Logo

Description automatically generated

**College of Professional Studies**

**Northeastern University San Jose**

**MPS Analytics**

**Course: ALY6050: Introduction to Enterprise Analytics**

**Assignment:**

Module 6 Project - Optimizing

**Submitted on:**

April 1, 2023

**Submitted to:**  **Submitted by:**

Prof: AZADEH MOBASHER NIKSHITA RANGANATHAN

# **ABSTRACT**

Nonlinear programming (NLP) is a branch of optimization that deals with finding the optimal values of decision variables in a nonlinear objective function, subject to constraints that can also be nonlinear in nature. In contrast to linear programming, where the objective function and constraints are linear, NLP problems involve nonlinear functions that can be more complex and difficult to solve.

NLP problems arise in many fields, including engineering, finance, economics, and operations research. They can involve a variety of nonlinear functions, such as polynomial, exponential, logarithmic, and trigonometric functions.

Solving NLP problems can be challenging due to the complexity of the objective function and constraints, as well as the potential for multiple local optima. Therefore, it is important to carefully formulate the problem and choose an appropriate solution method to ensure a reliable and efficient solution.

Quadratic programming (QP) is a mathematical optimization technique that deals with finding the minimum of a quadratic objective function subject to linear constraints. In other words, QP is the process of finding the best values for a set of decision variables that minimize a quadratic function, while satisfying a set of constraints.

A quadratic programming problem can be written in the general manner shown below:

minimize: ½ x^T Ax + x^T c

subject to: Ax <= b

x >= 0

where x is a vector of decision variables, Q is a positive semi-definite matrix, c is a vector of constants, A is a matrix of coefficients, and b is a vector of constraint values.

Many various industries, including finance, engineering, physics, and computer science, experience quadratic programming problems. They are often used to solve problems that involve quadratic forms, such as portfolio optimization, machine learning, and control systems.

One of the main advantages of quadratic programming is that it can handle non-linear objective functions that can be approximated by quadratic functions. Additionally, QP can handle a wide range of constraints, including linear, quadratic, and non-linear constraints.

In practice, solving a QP problem can be computationally intensive, especially for large-scale problems. Therefore, efficient algorithms and software libraries have been developed to solve QP problems efficiently.

**INTRODUCTION**

In Part 1 of this assignment, we will be helping the Rockhill Shipping & Transport Company minimize the total transportation cost of hazardous waste from six plants to three waste disposal sites, as they negotiate a new shipping contract with a chemical manufacturing company called Chimotoxic. We will use Linear Programming (LP) to determine the most cost-effective shipping routes, considering the costs of shipping from each plant to each disposal site, the weekly waste generated by each plant, and the maximum capacity of each disposal site. Our goal is to help Rockhill Transport Company optimize their shipping strategy while minimizing the risk of hazardous waste leaks.

**Given tables :**

Table

Description automatically generated

Table

Description automatically generated

A picture containing calendar

Description automatically generated

Table

Description automatically generated

For Part 2, we consider an investment portfolio consisting of six asset types, each with an expected return and covariance matrix. We aim to determine how to allocate $10,000 to the assets to achieve a minimum baseline expected return of 11% while minimizing risk. Additionally, we plot the expected portfolio return versus the minimized risk for eight different baseline return values and discuss any patterns or mathematical relationships that may exist between the two variables. Such insights can help inform our investment decisions and guide us towards building a well-diversified portfolio that balances risk and return.

**Given tables :**

Table

Description automatically generated

Table

Description automatically generated

**ANALYSIS & INTERPRETATION**

**PART-1**

**Direct shipment:**

To begin with, we will create a mathematical model that represents the problem using decision variables, objective function, and constraints.

There are 6 plants and 3 waste disposal sites, resulting in a total of 18 decision variables.

**Decision variables:**

X1 = Number of waste barrels transported from Denver to Orangeburg

X2 = Number of waste barrels transported from Denver to Florence

X3 = Number of waste barrels transported from Denver to Macon

X4 = Number of waste barrels transported from Morganton to Orangeburg

X5 = Number of waste barrels transported from Morganton to Florence

X6 = Number of waste barrels transported from Morganton to Macon

X7 = Number of waste barrels transported from Morrisville to Orangeburg

X8 = Number of waste barrels transported from Morrisville to Florence

X9 = Number of waste barrels transported from Morrisville to Macon

X10 = Number of waste barrels transported from Pineville to Orangeburg

X11 = Number of waste barrels transported from Pineville to Florence

X12 = Number of waste barrels transported from Pineville to Macon

X13 = Number of waste barrels transported from Rockhill to Orangeburg

X14 = Number of waste barrels transported from Rockhill to Florence

X15 = Number of waste barrels transported from Rockhill to Macon

X16 = Number of waste barrels transported from Statesville to Orangeburg

X17 = Number of waste barrels transported from Statesville to Florence

X18 = Number of waste barrels transported from Statesville to Macon

**Objective Function:**

The objective of the problem is to minimize the total transportation cost which is obtained by multiplying the shipping cost per-barrel by number of barrels transported.

**Constraints:**

There are 9 constraints in this problem

* The total number of barrels shipped from Denver must equal 45.
* The total number of barrels shipped from Morganton must equal 26.
* The total number of barrels shipped from Morrisville must equal 42.
* The total number of barrels shipped from Pineville must equal 53.
* The total number of barrels shipped from Rockhill must equal 29.
* The total number of barrels shipped from Statesville must equal 38.
* The total number of barrels shipped to Orangeburg must not exceed 65.
* The total number of barrels shipped to Florence must not exceed 80.
* The total number of barrels shipped to Macon must not exceed 105.

**Table

Description automatically generated**

**Figure 1-** **Optimal solution using Excel Solver (Direct Shipment)**

Using the Solver function in Excel, we can optimize the transportation cost by setting up the objective function and constraints and then allowing Solver to find the optimal solution. In this case, we set up the objective function to minimize the total transportation cost. We then added constraints.

The optimized solution generated by the Solver function indicates that by transporting waste barrels through the optimized routes, the Rockhill Transport company can save a significant amount of money on transportation costs. We obtained the Z value as **$2,988**.

The optimal solution being :

* The number of units of waste barrels to be sent from the plant in Denver to Orangeburg is 36 and to Florence is 9.
* The number of units of waste barrels to be sent from the plant in Morganton to Macon is 26.
* The number of units of waste barrels to be sent from the plant in Morrisville to Macon is 42.
* The number of units of waste barrels to be sent from the plant in Pineville to Florence is 53.
* The number of units of waste barrels to be sent from the plant in Rockhill to Orangeburg is 29.
* The number of units of waste barrels to be sent from the plant in Statesville to Florence is 18.

LP enables the Rockhill Transport company to make informed decisions about how to transport waste barrels in a cost-effective manner while meeting production requirements and waste disposal capacities. This leads to significant cost savings and better resource management. By incorporating LP into its decision-making processes, the company can gain a competitive advantage and stay ahead of its competitors in the market.

**Transshippment:**

Allen is considering whether it would be more cost-effective to use intermediate shipping points, in addition to transporting waste directly from the six plants to the three waste disposal sites. One of the key advantages of using intermediate shipping points is cost savings. By reducing the overall distance and time required to transport the waste, companies can achieve cost savings through more efficient use of resources and economies of scale.

Another advantage of using intermediate shipping points is improved flexibility. By using multiple modes of transportation and transfer points, companies can adjust their transportation routes and modes quickly in response to changes in demand, supply, or other factors. This can help companies to respond more effectively to market changes and customer needs.

We can note that all handling costs for the intermediate points would be handled by Chimotoxic, meaning that Rockhill would only need to cover transportation costs.

**Table

Description automatically generated**

**Figure 2- Model with cost (Transshippment)**

Excel's Solver function requires that all cells in the objective function and constraints have numerical values. In this case, when no transport cost value is provided, Excel's Solver function cannot use the cell for optimization. Therefore, to avoid errors and enable Solver to work properly, a default value of $3000 is assigned to such cells. This value is chosen to be large enough that it does not interfere with the optimization results but small enough to keep the transportation cost as close to the actual cost as possible. This approach ensures that the Solver function can be applied consistently and accurately to the problem, even when some values are missing.

Graphical user interface, application, table, Excel

Description automatically generated

**Figure 3- Optimal solution using Excel Solver (Transshippment)**

Optimal solution:

* Denver should ship 45 barrels to Morganton.
* Morganton should ship 42 barrels to Florence, and 46 barrels to Macon.
* Morrisville should ship 42 barrels to Macon.
* Pineville should ship 36 barrels to Rockhill, and 17 barrels to Morganton.
* Rockhill should ship 65 barrels to Orangeburg.
* Statesville should ship 38 barrels to Florence.
* The total amount of barrels that need to be shipped to Orangeburg is 65, to Florence is 80, and to Macon is 88.
* The minimum shipping cost is **$2,674**.

This approach of using intermediary nodes for transporting waste barrels allows for more efficient and cost-effective shipping. The optimal routes minimize costs while ensuring that the waste reaches its destination in compliance with local regulations. The use of different routes for each shipment also adds to the flexibility of the system, allowing for further optimization of transport costs.

**PART-2**

(i)

Table

Description automatically generated

**Figure 4- Updated Covariance matrix**

A covariance matrix provides a summary of the covariance between all pairs of variables in a data set. The diagonal elements of the covariance matrix contain the variances of each variable, while the off-diagonal elements contain the covariances between each pair of variables. The covariance matrix is often used in multivariate analysis to describe the relationships between multiple variables. It can also be used to calculate other statistical measures, such as correlation coefficients, which normalize the covariance values to fall between -1 and 1.

The covariance matrix is a symmetric, meaning that the values on the upper triangle are the same as those on the lower triangle. So, accordingly we have updated the given table (Figure 4).

Table

Description automatically generated

**Figure 5- Variance and Correlation**

**Variances:**

* Bonds have a variance of 0.001
* High Tech Stocks have a variance of 0.009
* Foreign Stocks have a variance of 0.008
* Call Options have a variance of 0.012
* Put Options have a variance of 0.012
* Gold has a variance of 0.005

**Co-variances:**

* The covariance between bonds and high-tech stocks is 0.0003.
* The covariance between bonds and foreign stocks is -0.0003.
* The covariance between bonds and call options is 0.00035.
* The covariance between bonds and put options is -0.00035.
* The covariance between bonds and gold is 0.0004.
* The covariance between high-tech stocks and foreign stocks is 0.0004.
* The covariance between high-tech stocks and call options is 0.0016.
* The covariance between high-tech stocks and put options is -0.0016.
* The covariance between high-tech stocks and gold is 0.0006.
* The covariance between foreign stocks and call options is 0.0015.
* The covariance between foreign stocks and put options is -0.0055.
* The covariance between foreign stocks and gold is -0.0007.
* The covariance between call options and put options is -0.0005.
* The covariance between call options and gold is 0.0008.
* The covariance between put options and gold is -0.0008.

**Decision variables:**

The decision variables are the weights assigned to each asset (Bonds, High Tech Stocks, Foreign Stocks, Call Options, Put Options, and Gold) in the portfolio. These weights determine the proportion of the portfolio's total value that is allocated to each asset.

**Objective:**

The objective function is a mathematical formula that represents the goal of the optimization problem, which is to minimize the portfolio variance while achieving a target return. In this case, the objective function is a sum product of the calculated variance and covariances (covariance matrix) and the variance and covariances (decision variables), which represent the overall risk of the portfolio.

**Constraints:**

This problem has two constraints. Firstly, the total proportion of investment across all assets must add up to 1,which means that the entire portfolio is fully invested. Secondly, the return percentage obtained by multiplying the weights assigned to each asset by their expected returns should be greater than or equal to 11% (given).

The solution to this optimization problem provides an optimal allocation of assets that can help investors make informed decisions regarding their investment choices.

Table

Description automatically generated

**Figure 6- Calculation for variance**

Table

Description automatically generated

**Figure 7- Variance and Standard deviation**

**Table

Description automatically generated**

**Figure 8- Weights using Excel solver , Investment amount**

**Variance – 0.000736**

**Standard Deviation – 0.027123**

We used Excel Solver to get the optimal values.

**Optimal Solution:**

* Weight of Bonds is 0.18981
* Weight of High-Tech Stocks is 0.10863
* Weight of Foreign Stocks is 0.27082
* Weight of Call Options is 0.04794
* Weight of Put Options is 0.25447
* Weight of Gold is 0.12832

Using the weights, we can calculate the allocation of the portfolio value for each asset as follows:

* Bonds: 0.18981 x $10,000 = $1,898.06
* High Tech-Stocks: 0.10863 x $10,000 = $1,086.31
* Foreign Stocks: 0.27082 x $10,000 = $2,708.28
* Call Options: 0.04794 x $10,000 = $479.43
* Put Options: 0.25447 x $10,000 = $2,544.70
* Gold: 0.12832 x $10,000 = $1,283.23

(ii)

**Table

Description automatically generated**

**Figure 9- Different returns and its respective variance**

The variance values are calculated using the Excel solver method.

Chart, line chart

Description automatically generated

**Figure 10- e vs r Plot**

The resulting data points have been plotted on the scatter plot to visually represent the relationship between the predicted portfolio return and the corresponding minimum risk value (variance).

We have used polynomial trendline to fit a curve to a scatter plot. The equation of the resulting curve is in the form of a polynomial equation. In this case, the equation of the curve is

**y = -5989.2x^2 + 34.587x + 0.0867**

where x represents the minimum risk (r) and y represents the portfolio return (e).

The constant term (0.0867) represents the predicted portfolio return when the minimum risk is zero, which is unlikely to be achievable in practice.

The coefficient of determination, denoted as R2, is a statistical measure that represents the proportion of the variance in the dependent variable (y) that is predictable from the independent variable (x) in a regression model. In this case, the R2 value is **0.9759**, which indicates that the regression is a good fit for the data and that 97.59% of the variability in the predicted portfolio returns can be explained by the variation in the minimum risk values.

In other words, the R2 value suggests that the minimum risk values have a strong influence on the portfolio returns. This implies that as the minimum risk values increase or decrease, the portfolio returns tend to follow.

Table

Description automatically generated

**Figure 11- Quadratic regression model (Fit)**

To find out the mathematical relationship between e and r, we should compute value

**(y - ax^2 + bx + c)^T(y - ax^2 + bx + c)**

where y represents the portfolio return, x represents the min risk, and a, b, and c are the coefficients obtained from the regression model.

**Table

Description automatically generated**The value in this case is 2.537e-05, which is a very small number. This indicates that the quadratic regression model provides an accurate fit to the data, as the differences between the predicted and actual portfolio returns are also very small.

**Figure 12- Quadratic regression model (Optimization)**

Another way to obtain the optimized values of a, b, and c for the quadratic equation is to use Excel solver. In this method, a, b, and c are treated as decision variables, and the objective is to minimize the sum of squared errors, which is represented by the expression (y - ax2 + bx + c)T(y - ax2 + bx + c). After applying the GRG nonlinear method in Excel solver, we obtained the optimized values of a, b, and c, which were very close to the values we obtained using the plot equation. The objective value was also very small, indicating that the model fits the data well.

**CONCLUSION**

In conclusion, the analysis conducted in this assignment provides valuable insights and solutions into two distinct business problems highlighting the importance of data-driven decision-making in business.

The first scenario involved waste disposal for a company and compared the costs of direct shipping versus transshipment. Opting for the transshipment approach would be a wiser choice for the manager as it offers a more cost-effective solution compared to direct shipping from the plants to the waste disposal sites. The direct shipping cost amounts to $2,988, whereas the transshipment plan only costs $2,674 by using various plants and waste sites as intermediary nodes.

This would result in a significant savings of $314 per week for the company which is a considerable amount of money in the long run. These savings could be utilized to invest in other areas of the business, such as expanding the company's operations or improving the overall waste disposal process.

In the second problem, we have provided a portfolio allocation strategy that suggests investing in a combination of bonds, high tech stocks, foreign stocks, call options, put options, and gold. This allocation has been optimized to provide an expected return of 11% and a low variance of 0.000736. Additionally, we have used a quadratic regression analysis to determine the relationship between risk and return for the selected assets, and the results indicate a good fit of the regression equation.

**REFERENCES**

Bluman, A. G. (2018). Elementary Statistics, 10th ed. McGraw Hill.

Kabacoff, R. I. (2011). R in action: Data analysis and graphics with R. Manning Publications Co.