Michael Barrett

Dr. Fatih Celiker

MAT 5770

April 30, 2018

Nonlinear Programming Case Study

Case 13.1 in the tenth edition of *Introduction to Operations Research* is titled Savvy Stock Selection. In this case, Lydia is looking to build a portfolio from six stocks that she expects to outperform the market, all the while maximizing return and minimizing risk. Let stock 1 be BIGBELL (symbol BB), stock 2 be LOTSOFPLACE (symbol LOP), stock 3 be INTERNETLIFE (symbol ILI), stock 4 be HEALTHTOMORROW (symbol HEAL), stock 5 be QUICKY (symbol QUI), and stock 6 be AUTOMOBILE ALLIANCE (symbol AUA). Data is collected and given for each of the stocks, and includes current price, either predicted price or predicted percent increase, historical variance of return, and covariances with each of the other five stocks. Using this data, Lydia will consider several strategies to build her portfolio.

The first strategy that Lydia tries is to ignore the risk of all the investments, and calculate an optimal investment portfolio. This can easily be modeled as a linear programming problem. If we let R(**x**) be the expected return on her portfolio, μ_i be the predicted percent increase for stock i, and x_i be the fraction of her money invested into stock i, so the linear programming problem becomes:

Maximize: $R(\boldsymbol{x})=\sum_{i=1}^{6}\mu_i x_i$

Such that: $\sum_{i=1}^{6}x_i\leq 1$

And: $x_i\geq 0,(i=1,2,\ldots,6)$

This is a simple problem which is maximized by finding i such that $\mu_i = \max(\boldsymbol{\mu})$, then setting $x_i = 1$, and $x_j = 0$ ($j \neq i$). In other words, Lydia puts all her money into the one stock she expects to have the highest percent increase. That stock is ILI, with an expected increase of 100%. Thus, the expected return on her portfolio is also 100%.

To understand the total risk of this portfolio, we calculate its variance using the formula:

$$Var[\boldsymbol{x}] = \sum_{i=1}^{6} \sum_{j=1}^{6} \sigma_{ij} x_i x_j$$

And for our solution, this simplifies to the variance of the percent increase of ILI, which is 0.333. Along with having the greatest expected return, this stock has the greatest variance of the rest, making it the riskiest.

The second strategy that Lydia tries still ignores risk, but limits the fraction of her money invested in any individual stock to 40%. Choosing each x_i for this new portfolio is still a linear programming problem. In fact, it is almost the exact same as the problem from her last portfolio, only with the added constraints:

$$x_i \leq 0.4, \ (i = 1, 2, \ldots, 6)$$

With the added constraints, this remains a simple problem that doesn't require very advanced methods.

The function is maximized at $\mathbf{x^*} = (0, 0, 0.4, 0.4, 0.2, 0)^T$, which implies Lydia puts 40% of her money into ILI (with an expected increase of 100%), 40% of her money into HEAL (with an expected increase of 50%), and 20% of her money into QUI (with an expected increase of 46%). These are the three stocks with the highest expected increases, in order. Thus, the expected return on her portfolio is 69.2%.

Again, we want to understand the risk of this portfolio, so we calculate the variance. The variance is 0.06068. This is much lower than the variance of the previous portfolio, even though the two riskiest stocks make up 80% of this portfolio. In fact, the variance of this portfolio is lower than the individual variances of 5 out of the 6 stocks, including each of the ones used.

The third strategy that Lydia tries accounts for the risk of her investment. She wants to minimize the risk (variance) of her portfolio such that her portfolio has an expected return greater than or equal to some constant. Let that constant be represented by r. Thus, Var[**x**] becomes the objective function, which is not a linear function of **x**. To solve this problem, we must introduce nonlinear programming. Since this function contains terms $x_i x_j$ ($i, j = 1, 2,\ldots, 6$), this can be modeled as a quadratic programming problem, as such:

Minimize: $$Var[\boldsymbol{x}]=\sum_{i=1}^{6}\sum_{j=1}^{6}\sigma_{ij}x_i x_j$$

Such that: $$\sum_{i=1}^{6}\mu_i x_i \geq r$$

$$\sum_{i=1}^{6}x_i \leq 1$$

And: $$x_i \geq 0, (i=1,2,\ldots,6)$$

The objective function can be re-written as a Var[**x**] = $\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x}$, where **Q** is a matrix such that $q_{ij} = \sigma_{ij}$, for $i, j = 1, 2,\ldots, 6$, and $i \leq j$, and $q_{ij} = 0$ for $i, j = 1, 2,\ldots, 6$, and $i > j$. Thus, we have the matrix:

$$\boldsymbol{Q}=\begin{bmatrix} 0.032 & 0.005 & 0.03 & -0.031 & -0.027 & 0.01 \\ 0 & 0.1 & 0.085 & -0.07 & -0.05 & 0.02 \\ 0 & 0 & 0.333 & -0.11 & -0.02 & 0.042 \\ 0 & 0 & 0 & 0.125 & 0.05 & -0.06 \\ 0 & 0 & 0 & 0 & 0.065 & -0.02 \\ 0 & 0 & 0 & 0 & 0 & 0.08 \end{bmatrix}$$

Furthermore, the constraints can be re-written as $\mathbf{Ax} \leq \mathbf{b}$, where $a_{11}, a_{12},\ldots, a_{16}$ are $-\mu_1, -\mu_2,\ldots, -\mu_6$, and where $a_{21}, a_{22},\ldots, a_{26}$ all equal 1. For the vector **b**, b_1 is $-r$, and b_2 is 1. Thus we have:

$$\boldsymbol{A}=\begin{bmatrix} -0.20 & -0.42 & -1.00 & -0.50 & -0.46 & -0.30 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} -r \\ 1 \end{bmatrix}$$

Since this problem is subject to more constraints than just the non-negativity constraints, we cannot find an optimal solution by simply creating a system of equations by setting each partial derivative of the objective function equal to zero, then evaluating the system. Also, we cannot utilize either the gradient search procedure nor the general version of Newton's method, both of which are numerical search procedures which are available to solve an unconstrained problem numerically.

Since it is a constrained problem, with a concave objective function and convex constraints, the simplest way to solve the problem is using the Karush-Kuhn-Tucker conditions (or KKT conditions) for optimality. The basic result of these conditions is the following theorem:

> Assume that $f(\mathbf{x})$, $g_1(\mathbf{x})$, $g_2(x)$,…, $g_m(\mathbf{x})$ are differentiable functions satisfying certain regularity conditions. Then
>
> $$\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)^T$$
>
> can be an optimal solution for the nonlinear programming problem only if there exist m numbers $u_1, u_2, \ldots, u_m$ such that all the following KKT conditions are satisfied:

1. $\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \leq 0$, at $\mathbf{x} = \mathbf{x}^*$, for $j = 1, 2, \ldots, n$.

2. $x_j^* \left(\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \right) = 0$, at $\mathbf{x} = \mathbf{x}^*$, for $j = 1, 2, \ldots, n$.

3. $g_i(\mathbf{x}^*) - b_i \leq 0$, for $i = 1, 2, \ldots, m$.

4. $u_i [g_i(\mathbf{x}^*) - b_i] = 0$, for $i = 1, 2, \ldots, m$.

5. $x_j^* \geq 0$, for $j = 1, 2, \ldots, n$.

6. $u_i \geq 0$, for $i = 1, 2, \ldots, m$.

Applying this theorem to Lydia's portfolio problem, we can see that $f(\mathbf{x}) = \mathrm{Var}[\mathbf{x}] = \mathbf{x}^T\mathbf{Q}\mathbf{x}$, and that $\mathbf{g}(\mathbf{x}) = \mathbf{A}\mathbf{x}$. Clearly, $\mathbf{b} = \mathbf{b}$. We can introduce non-negative slack variables $\mathbf{y}$ and $\mathbf{v}$ to express the inequalities in conditions 1 and 3 as equations, as such:

1. $\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} + y_j = 0$, for $j = 1, 2,\ldots, n$.

3. $g_i(\mathbf{x}) - b_i + v_i = 0$, for $i = 1, 2,\ldots, m$.

Condition 2 can now be expressed as requiring either x_j or y_j equal zero, and condition 4 can be expressed as requiring either u_i or v_i equal zero, as such:

2. $x_j y_j = 0$ for $j = 1, 2,\ldots, n$.

4. $u_i v_i = 0$ for $i = 1, 2,\ldots, m$.

Finally, conditions 2 and 4 can be combined into one constraint, called the complementarity constraint:

$$\sum_{j=1}^{n} x_j y_j + \sum_{i=1}^{m} u_i v_i = 0$$

These KKT conditions are now linear programming constraints, with the exception of the complementarity constraint, and can all be expressed in matrix/vector notation:

$$\mathbf{Qx} + \mathbf{A}^T\mathbf{u} - \mathbf{y} = \mathbf{0},$$

$$\mathbf{Ax} + \mathbf{v} = \mathbf{b},$$

$$\mathbf{x} \geq \mathbf{0},\ \mathbf{u} \geq \mathbf{0},\ \mathbf{y} \geq \mathbf{0},\ \mathbf{v} \geq \mathbf{0},$$

$$\mathbf{x}^T\mathbf{y} + \mathbf{u}^T\mathbf{v} = 0,$$

A corollary to the given theorem states that if the problem is a convex programming problem and $f(\mathbf{x})$, $g_1(\mathbf{x})$, $g_2(x)$,…, $g_m(\mathbf{x})$ all satisfy the regularity conditions, then $\mathbf{x^*} = (x_1^*, x_2^*,\ldots, x_n^*)^T$ is an optimal solution if and only if all the conditions of the theorem are satisfied. This corollary implies that **x** is optimal if and only if there exist values of **y**, **u**, and **v** such that all four vectors satisfy the all the above constraints. This implies that we can reduce the original problem to simply finding a feasible solution subject to these constraints.

Now that we have simplified the problem to finding a feasible solution subject to the above constraints, we can return to building Lydia's portfolio. She wants to minimize Var[**x**] such that R(**x**) is greater than or equal to *r*. She wants to test 3 values for *r*, which are 25%, 35%, and 40%. When we plug these three values into b_1 (in decimal form, to be consistent with **A**), we can solve for an optimal **x***. The linear constraints can be expressed as:

$$\begin{bmatrix} \boldsymbol{Q} \\ \boldsymbol{A} \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} \boldsymbol{A}^T \\ \boldsymbol{0}_{2,2} \end{bmatrix} \boldsymbol{u} - \begin{bmatrix} \boldsymbol{I} \\ \boldsymbol{0}_{2,6} \end{bmatrix} \boldsymbol{y} + \begin{bmatrix} \boldsymbol{0}_{6,2} \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{v} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b} \end{bmatrix} \quad ,$$

Which is equivalent to:

$$\begin{bmatrix} 0.032 & 0.005 & 0.03 & -0.031 & -0.027 & 0.01 & -0.20 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.085 & -0.07 & -0.05 & 0.02 & -0.42 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.333 & -0.11 & -0.02 & 0.042 & -1.00 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.125 & 0.05 & -0.06 & -0.50 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.065 & -0.02 & -0.46 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.08 & -0.30 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -0.20 & -0.42 & -1.00 & -0.50 & -0.46 & -0.30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_6 \\ u_1 \\ u_2 \\ y_1 \\ \vdots \\ y_6 \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -r \\ 1 \end{bmatrix}$$

I chose to evaluate this problem using computer software, as solving the problem by hand has proven difficult (specifically with regards to the complementarity constraint). The software I chose to use is GNU Octave. I evaluated the problem using the "qp" function. I entered an initial guess for **x** (**0**), the matrix **Q**, the linear coefficients for **x** (**0**), the constraint matrix **A**, the vector **b**, the lower bound for **x** (**0**), and the upper bound for **x** $(1, 1, 1, 1, 1, 1)^T$.

For r = 0.25, I find a solution of:

x* = $(0.7325, 0.0042, 0.0000, 0.0383, 0.0944, 0.1307)^T$, R(**x***) = 0.2500, Var[**x***] = 0.0172.

For r = 0.35, I find a solution of:

x* = $(0.3575, 0.1520, 0.0077, 0.1379, 0.2158, 0.1292)^T$, R(**x***) = 0.3500, Var[**x***] = 0.0075.

For r = 0.40, I find a solution of:

x* = $(0.2669, 0.1621, 0.0532, 0.1683, 0.2270, 0.1225)^T$, R(**x***) = 0.4000, Var[**x***] = 0.0084.

Having checked the inputs for errors, and checked my answers by running the problem in other solvers, I have no explanation for how the variance of the first solution is the greatest, when it has the least restrictive constraints. What I did expect is that every x_i either increases or decreases as we increase r. This implies that more aggressive portfolios will favor some stocks over others. Problems I see with this method mainly have to do with how difficult it is to obtain an answer analytically. This means that we need to resort to numerical methods to approximate a solution. Nevertheless, using good computer software, we can get solutions which are very close to optimal.