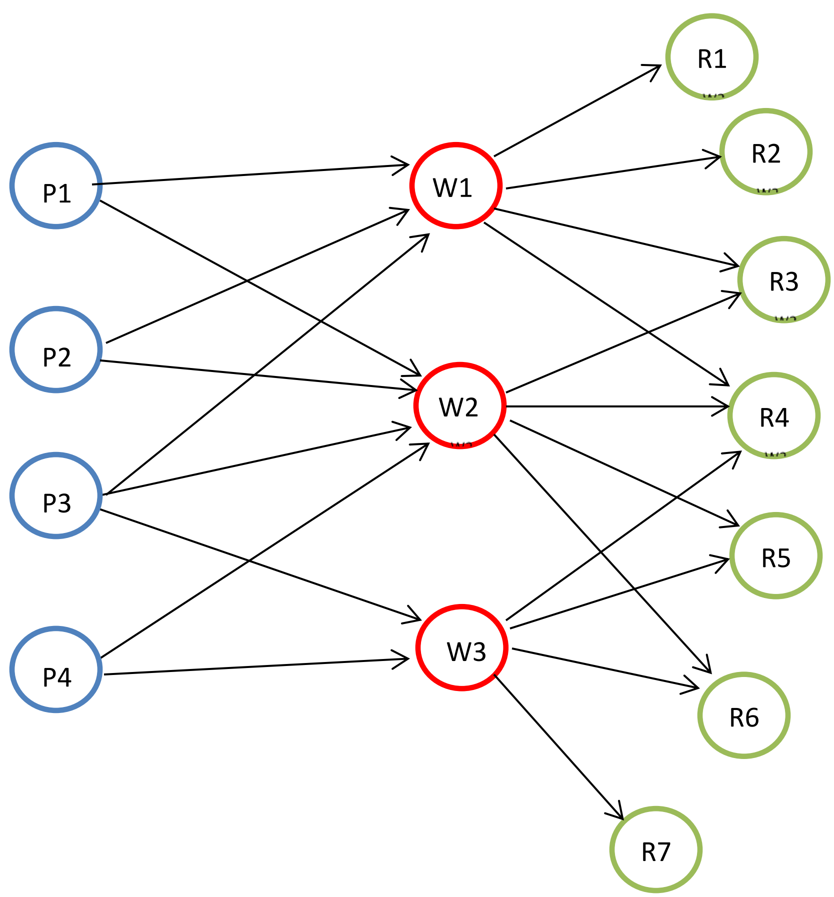
CS 325 Project 3: Linear Programming

For this project, you will model the following problems as linear programs and solve them using a language/linear programming /mathematical software of your choice (problems may be solved using different methods) . Include in your report the solutions to the problems along with any code/files used. This project will only be submitted in Canvas only.

Problem 1: Transshipment Model

This is an extension of the transportation model. There are now intermediate transshipment points added between the sources (plants) and destinations (retailers). Items being shipped from a Plant (p_i) must be shipped to a Warehouse (w_j) before being shipped to the Retailer (r_k). Each Plant will have an associated supply (s_i) and each Retailer will have a demand (d_k). The number of plants is n, number of warehouses is q and the number of retailers is m. The edges (i,j) from plant (p_i)to warehouse (w_j) have costs associated denoted cp(i,j). The edges (j,k) from a warehouse (w_j)to a retailer (r_k) have costs associated denoted cw(j,k).

The graph below shows the transshipment map for a manufacturer of refrigerators. Refrigerators are produced at four plants and then shipped to a warehouse (weekly) before going to the retailer.



CS 325 Project 3: Linear Programming

Below are the costs of shipping from a plant to a warehouse and then a warehouse to a retailer. If it is impossible to ship between the two locations an X is placed in the table.

cost	W1	W2	W3
P1	$10	$15	X
P2	$11	$8	X
P3	$13	$8	$9
P4	X	$14	$8

cost	R1	R2	R3	R4	R5	R6	R7
W1	$5	$6	$7	$10	X	X	X
W2	X	X	$12	$8	$10	$14	X
W3	X	X	X	$14	$12	$12	$6

The tables below give the capacity of each plant (supply) and the demand for each retailer (per week).

	P1	P2	P3	P4
Supply	150	450	250	150

	R1	R2	R3	R4	R5	R6	R7
Demand	100	150	100	200	200	150	100

Part A: Determine the number of refrigerators to be shipped plants to warehouses and then warehouses to retailers to minimize the cost.

i. Formulate the problem as a linear program with an objective function and all constraints.
ii. Determine the optimal solution for the linear program using any software you want. Include a copy of the code/file in the report.
iii. What are the optimal shipping routes and minimum cost.

Part B: Due to old infrastructure Warehouse 2 is going to close eliminating all of the associated routes. What is the optimal solution for this modified model? Is it feasible to ship all the refrigerators to either warehouse 1 or 3 and then to the retailers without using warehouse 2? Why or why not?

Part C: Instead of closing Warehouse 2 management has decide to keep a portion of it open but limit shipments to 100 refrigerators per week. Is this feasible? If so what is the optimal solution when warehouse 2 is limited to 100 refrigerators?

Part D: Formulate a generalized linear programming model for the transshipment problem. Give the objective function and constraints as mathematical formulas.

Problem 2: Modified from DPV 7.16

Veronica the owner of Very Veggie Vegeria is creating a new healthy salad that is low in calories but meets certain nutritional requirements. A salad is any combination of the following ingredients:
Tomato, Lettuce, Spinach, Carrot, Smoked Tofu, Sunflower Seeds, Chickpeas, Oil

Each salad must contain:

- At least 15 grams of protein
- At least 2 and at most 8 grams of fat
- At least 4 grams of carbohydrates
- At most 200 milligrams of sodium
- At least 40% leafy greens by mass.

The nutritional contents of these ingredients (per 100 grams) and cost are

Ingredient	Energy (kcal)	Protein (grams)	Fat (grams)	Carbohydrate (grams)	Sodium (mg)	Cost (100g)
Tomato	21	0.85	0.33	4.64	9.00	$1.00
Lettuce	16	1.62	0.20	2.37	28.00	$0.75
Spinach	40	2.86	0.39	3.63	65.00	$0.50
Carrot	41	0.93	0.24	9.58	69.00	$0.50
Sunflower Seeds	585	23.4	48.7	15.00	3.80	$0.45
Smoked Tofu	120	16.00	5.00	3.00	120.00	$2.15
Chickpeas	164	9.00	2.6	27.0	78.00	$0.95
Oil	884	0	100.00	0	0	$2.00

Part A: Determine the combination of ingredients that minimizes calories but meets all nutritional requirements.

i. Formulate the problem as a linear program with an objective function and all constraints.
ii. Determine the optimal solution for the linear program using any software you want. Include a copy of the code/file in the report.
iii. What is the cost of the low calorie salad?

Part B: Veronica realizes that it is also important to minimize the cost associated with the new salad. Unfortunately some of the ingredients can be expensive. Determine the combination of ingredients that minimizes cost.

i. Formulate the problem as a linear program with an objective function and all constraints.

ii. Determine the optimal solution for the linear program using any software you want. Include a copy of the code/file in the report.
iii. How many calories are in the low cost salad?

Part C: Compare the results from part A and B. Veronica's goal is to create a Very Veggie Salad that is both low calorie and low cost. She would like to sell the salad for $5.00 and still have a profit of at least $3.00. However if she can advertise that the salad has under 250 calories then she may be able to sell more.

i. Suggest some possible ways that she select a combination of ingredients that is "near optimal" for both objectives. This is a type of multi-objective optimization.
ii. What combination of ingredient would you suggest and what is the associated cost and calorie.
iii. Note: There is not one "right" answer. Discuss how you derived your solution.

<u>**Problem 3:**</u> Solving a regression problem by linear programming

Background

In statistics one very common approach for modeling the relationship between a dependent variable y and one or more independent variables x is linear regression. Simple linear regression is used to fit a straight line to statistical data. If (x_1, y_1), (x_2, y_2), …, (x_n, y_n) are data points then regression analysis finds a regression line: $\hat{y} = a_1 x + a_0$, where a_1 (slope) and a_0 (y intercept) are the parameters of the regression line, that "fits" the data. It is important to remember that $\hat{y}$ is a predicted value (for y) calculated from the regression line formula. There will be deviation (error) associated with $\hat{y}$ since in most cases not all the data points fall on a single line. The deviation fr any data point is $(y_i - \hat{y}_i)$, the observed value minus the predicted value.

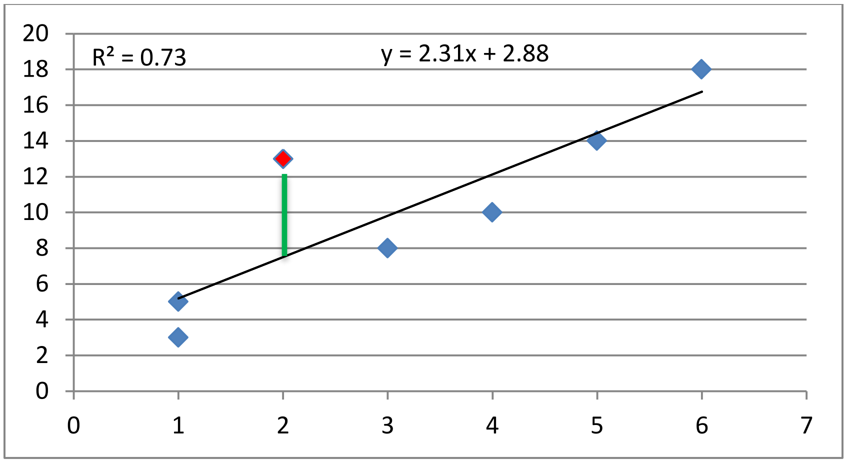
There are several methods that can be used to calculate the regression line. The most common technique is the method of least squares which minimizes the sum of squared deviations. That is

$$\text{minimize } \sum_1^n (y_i - \hat{y}_i)^2$$

There are algebraic formulas for calculating the parameters a_1 and a_0 for the least squares regression line.

Example: Consider the data set (x,y) = { (1,5), (1, 3), (2, 13), (3, 8), (4,10), (5, 14), (6, 18) }. These points were plotted in Excel along with the least squares regression line (trend line) $\hat{y} = 2.31 + 2.88$.

Notice that for x = 2, the line predicts $\hat{y} = 2.31(2) + 2.88 = 7.5$. However the actual value of y corresponding to 2 is 13. The difference between 13 and 7.5 is the deviation (13-7.5) = 5.5. The corresponding squared deviation is $5.5^2 = 30.25$. If we added up the squared deviations for all the data points, the least squares lines would have the smallest sum of all possible lines.



For most data sets the least squares line is a “good” fit. However there are some cases when it is beneficial to calculate the regression line using another method.

Part A: One alternative to the least squares line is the Least Absolute Deviations (LAD). Formulate a linear program whose optimal solution minimizes the sum of the absolute deviations of the data from the line. That is formulate

$$\min \sum_{i=1}^{n} \left| y_i - (a_1 x_i + a_0) \right|$$

as an LP and solve for the a_0 and a_1 that minimize the sum of absolute deviations.

i. Write the linear program for the general problem written as an objective and set of constraints
ii. Use the linear program to find the LAD regression line for the data set

(x,y) = { (1,5), (1, 3), (2, 13), (3, 8), (4,10), (5, 14), (6, 18) }

What was the sum of absolute deviations?

iii. Plot the points and graph your LAD line and the least squares line. Comment.

Part B: Another alternative to the least squares method is to find a line that minimizes of the maximum absolute deviation (MMAD). That is formulate

$$\min \max_i \left| y_i - (a_1 x_i + a_0) \right|$$

as an LP.

i. Write the linear program for the general problem written as an objective and set of constraints
ii. Use the linear program to find the MMAD regression line for the data set

(x,y) = { (1,5), (1, 3), (2, 13), (3, 8), (4,10), (5, 14), (6, 18) }

What was the min of the max absolute deviations?

iii. Plot the points and graph the MMAD line and the least squares line. Compare.
iv. Can you create a data set for which all three methods of regression (least squares, LAD, MMAD) compute the same line.

Part C: Multiple Linear Regression. Generalize the simple linear regression model to allow for two independent variables (x_1 and x_2).

$$\hat{y} = a_2x_2 + a_1x_1 + a_0,$$

The model is called multiple linear not because the result is a line but because all variables are 1^{st} degree. Extend the techniques from Part A to find the least absolute deviations regression equation. Use linear programming to fit a LAD multiple linear regression model to the data set below:

x_1	x_2	y
1	1	5
1	2	9
2	2	12
0	1	3
0	0	0
1	3	11

You do not have to graph this since it is three dimensional.