and the vector of unknowns, respectively.
The problem is constrained as:

$$A\mathbf{x} \leq \mathbf{b}$$

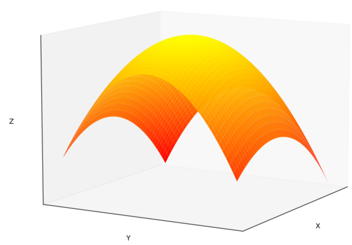
To solve it, different methodologies exists, and only to give a general overview the more basic ones will be addressed here.

A.3.1 Convexity

Checking the convexity of the quadratic function is really helpful in understanding the presence of a minimum or maximum. Indeed, the way the Q matrix is definite is a sufficient but non-necessary condition.
Symmetric matrices ($A = A^T$) can be classified as follows, depending on the sign of their eigenvalues:

Positive definite: all the eigenvalues are strictly positive;

Negative definite: all the eigenvalues are strictly negative;



Semi-Positive definite: non-negative eigenvalues;

Semi-Negative definite: non-positive eigenvalues;

Non definite: some eigenvalues are positive, some eigenvalues are negative.

So, being in this specific case the matrix Q equivalent to the Hessian matrix - that is, with entrees the second partial derivatives of the function - the eigenvalues to be considered are those of the Q matrix. The sign of the eigenvalues is expressing the direction toward which you have to move, thus if are all positive, you have to move upwards, meaning you are in a minimum point (*convex function*); with the equivalent reasoning can identify a maximum (*concave function*).

A.3.2 Karush-Kuhn-Tucker

It is a method that can be used with any non-linear programming problem, and so also with QP cases.
First, it requires the definition of the Lagrangian function:

$$L = f(\mathbf{x}) \ - \ \sum_{i=1}^{m} \mu_i g_i \ - \sum_{j=1}^{n} \lambda_i h_i$$

where $f(\mathbf{x})$ is the objective function, $g_i(\mathbf{x})$ refers to the constraints in inequality form (≤ 0), while $h_j(\mathbf{x})$ is referred to the equality constraints ($= 0$). They are multiplied by the Lagrangian multipliers: μ_i associated to inequalities and λ_i associated to equality restraints.
Then, the Lagrangian function has to be derived with respect to each entry of the vector $\mathbf{x}$ (that is, each variable of the function), and setting all the equations equal to zero. In the system, linear considering $f(\mathbf{x})$ as quadratic, has to be included the sign constraint of the Lagrangian multipliers: μ_i have to be *non*-negative, while λ_i can be any real number.
So, the system has to be solved, finding the optimal solutions as the constraints are activated (the advantage is to decide, according on the sign of the multiplier, which constraint to consider), and concluding evaluating which points are maxima or minima.

The Quadratic Programming addressed here is only a specific case of Non-Linear Programming, as already said, but more specifically is a subset of Quadratically Constrained Quadratic Programming. It is presented in the form:

$$minimize \ \ f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} \ + \ \mathbf{c}^T\mathbf{x} \ + \ cost$$

$$s.t. \ \ \frac{1}{2}\mathbf{x}^T P_i\mathbf{x} \ + \ \mathbf{q}_i^T\mathbf{x} \ + d_i \ \leq 0 \ \ \forall i = 1, ...m$$

$$A\mathbf{x} = \mathbf{b}$$

That is, inequalities constraints are in a quadratic form as long as $P_i \neq 0$. In this case, the KKT algorithm can still be performed, in a similar way as described for a Quadratic Programming case: building the Lagrangian, deriving it with respect to the variables and building a linear system with equalities set equal to zero; solve it ensuring the satisfaction of the sign constraints associated to the lagrangian multipliers.

A.4 Python Code for Stigler Diet Problem

```
 #import required library
 from ortools.linear_solver import pywraplp

 # Nutrient minimums.
nutrients = [
    ["Calories (kcal)", 3],
    ["Protein (g)", 70],
    ["Calcium (g)", 0.8],
    ["Iron (mg)", 12],
    ["Vitamin A (KIU)", 5],
    ["Vitamin B1 (mg)", 1.8],
    ["Vitamin B2 (mg)", 2.7],
    ["Niacin (mg)", 18],
    ["Vitamin C (mg)", 75],
]

# Commodity, Unit, 1939 price (cents), Calories (kcal), Protein (g),
# Calcium (g), Iron (mg), Vitamin A (KIU), Vitamin B1 (mg), Vitamin B2 (mg
   ),
# Niacin (mg), Vitamin C (mg)
data = [
    # fmt: off
  ['Wheat Flour (Enriched)', '10 lb.', 36, 44.7, 1411, 2, 365, 0, 55.4,
   33.3, 441, 0],
  ['Macaroni', '1 lb.', 14.1, 11.6, 418, 0.7, 54, 0, 3.2, 1.9, 68, 0],
  ['Wheat Cereal (Enriched)', '28 oz.', 24.2, 11.8, 377, 14.4, 175, 0,
   14.4, 8.8, 114, 0],
  ['Corn Flakes', '8 oz.', 7.1, 11.4, 252, 0.1, 56, 0, 13.5, 2.3, 68, 0],
  ['Corn Meal', '1 lb.', 4.6, 36.0, 897, 1.7, 99, 30.9, 17.4, 7.9, 106,
   0],
  ['Hominy Grits', '24 oz.', 8.5, 28.6, 680, 0.8, 80, 0, 10.6, 1.6, 110,
   0],
  ['Rice', '1 lb.', 7.5, 21.2, 460, 0.6, 41, 0, 2, 4.8, 60, 0],
  ['Rolled Oats', '1 lb.', 7.1, 25.3, 907, 5.1, 341, 0, 37.1, 8.9, 64, 0],
  ['White Bread (Enriched)', '1 lb.', 7.9, 15.0, 488, 2.5, 115, 0, 13.8,
   8.5, 126, 0],
```

```
['Whole Wheat Bread', '1 lb.', 9.1, 12.2, 484, 2.7, 125, 0, 13.9, 6.4,
 160, 0],
['Rye Bread', '1 lb.', 9.1, 12.4, 439, 1.1, 82, 0, 9.9, 3, 66, 0],
['Pound Cake', '1 lb.', 24.8, 8.0, 130, 0.4, 31, 18.9, 2.8, 3, 17, 0],
['Soda Crackers', '1 lb.', 15.1, 12.5, 288, 0.5, 50, 0, 0, 0, 0, 0],
['Milk', '1 qt.', 11, 6.1, 310, 10.5, 18, 16.8, 4, 16, 7, 177],
['Evaporated Milk (can)', '14.5 oz.', 6.7, 8.4, 422, 15.1, 9, 26, 3,
 23.5, 11, 60],
['Butter', '1 lb.', 30.8, 10.8, 9, 0.2, 3, 44.2, 0, 0.2, 2, 0],
['Oleomargarine', '1 lb.', 16.1, 20.6, 17, 0.6, 6, 55.8, 0.2, 0, 0, 0],
['Eggs', '1 doz.', 32.6, 2.9, 238, 1.0, 52, 18.6, 2.8, 6.5, 1, 0],
['Cheese (Cheddar)', '1 lb.', 24.2, 7.4, 448, 16.4, 19, 28.1, 0.8, 10.3,
  4, 0],
['Cream', '1/2 pt.', 14.1, 3.5, 49, 1.7, 3, 16.9, 0.6, 2.5, 0, 17],
['Peanut Butter', '1 lb.', 17.9, 15.7, 661, 1.0, 48, 0, 9.6, 8.1, 471,
 0],
['Mayonnaise', '1/2 pt.', 16.7, 8.6, 18, 0.2, 8, 2.7, 0.4, 0.5, 0, 0],
['Crisco', '1 lb.', 20.3, 20.1, 0, 0, 0, 0, 0, 0, 0, 0],
['Lard', '1 lb.', 9.8, 41.7, 0, 0, 0, 0.2, 0, 0.5, 5, 0],
['Sirloin Steak', '1 lb.', 39.6, 2.9, 166, 0.1, 34, 0.2, 2.1, 2.9, 69,
 0],
['Round Steak', '1 lb.', 36.4, 2.2, 214, 0.1, 32, 0.4, 2.5, 2.4, 87, 0],
['Rib Roast', '1 lb.', 29.2, 3.4, 213, 0.1, 33, 0, 0, 2, 0, 0],
['Chuck Roast', '1 lb.', 22.6, 3.6, 309, 0.2, 46, 0.4, 1, 4, 120, 0],
['Plate', '1 lb.', 14.6, 8.5, 404, 0.2, 62, 0, 0.9, 0, 0, 0],
['Liver (Beef)', '1 lb.', 26.8, 2.2, 333, 0.2, 139, 169.2, 6.4, 50.8,
 316, 525],
['Leg of Lamb', '1 lb.', 27.6, 3.1, 245, 0.1, 20, 0, 2.8, 3.9, 86, 0],
['Lamb Chops (Rib)', '1 lb.', 36.6, 3.3, 140, 0.1, 15, 0, 1.7, 2.7, 54,
 0],
['Pork Chops', '1 lb.', 30.7, 3.5, 196, 0.2, 30, 0, 17.4, 2.7, 60, 0],
['Pork Loin Roast', '1 lb.', 24.2, 4.4, 249, 0.3, 37, 0, 18.2, 3.6, 79,
 0],
['Bacon', '1 lb.', 25.6, 10.4, 152, 0.2, 23, 0, 1.8, 1.8, 71, 0],
['Ham, smoked', '1 lb.', 27.4, 6.7, 212, 0.2, 31, 0, 9.9, 3.3, 50, 0],
['Salt Pork', '1 lb.', 16, 18.8, 164, 0.1, 26, 0, 1.4, 1.8, 0, 0],
['Roasting Chicken', '1 lb.', 30.3, 1.8, 184, 0.1, 30, 0.1, 0.9, 1.8,
 68, 46],
['Veal Cutlets', '1 lb.', 42.3, 1.7, 156, 0.1, 24, 0, 1.4, 2.4, 57, 0],
['Salmon, Pink (can)', '16 oz.', 13, 5.8, 705, 6.8, 45, 3.5, 1, 4.9,
 209, 0],
['Apples', '1 lb.', 4.4, 5.8, 27, 0.5, 36, 7.3, 3.6, 2.7, 5, 544],
['Bananas', '1 lb.', 6.1, 4.9, 60, 0.4, 30, 17.4, 2.5, 3.5, 28, 498],
['Lemons', '1 doz.', 26, 1.0, 21, 0.5, 14, 0, 0.5, 0, 4, 952],
['Oranges', '1 doz.', 30.9, 2.2, 40, 1.1, 18, 11.1, 3.6, 1.3, 10, 1998],
['Green Beans', '1 lb.', 7.1, 2.4, 138, 3.7, 80, 69, 4.3, 5.8, 37, 862],
```

```
['Cabbage', '1 lb.', 3.7, 2.6, 125, 4.0, 36, 7.2, 9, 4.5, 26, 5369],
['Carrots', '1 bunch', 4.7, 2.7, 73, 2.8, 43, 188.5, 6.1, 4.3, 89, 608],
['Celery', '1 stalk', 7.3, 0.9, 51, 3.0, 23, 0.9, 1.4, 1.4, 9, 313],
['Lettuce', '1 head', 8.2, 0.4, 27, 1.1, 22, 112.4, 1.8, 3.4, 11, 449],
['Onions', '1 lb.', 3.6, 5.8, 166, 3.8, 59, 16.6, 4.7, 5.9, 21, 1184],
['Potatoes', '15 lb.', 34, 14.3, 336, 1.8, 118, 6.7, 29.4, 7.1, 198, 2522],
['Spinach', '1 lb.', 8.1, 1.1, 106, 0, 138, 918.4, 5.7, 13.8, 33, 2755],
['Sweet Potatoes', '1 lb.', 5.1, 9.6, 138, 2.7, 54, 290.7, 8.4, 5.4, 83, 1912],
['Peaches (can)', 'No. 2 1/2', 16.8, 3.7, 20, 0.4, 10, 21.5, 0.5, 1, 31, 196],
['Pears (can)', 'No. 2 1/2', 20.4, 3.0, 8, 0.3, 8, 0.8, 0.8, 0.8, 5, 81],
['Pineapple (can)', 'No. 2 1/2', 21.3, 2.4, 16, 0.4, 8, 2, 2.8, 0.8, 7, 399],
['Asparagus (can)', 'No. 2', 27.7, 0.4, 33, 0.3, 12, 16.3, 1.4, 2.1, 17, 272],
['Green Beans (can)', 'No. 2', 10, 1.0, 54, 2, 65, 53.9, 1.6, 4.3, 32, 431],
['Pork and Beans (can)', '16 oz.', 7.1, 7.5, 364, 4, 134, 3.5, 8.3, 7.7, 56, 0],
['Corn (can)', 'No. 2', 10.4, 5.2, 136, 0.2, 16, 12, 1.6, 2.7, 42, 218],
['Peas (can)', 'No. 2', 13.8, 2.3, 136, 0.6, 45, 34.9, 4.9, 2.5, 37, 370],
['Tomatoes (can)', 'No. 2', 8.6, 1.3, 63, 0.7, 38, 53.2, 3.4, 2.5, 36, 1253],
['Tomato Soup (can)', '10 1/2 oz.', 7.6, 1.6, 71, 0.6, 43, 57.9, 3.5, 2.4, 67, 862],
['Peaches, Dried', '1 lb.', 15.7, 8.5, 87, 1.7, 173, 86.8, 1.2, 4.3, 55, 57],
['Prunes, Dried', '1 lb.', 9, 12.8, 99, 2.5, 154, 85.7, 3.9, 4.3, 65, 257],
['Raisins, Dried', '15 oz.', 9.4, 13.5, 104, 2.5, 136, 4.5, 6.3, 1.4, 24, 136],
['Peas, Dried', '1 lb.', 7.9, 20.0, 1367, 4.2, 345, 2.9, 28.7, 18.4, 162, 0],
['Lima Beans, Dried', '1 lb.', 8.9, 17.4, 1055, 3.7, 459, 5.1, 26.9, 38.2, 93, 0],
['Navy Beans, Dried', '1 lb.', 5.9, 26.9, 1691, 11.4, 792, 0, 38.4, 24.6, 217, 0],
['Coffee', '1 lb.', 22.4, 0, 0, 0, 0, 0, 4, 5.1, 50, 0],
['Tea', '1/4 lb.', 17.4, 0, 0, 0, 0, 0, 0, 2.3, 42, 0],
['Cocoa', '8 oz.', 8.6, 8.7, 237, 3, 72, 0, 2, 11.9, 40, 0],
['Chocolate', '8 oz.', 16.2, 8.0, 77, 1.3, 39, 0, 0.9, 3.4, 14, 0],
['Sugar', '10 lb.', 51.7, 34.9, 0, 0, 0, 0, 0, 0, 0, 0],
```

```
  ['Corn Syrup', '24 oz.', 13.7, 14.7, 0, 0.5, 74, 0, 0, 0, 5, 0],
  ['Molasses', '18 oz.', 13.6, 9.0, 0, 10.3, 244, 0, 1.9, 7.5, 146, 0],
  ['Strawberry Preserves', '1 lb.', 20.5, 6.4, 11, 0.4, 7, 0.2, 0.2, 0.4, 3, 0],
    # fmt: on
]

 # Instantiate a Glop solver and naming it.
solver = pywraplp.Solver.CreateSolver("GLOP")
if not solver:
    return

# Declare an array to hold our variables.
foods = [solver.NumVar(0.0, solver.infinity(), item[0]) for item in data]

print("Number of variables =", solver.NumVariables())

# Create the constraints, one per nutrient.
constraints = []
for i, nutrient in enumerate(nutrients):
    constraints.append(solver.Constraint(nutrient[1], solver.infinity()))
    for j, item in enumerate(data):
        constraints[i].SetCoefficient(foods[j], item[i + 3])

print("Number of constraints =", solver.NumConstraints())

# Objective function: Minimize the sum of (price-normalized) foods.
objective = solver.Objective()
for food in foods:
    objective.SetCoefficient(food, 1)
objective.SetMinimization()

#Compute the solution
print(f"Solving with {solver.SolverVersion()}")
status = solver.Solve()

# Check that the problem has an optimal solution.
if status != solver.OPTIMAL:
    print("The problem does not have an optimal solution!")
    if status == solver.FEASIBLE:
        print("A potentially suboptimal solution was found.")
    else:
        print("The solver could not solve the problem.")
        exit(1)

# Display the amounts (in dollars) to purchase of each food.
```

```
nutrients_result = [0] * len(nutrients)
print("\nAnnual Foods:")
for i, food in enumerate(foods):
    if food.solution_value() > 0.0:
        print("{}: ${}".format(data[i][0], 365.0 * food.solution_value()))
        for j, _ in enumerate(nutrients):
            nutrients_result[j] += data[i][j + 3] * food.solution_value()
print("\nOptimal annual price: ${:.4f}".format(365.0 * objective.Value()))

print("\nNutrients per day:")
for i, nutrient in enumerate(nutrients):
    print(
        "{}: {:.2f} (min {})".format(nutrient[0], nutrients_result[i],
    nutrient[1]))
```

Bibliography

[1] G. J. Stigler, "The cost of subsistence," *Journal of farm economics*, vol. 27, no. 2, pp. 303–314, 1945.

[2] G. B. Dantzig, "The diet problem," *Interfaces*, vol. 20, no. 4, pp. 43–47, 1990.

[3] S. G. Garille and S. I. Gass, "Stigler's diet problem revisited," *Operations Research*, vol. 49, no. 1, pp. 1–13, 2001.

[4] N. I. of Health. "Nutrient recommendations: Dietary reference intakes (dri)." (), [Online]. Available: `https://ods.od.nih.gov/HealthInformation/nutrientrecommendations.aspx`. (accessed: 20.03.2025).

[5] M. Maillot, F. Vieux, M. J. Amiot, and N. Darmon, "Individual diet modeling translates nutrient recommendations into realistic and individual-specific food choices," *The American journal of clinical nutrition*, vol. 91, no. 2, pp. 421–430, 2010.

[6] N. Darmon, E. L. Ferguson, and A. Briend, "A cost constraint alone has adverse effects on food selection and nutrient density: An analysis of human diets by linear programming," *The Journal of nutrition*, vol. 132, no. 12, pp. 3764–3771, 2002.

[7] L. H. Allen, "Anemia and iron deficiency: Effects on pregnancy outcome123," *The American Journal of Clinical Nutrition*, vol. 71, no. 5, 1280S–1284S, 2000, ISSN: 0002-9165. DOI: `https://doi.org/10.1093/ajcn/71.5.1280s`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0002916523071563`.

[8] R. K. Rude, "Magnesium deficiency: A cause of heterogenous disease in humans," *Journal of bone and mineral Research*, vol. 13, no. 4, pp. 749–758, 1998.

[9] P. Pravina, D. Sayaji, M. Avinash, *et al.*, "Calcium and its role in human body," *International Journal of Research in Pharmaceutical and Biomedical Sciences*, vol. 4, no. 2, pp. 659–668, 2013.

[10] C. Van Dooren, M. Tyszler, G. F. Kramer, and H. Aiking, "Combining low price, low climate impact and high nutritional value in one shopping basket through diet optimization by linear programming," *Sustainability*, vol. 7, no. 9, pp. 12 837–12 855, 2015.

[11] FAO. "Dietary guidelines and sustainability." (), [Online]. Available: https://www.fao.org/nutrition/education/dietary-guidelines/background/sustainable-dietary-guidelines/en/. (accessed: 20.03.2025).

[12] E. Commission. "Life cycle assessment (lca)." (), [Online]. Available: https://eplca.jrc.ec.europa.eu/lifecycleassessment.html. (accessed: 20.03.2025).

[13] F. Barilla. "Double pyramid 2016. a more sustainable future depends on us." (), [Online]. Available: https://www.fondazionebarilla.com/wp-content/uploads/2016/02/doublepyramid2016-more-sustainable-future-depends-on-us.pdf. (accessed: 20.03.2025).

[14] X. Liu, L. Xin, and X. Li, "The global nutrition can be greatly improved with diet optimization," *Resources, Conservation and Recycling*, vol. 202, p. 107 343, 2024.

[15] U. News and W. Report. "Mediterranean diet: Best in 2024." (), [Online]. Available: https://health.usnews.com/best-diet/best-diets-overall. (accessed: 7.05.2025).

[16] R. J. Widmer, A. J. Flammer, L. O. Lerman, and A. Lerman, "The mediterranean diet, its components, and cardiovascular disease," *The American journal of medicine*, vol. 128, no. 3, pp. 229–238, 2015.

[17] N. Darmon, E. Ferguson, and A. Briend, "Linear and nonlinear programming to optimize the nutrient density of a population's diet: An example based on diets of preschool children in rural malawi12," *The American journal of clinical nutrition*, vol. 75, no. 2, pp. 245–253, 2002.

[18] N. Abbaspour, R. Hurrell, and R. Kelishadi, "Review on iron and its importance for human health," *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, vol. 19, no. 2, p. 164, 2014.

[19] P. T. Lieu, M. Heiskala, P. A. Peterson, and Y. Yang, "The roles of iron in health and disease," *Molecular aspects of medicine*, vol. 22, no. 1-2, pp. 1–87, 2001.

[20] R. Hurrell and I. Egli, "Iron bioavailability and dietary reference values," *The American journal of clinical nutrition*, vol. 91, no. 5, 1461S–1467S, 2010.

[21] N. I. of Health. "Iron." (), [Online]. Available: https://ods.od.nih.gov/factsheets/Iron-Consumer/. (accessed: 18.04.2025).

[22] I. Salute. "Food rich in iron combined with other nutrients to favour the absorption." (), [Online]. Available: https://www.issalute.it/index.php/la-salute-dalla-a-alla-z-menu/f/ferro-nella-dieta. (accessed 13.06.2025).

[23] M. Jalava, M. Kummu, M. Porkka, S. Siebert, and O. Varis, "Diet change—a solution to reduce water use?" *Environmental research letters*, vol. 9, no. 7, p. 074 016, 2014.

[24] A. Chaudhary and V. Krishna, “Country-specific sustainable diets using optimization algorithm,” *Environmental science & technology*, vol. 53, no. 13, pp. 7694–7703, 2019.

[25] R. Green, J. Milner, A. D. Dangour, *et al.*, “The potential to reduce greenhouse gas emissions in the uk through healthy and realistic dietary change,” *Climatic Change*, vol. 129, pp. 253–265, 2015.

[26] M. Nordman, A. D. Lassen, A. Stockmarr, P. van ‘t Veer, S. Biesbroek, and E. Trolle, “Exploring healthy and climate-friendly diets for danish adults: An optimization study using quadratic programming,” *Frontiers in Nutrition*, vol. 10, p. 1 158 257, 2023.

[27] M. Layrisse, C. Martinez-Torres, and M. Roche, “Effect of interaction of various foods on iron absorption,” *The American journal of clinical nutrition*, vol. 21, no. 10, pp. 1175–1183, 1968.

[28] M. Donati, D. Menozzi, C. Zighetti, A. Rosi, A. Zinetti, and F. Scazzina, “Towards a sustainable diet combining economic, environmental and nutritional objectives,” *Appetite*, vol. 106, pp. 48–57, 2016.

[29] G. B. Dantzig, “Linear programming and extensions,” in *Linear programming and extensions*, Princeton university press, 2016.