Portfolio Optimization

Justin Wyss-Gallifent

September 19, 2023

14.1   Introduction

Suppose we are investing in the stock market and we are examining several companies. Each company has stock which we can buy and sell. Suppose for each stock we know both the average weekly rate of return for the stock, as well as the variance in that rate of return.

At this point if we had to invest we might pick the stock with the highest return (if we were willing to take the corresponding risk) or with the lowest variance (if we were risk-averse). Or we might invest some of our money in one stock and the rest in another.

However to complicate things suppose the performance of the companies affect one another. Perhaps when one company's stock goes up, another goes down, and perhaps the third goes up but not by as much.

We will address two questions:

(I) Suppose we have no interest in the return but simply want to invest our money to minimize the risk. How should we do this?

(II) Suppose we want a particular rate of return. How can we achieve this while minimizing the risk?

14.2   A Brief Review of Statistics

14.2.1   Random Variables

Definition 14.2.1.1. A *random variable* is a variable whose outcomes follow some unknown (perhaps random) pattern. Typically random variables are denoted by capital letters such as X, Y, X_1, etc.

Example 14.1. If a fair die is rolled and the result is assigned to the random variable X. Then $X \in \{1, 2, 3, 4, 5, 6\}$.

The word "random" is slightly inaccurate.

Example 14.2. The daily percentage return on a stock, can be treated as a random variable. It's not really random (it's affected by company performance, investor behavior, etc.) but it is unpredictable and we can ask questions about its value.

14.2.2   Expected Value

Definition 14.2.2.1. The *expected value* of a random variable X, denoted $E(X)$ (or sometimes $\mu(X)$ or μ_X or just μ when it's clear), is the long-term

average of that variable, meaning if the variable took on values over and over again forever what the average would be.

If each outcome is equally likely then the expected value is simply the average.

Example 14.3. A die is rolled and the result is assigned to the random variable X. Then $E(X) = (1 + 2 + 3 + 4 + 5 + 6)/6 = 3.5$.

The expected value of the daily return of a stock can be approximated simply be averaging the returns over a wide sample of days, perhaps the last 30 or 60 or 90 days. Of course this isn't perfect since the stock might change in behavior but it gives us something to work with.

The expected value isn't really a time-related thing since we can imagine rolling infinitely many dice all at once instead of rolling one die infinitely many times but often it is time-related or can be thought of that way.

14.2.3   Variance

Definition 14.2.3.1. The *variance* of a random variable X, denoted $Var(X)$, is the long-term average of the square of the difference between the variable and its long-term average. In other words:

$$Var(X) = E\left((X - E(X))^2\right)$$

Note: The square root of variance is the *standard deviation* and is denoted $\sigma(X)$ or σ_X or just σ. Consequently sometimes the variance is denoted $\sigma(X)^2$ or σ_X^2 or just σ^2 when it's clear.

If each outcome is equally likely then so is each $X - E(X)$ so then the variance is simply the average of $(X - E(X))^2$ taken over all possible X.

Example 14.4. The variance for the die we rolled is:

$$
\begin{aligned}
Var(X) &= \frac{(1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2}{6} \\
&= \frac{17.5}{6} \\
&= \frac{35}{12} \\
&\approx 2.9167
\end{aligned}
$$

Basically if a random variable X has a high variance this means that the values it takes on can be more spread out away from the average.

As far as stocks are concerned this can be understood as a measurement of risk.

14.2.4 Covariance

Suppose we have two random variables which take on values together so instead of looking at what each does independently we look at what they do in pairs. We might want to measure the connection between them.

Example 14.5. Suppose $X \in \{1, 2, 3\}$ and $Y \in \{1, 2, 3\}$ and we find that in real world data we see the three pairs $(1, 1)$, $(2, 2)$ and $(3, 3)$. We notice that smaller X values correspond to smaller Y values and larger X values correspond to larger Y values.

Definition 14.2.4.1. The *covariance* of two random variables X and Y, denoted $Cov(X, Y)$, is the long-term average of the product of the differences between the two variables and their long-term averages. In other words:

$$Cov(X, Y) = E\left((X - E(X))(Y - E(Y))\right)$$

Note: Sometimes the covariance is denoted $\sigma(X, Y)^2$ or σ_{XY}^2 in order to match the notation of variance, even though neither $\sigma(X, Y)$ or σ_{XY} is a meaningful value other than square root of covariance.

If all outcomes of the pairs that appear are equally likely then the covariance is just the average of $(X - E(X))(Y - E(Y))$ over all possible pairs.

Example 14.6. In our example above we have $E(X) = 2$ and $E(Y) = 2$ and so we average $(X - 2)(Y - 2)$ over our three pairs:

$$Cov(X, Y) = \frac{(1 - 2)(1 - 2) + (2 - 2)(2 - 2) + (3 - 2)(3 - 2)}{3} = \frac{2}{3}$$

Example 14.7. If the three pairs that we observed had been $(1, 3)$, $(2, 2)$ and $(3, 1)$ then we find:

$$Cov(X, Y) = \frac{(1 - 2)(3 - 2) + (2 - 2)(2 - 2) + (1 - 2)(3 - 2)}{3} = -\frac{2}{3}$$

which makes sense because now larger X values correspond to smaller Y values and vice-versa.

A positive covariance means that as one of X and Y increases, so does the other, whereas a negative covariance means that as one of X and Y increases, the other decreases. A covariance of zero means that a change in one has no impact on the other.

Notice that this does not imply that either of them directly influences the other but that they tend to act together, for whatever reason.

This is clear from the formula since if X and Y tend to be both larger or smaller than their individual expected values at the same time then $(X - E(X))(Y - E(Y))$ will tend to be positive, giving a positive expected value for that product, whereas if one tends to be larger while the other is smaller then that product will tend to be negative, giving an negative expected value for that product.

All the variances and covariances can all be placed in a handy matrix:

Definition 14.2.4.2. Given random variables $X_1, ..., X_n$ we define the *covariance matrix* :

$$\Sigma = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & \ldots & Cov(X_1, X_n) \\ Cov(X_1, X_2) & Var(X_2) & \ldots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_1, X_n) & Cov(X_2, X_n) & \ldots & Var(X_n) \end{bmatrix}$$

14.2.5 Associated Theorems in Brief

Here is a brief summary of the theorems we will need:

Theorem 14.2.5.1. We have

$$E(a_1 X_1 + ... + a_n X_n) = a_1 E(X_1) + ... + a_n E(X_n)$$

Theorem 14.2.5.2. Variance may also be calculated with the formula

$$Var(X) = E(X^2) - E(X)^2$$

Theorem 14.2.5.3. The variance of a variable is just its covariance with itself

$$Var(X) = Cov(X, X)$$

Theorem 14.2.5.4. Covariance may also be calculated with the formula

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Theorem 14.2.5.5. We have

$$Cov(aX, bY) = abCov(X, Y)$$

Theorem 14.2.5.6. We have

$$Var\left(\sum_i X_i\right) = \sum_{i,j} Cov(X_i, X_j)$$

Theorem 14.2.5.7. We have

$$Var\left(\sum_i a_i X_i\right) = \sum_{i,j} a_i a_j Cov(X_i, X_j)$$

Theorem 14.2.5.8. We have

$$Var\left(\sum_i a_i X_i\right) = \bar{a}^T \Sigma \bar{a}$$

14.2.6   Associated Theorems with Proofs

Theorem 14.2.6.1. We have

$$E(a_1 X_1 + ... + a_n X_n) = a_1 E(X_1) + ... + a_n E(X_n)$$

Proof. The proof of this is fairly obvious for simple random variables and not so obvious for more complicated random variables. For a simple case suppose we have two random variables X, Y with $X \in \{x_1, x_2, ..., x_n\}$ and $Y \in \{y_1, y_2, ..., y_m\}$ with each outcome being equally likely. Then for constants a, b we have $aX \in \{ax_1, ax_2, ..., ax_n\}$ and $bY \in \{by_1, by_2, ..., by_m\}$ and $aX + bY \in \{ax_i + by_j \,|\, 1 \leq i \leq n, 1 \leq j \leq m\}$. The probability of each $ax_i + bj_j$ is then $(1/n)(1/m)$ and so:

$$E(aX + bY) = \sum_{i,j} \frac{1}{nm}(ax_i + by_j)$$

$$= a\sum_{i,j} \frac{1}{nm} x_i + b\sum_{i,j} \frac{1}{nm} y_i$$

$$= a\sum_i \frac{1}{n} x_i + b\sum_j \frac{1}{m} y_i$$

$$= aE(X) + bE(y)$$

$\square$

Theorem 14.2.6.2. Variance may also be calculated with the formula

$$Var(X) = E(X^2) - E(X)^2$$

Proof. Obvious after the next two proofs. □

Theorem 14.2.6.3. The variance of a variable is just its covariance with itself

$$Var(X) = Cov(X, X)$$

Proof. Well, $Cov(X, X) = E((X - E(X))(X - E(X))) = Var(X)$. □

Theorem 14.2.6.4. Covariance may also be calculated with the formula

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Proof. We have:

$$\begin{aligned}
Cov(X, Y) &= E\left((X - E(X))(Y - E(Y))\right) \\
&= E\left(XY - XE(Y) - YE(X) - E(X)E(Y)\right) \\
&= E(XY) - E(X)E(Y) - E(Y)E(X) - E(X)E(Y) \\
&= E(XY) - E(X)E(Y)
\end{aligned}$$

Notice that line 3 follow from line 2 by the linearity of E, noting that $E(Y)$ and $E(X)$ are constants. □

Theorem 14.2.6.5. We have

$$Cov(aX, bY) = abCov(X, Y)$$

Proof. We have:

$$\begin{aligned}
Cov(aX, bY) &= E(aXbY) - E(aX)E(bY) \\
&= abE(XY) - abE(X)E(Y) \\
&= ab\, Cov(X, Y)
\end{aligned}$$

□

Theorem 14.2.6.6. We have

$$Var\left(\sum_i X_i\right) = \sum_{i,j} Cov(X_i, X_j)$$

Proof. We have:

$$Var\left(\sum_i X_i\right) = E\left(\left(\sum_i X_i\right)^2\right) - E\left(\sum_i X_i\right)^2$$

$$= E\left(\sum_{i,j} X_i X_j\right) - \left(\sum_{i,j} E(X_i)\right)^2$$

$$= \sum_{i,j} E(X_i X_j) - \sum_{i,j} E(X_i)E(X_j)$$

$$= \sum_{i,j} [E(X_i X_j) - E(X_i)E(X_j)]$$

$$= \sum_{i,j} Cov(X_i, X_j)$$

$\square$

Theorem 14.2.6.7. We have

$$Var\left(\sum_i a_i X_i\right) = \sum_{i,j} a_i a_j Cov(X_i, X_j)$$

Proof. Follows from previous theorems. $\square$

Theorem 14.2.6.8. We have

$$Var\left(\sum_i a_i X_i\right) = \bar{a}^T \Sigma \bar{a}$$

Proof. Brute force matrix calculation. $\square$

14.3   A Brief Review of Lagrange Multipliers

Typically Lagrange Multipliers are approached this way for the two variable case:

We have a function $f(x, y)$ for which we wish to find a minimum or maximum (which we know exists) subject to a constraint $g(x, y) = 0$.

We solve the system of equations:

$$f_x = \lambda g_x$$
$$f_y = \lambda g_y$$
$$g(x, y) = 0$$

Then we test f at each resulting (x, y) and choose whichever (minimum or maximum) we wanted.

Another classic way to approach this is to define the *Lagrange Function*:

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

And then solving the original system is exactly the same as solving:

$$L_x = 0$$
$$L_y = 0$$
$$L_\lambda = 0$$

The only difference being that we use $-\lambda$ instead of λ but that doesn't matter.

The advantage of this second approach is that it generalizes easily not just to more variables but to additional constraints.

Theorem 14.3.0.1. General Method of Lagrange Mulipliers: Suppose we wish to maximize or minimize the function $f(\bar{x})$ subject to the set of constraints $g_1(\bar{x}) = ... = g_k(\bar{x}) = 0$. If we define the Lagrange function:

$$L(\bar{x}, \lambda_1, ..., \lambda_k) = f(\bar{x}) + \lambda_1 g_1(\bar{x}) + ... + \lambda_k g_k(\bar{x})$$

then the maximum and minimum will occur at a solution to the system

$$L_{x_1} = 0$$
$$\vdots \quad \vdots$$
$$L_{x_n} = 0$$
$$L_{\lambda_1} = 0$$
$$\vdots \quad \vdots$$
$$L_{\lambda_k} = 0$$

Proof. Omitted. $\qquad\square$

Notice that the system can be succinctly written as $\nabla L = \bar{0}$ where ∇L is the gradient.

Example 14.8. To find the assumed minimum value of the function $f(x, y, z) = x^2 + x + y^2 + z^2 + 2z$ subject to the constraints $x + y - z = 2$ and $2x - y + z = 0$ we set $g_1(x, y, z) = x + y - z - 2$ and $g_2(x, y, z) = 2x - y + z$ and then we set:

$$L(x, y, z, \lambda_1, \lambda_2) = x^2 + x + y^2 + z^2 + 2z + \lambda_1(x + y - z - 2) + \lambda_2(2x - y + z)$$

and solve the system $\nabla L = \bar{0}$ which is:

$$2x + 1 + \lambda_1 + 2\lambda_2 = 0$$
$$2y + \lambda_1 - \lambda_2 = 0$$
$$2z + 2 - \lambda_1 + \lambda_2 = 0$$
$$x + y - z - 2 = 0$$
$$2x - y + z = 0$$

If we solve these we get $(x, y, z, \lambda_1, \lambda_2) = \left(\frac{2}{3}, \frac{1}{6}, -\frac{7}{6}, -1, -\frac{2}{3}\right)$

Since there's only one answer it must be our minimum and so the minimum occurs at $\left(\frac{2}{3}, \frac{1}{6}, -\frac{7}{6}\right)$ and equals

$$f\left(\frac{2}{3}, \frac{1}{6}, -\frac{7}{6}\right) = \frac{1}{6}$$

14.4 Portfolio Optimization

14.4.1 Introduction

Let's begin with a simple examle which will help us see why linear algebra is useful.

We are assuming the *Constant Expected Return (CER) Model*. This is a standard model used in asset return theory which makes many assumptions about the behavior of the returns for the sake of simplicity.

For our purposes the only assumptions we need to note is that the expected value and variance of a return is constant over time and therefore we can get reasonable values by collecting a large sample.

So now let's suppose we have three companies each with a stock. Each stock has a return which is a random variable so we will denote these R_1, R_2 and R_3.

Suppose then we collect data over a long period of time and we calculate the monthly expected values which we denote $\mu_1 = E(R_1)$, $\mu_2 = E(R_2)$, $\mu_3 = E(R_3)$, we calculate the monthly variances which we denote σ_1^2, σ_2^2, σ_3^2, and we calculate the monthly covariances which we denote σ_{12}^2, σ_{13}^2, σ_{23}^2.

Note we can also put the monthly expected values in a vector $\bar{\mu}$.

Suppose we have a total amount to invest and we wish to split this amount between the three stocks. Rather than deal with a specific total amount we will just look at the proportion of our amount which should be invested in each stock.

Call those proportions x_1, x_2, x_3 and so we have $x_1 + x_2 + x_3 = 1$. Note we can also put these proportions in a vector $\bar{x}$.

The return for our portfolio will then have random variable:

$$R = x_1 R_1 + x_2 R_2 + x_3 R_3$$

The expected return for our portfolio, will then be:

$$E(R) = x_1 \mu_1 + x_2 \mu_2 + x_3 \mu_3 = \bar{\mu}^T \bar{x} = \bar{x}^T \bar{\mu}$$

The variance for our portfolio will then be:

$$
\begin{aligned}
Var(R) &= Var(x_1 R_1 + x_2 R_2 + x_3 R_3) \\
&= \sum_{i,j} Cov(x_i R_i, x_j R_j) \\
&= \sum_{i,j} x_i x_j Cov(R_i, R_j) \\
&= x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + x_3^2 \sigma_3^2 + 2 x_1 x_2 \sigma_{12}^2 + 2 x_1 x_3 \sigma_{13}^2 + 2 x_2 x_3 \sigma_{23}^2 \\
&= \bar{x}^T \Sigma \bar{x}
\end{aligned}
$$

14.4.2 Global Minimum Variance Portfolio

Suppose we wish to invest our portfolio in a way that minimizes the variance without regard for the return. Basically you've got to put your money somewhere but you want as little risk as possible.

What this means is that we wish to choose x_1, x_2, x_3 which minimizes:

$$Var(R) = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + 2x_1x_2\sigma_{12}^2 + 2x_1x_3\sigma_{13}^2 + 2x_2x_3\sigma_{23}^2$$

subject to the constraint:

$$x_1 + x_2 + x_3 = 1$$

We define the Lagrange function:

$$
\begin{aligned}
L &= Var(X) + \lambda(x_1 + x_2 + x_3 - 1) \\
&= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + 2x_1x_2\sigma_{12}^2 + 2x_1x_3\sigma_{13}^2 + 2x_2x_3\sigma_{23}^2 \\
&\quad + \lambda(x_1 + x_2 + x_3 - 1)
\end{aligned}
$$

and solve the system $\nabla L = \bar{0}$ which is:

$$
\begin{aligned}
2\sigma_1^2 x_1 + 2\sigma_{12}^2 x_2 + 2\sigma_{13}^2 x_3 + \lambda &= 0 \\
2\sigma_{12}^2 x_1 + 2\sigma_2^2 x_2 + 2\sigma_{23}^2 x_3 + \lambda &= 0 \\
2\sigma_{13}^2 x_1 + 2\sigma_{23}^2 x_2 + 2\sigma_3^2 x_3 + \lambda &= 0 \\
x_1 + x_2 + x_3 - 1 &= 0
\end{aligned}
$$

Notice that this can be rewritten as a matrix equation:

$$
\begin{bmatrix}
2\sigma_1^2 & 2\sigma_{12}^2 & 2\sigma_{13}^2 & 1 \\
2\sigma_{12}^2 & 2\sigma_2^2 & 2\sigma_{23}^2 & 1 \\
2\sigma_{13}^2 & 2\sigma_{23}^2 & 2\sigma_3^2 & 1 \\
1 & 1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ \lambda
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 1
\end{bmatrix}
$$

which can be rewritten much more simply using the covariance matrix as:

$$
\begin{bmatrix}
2\Sigma & \bar{1} \\
\bar{1}^T & 0
\end{bmatrix}
\begin{bmatrix}
\bar{x} \\ \lambda
\end{bmatrix}
=
\begin{bmatrix}
\bar{0} \\ 1
\end{bmatrix}
$$

which is especially nice since it generalizes to more than three options and can be solved via a matrix inverse:

$$
\begin{bmatrix}
\bar{x} \\ \lambda
\end{bmatrix}
=
\begin{bmatrix}
2\Sigma & \bar{1} \\
\bar{1}^T & 0
\end{bmatrix}^{-1}
\begin{bmatrix}
\bar{0} \\ 1
\end{bmatrix}
$$

The questions as to why this matrix is invertible and why this yields a solution of minimum variance will be left for later.

Observe that it's not necessary to know $\bar{\mu}$ to find the Global Minimum Variance Portfolio but it is necessary if we wish to know the expected return.

Example 14.9. Suppose three companies have expected monthly earnings:

$$\bar{\mu} = \begin{bmatrix} 0.0385 \\ 0.0021 \\ 0.0202 \end{bmatrix}$$

And they have variances and covariances given by:

$$\Sigma = \begin{bmatrix} 0.0110 & 0.0015 & 0.0008 \\ 0.0015 & 0.0121 & 0.0016 \\ 0.0008 & 0.0016 & 0.0218 \end{bmatrix}$$

Then the global minimum variance portfolio can be found via:

$$\begin{bmatrix} \bar{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} 2\Sigma & \bar{1} \\ \bar{1}^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{0} \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.0220 & 0.0030 & 0.0016 & 1.0000 \\ 0.0030 & 0.0242 & 0.0032 & 1.0000 \\ 0.0016 & 0.0032 & 0.0436 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.4271 \\ 0.3672 \\ 0.2057 \\ -0.0108 \end{bmatrix}$$

Thus to minimize our risk we should set:

$$\bar{x} = \begin{bmatrix} 0.4271 \\ 0.3672 \\ 0.2057 \end{bmatrix}$$

meaning proportionally speaking we should invest 0.4271 of our portfolio in Company 1, 0.3672 of our portfolio in Company 2, and 0.2057 of our portfolio in Company 3.

The variance in this case would be:

$$Var(R) = \bar{x}^T \Sigma \bar{x} = 0.0054$$

13

and this is the minimum variance we can ever achieve. Notice that it is significantly lower than any of the individual company variances (the diagonal entries in Σ) because by spreading out our portfolio we have taken advantage of the covariances.

The expected return would be:

$$E(R) = \bar{x}^T \bar{\mu} = 0.0214$$

14.4.3   Minimum Variance Portfolio

Suppose now we wish to invest our portfolio by specifying a desired return and then minimizing the variance. The only change that occurs is that we gain a new condition.

Now we wish to choose x_1, x_2, x_3 which minimizes:

$$Var(R) = x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + 2x_1x_2\sigma_{12}^2 + 2x_1x_3\sigma_{13}^2 + 2x_2x_3\sigma_{23}^2$$

subject to the two constraints:

$$x_1 + x_2 + x_3 = 1$$

and:

$$E(R) = \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 = \mu_0$$

where μ_0 is the desired return.

We define the Lagrange function:

$$\begin{aligned}
L &= x_1^2\sigma_1^2 + x_2^2\sigma_2^2 + x_3^2\sigma_3^2 + 2x_1x_2\sigma_{12}^2 + 2x_1x_3\sigma_{13}^2 + 2x_2x_3\sigma_{23}^2 \\
&+ \lambda_1(x_1\mu_1 + x_2\mu_2 + x_3\mu_3 - \mu_0) \\
&+ \lambda_2(x_1 + x_2 + x_3 - 1)
\end{aligned}$$

and solve the system $\nabla L = \bar{0}$ which is:

$$\begin{aligned}
2\sigma_1^2 x_1 + 2\sigma_{12}^2 x_2 + 2\sigma_{13}^2 x_3 + \mu_1\lambda_1 + \lambda_2 &= 0 \\
2\sigma_{12}^2 x_1 + 2\sigma_2^2 x_2 + 2\sigma_{23}^2 x_3 + \mu_2\lambda_1 + \lambda_2 &= 0 \\
2\sigma_{13}^2 x_1 + 2\sigma_{23}^2 x_2 + 2\sigma_3^2 x_3 + \mu_3\lambda_1 + \lambda_2 &= 0 \\
\mu_1 x_2 + \mu_2 x_2 + \mu_3 x_3 - \mu_0 &= 0 \\
x_1 + x_2 + x_3 - 1 &= 0
\end{aligned}$$

Notice that this can be rewritten as a matrix equation:

$$
\begin{bmatrix}
2\sigma_1^2 & 2\sigma_{12}^2 & 2\sigma_{13}^2 & \mu_1 & 1 \\
2\sigma_{12}^2 & 2\sigma_2^2 & 2\sigma_{23}^2 & \mu_2 & 1 \\
2\sigma_{13}^2 & 2\sigma_{23}^2 & 2\sigma_3^2 & \mu_3 & 1 \\
\mu_1 & \mu_2 & \mu_3 & 0 & 0 \\
1 & 1 & 1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ \mu_0 \\ 1
\end{bmatrix}
$$

which can be rewritten much more simply using the covariance matrix as:

$$
\begin{bmatrix}
2\Sigma & \bar{\mu} & \bar{1} \\
\bar{\mu}^T & 0 & 0 \\
\bar{1}^T & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\bar{x} \\ \lambda_1 \\ \lambda_2
\end{bmatrix}
=
\begin{bmatrix}
\bar{0} \\ \mu_0 \\ 1
\end{bmatrix}
$$

which is again especially nice since it generalizes to more than three options and can be solved via a matrix inverse:

$$
\begin{bmatrix}
\bar{x} \\ \lambda_1 \\ \lambda_2
\end{bmatrix}
=
\begin{bmatrix}
2\Sigma & \bar{\mu} & \bar{1} \\
\bar{\mu}^T & 0 & 0 \\
\bar{1}^T & 0 & 0
\end{bmatrix}^{-1}
\begin{bmatrix}
\bar{0} \\ \mu_0 \\ 1
\end{bmatrix}
$$

Example 14.9 Revisited.

Again, the questions as to why this matrix is invertible and why this yields a solution of minimum variance will be left for later.

To continue our example above suppose that we wish for a return of $\mu_0 = 0.0305$. Then the corresponding minimum variance portfolio can be found via:

$$
\begin{bmatrix}
\bar{x} \\ \lambda_1 \\ \lambda_2
\end{bmatrix}
=
\begin{bmatrix}
2\Sigma & \bar{\mu} & \bar{1} \\
\bar{\mu}^T & 0 & 0 \\
\bar{1}^T & 0 & 0
\end{bmatrix}^{-1}
\begin{bmatrix}
\bar{0} \\ \mu_0 \\ 1
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0.0220 & 0.0030 & 0.0016 & 0.0385 & 1.0000 \\
0.0030 & 0.0242 & 0.0032 & 0.0021 & 1.0000 \\
0.0016 & 0.0032 & 0.0436 & 0.0202 & 1.0000 \\
0.0385 & 0.0021 & 0.0202 & 0 & 0 \\
1.0000 & 1.0000 & 1.0000 & 0 & 0
\end{bmatrix}^{-1}
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0.0305 \\ 1
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0.6770 \\ 0.1154 \\ 0.2076 \\ -0.2770 \\ -0.0049
\end{bmatrix}
$$

Thus to minimize our risk we should set:

$$\bar{x} = \begin{bmatrix} 0.6770 \\ 0.1154 \\ 0.2076 \end{bmatrix}$$

meaning proportionally speaking we should invest 0.6770 of our portfolio in Company 1, 0.1154 of our portfolio in Company 2, and 0.2076 of our portfolio in Company 3.

Notice that the variance in this case equals:

$$Var(R) = \bar{x}^T \Sigma \bar{x} = 0.0067$$

Any other $\bar{x}$ which achieves the same return would have a larger variance.

Notice also that this variance is higher than the global minimimum variance. In this case we are getting a higher return , our desired 0.0305 rather than the 0.0214 return from the global minimum variance portfolio return, in exchange for a higher variance, 0.0067 rather than 0.0054.

14.4.4   Extreme Examples and Interpretations

The example in the previous section was fairly straightforward but it's worth adjusting the example a couple of different ways to see the results:

Example 14.9 Revisited.

Suppose we want a return of 0.035. Let's see what happens.

$$\begin{bmatrix} \bar{x} \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 2\Sigma & \bar{\mu} & \bar{1} \\ \bar{\mu}^T & 0 & 0 \\ \bar{1}^T & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{0} \\ 0.035 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.8002 \\ -0.0087 \\ 0.2085 \\ -0.4136 \\ -0.0020 \end{bmatrix}$$

This suggests that we should allocate our portfolio according to:

$$\bar{x} = \begin{bmatrix} 0.8002 \\ -0.0087 \\ 0.2085 \end{bmatrix}$$

Before wondering about the negative allocation, note the risk here:

$$Var(R) = \bar{x}^T \Sigma \bar{x} = 0.0082$$

This is higher than either of the previous.

Okay, so what's going on here with the negative allocation? How can we have a negative amount of Company 2 and how can our two positive amounts add to more than 1?

The answer to this first question is initially that the mathematics doesn't know that, it simply finds the $\bar{x}$ which does the job. However this is not totally unreasonable. It's possible when investing to hold a *short position* on a stock. This works as follows:

Example 14.10. Suppose a stock costs $10 per share. You believe it's going to go down soon so clearly you would not buy it. However because you believe it's going to go down you borrow 100 shares from your broker and sell them, earning a quick $1000. Suppose later when the stock drops to $5 per share you buy 100 shares for $500 and give them back to the broker. You have earned $500. This is called closing the short position.

Before this final sale you are said to have a short (negative) position on the stock since you technically purchased -100 shares when you borrowed and sold them.

When the stock goes down you are happy because you are gaining money relative to if you closed the position and when the stock goes up you are unhappy because you are losing money relative to if you closed the position.

A negative value in $\bar{x}$ can be interpreted simply as that, holding a short position on the stock.

The answer to this second question is that since you earned money from your short position you can allocate more money to the other two. The net total portfolio value is still 1.

Not all brokers allow short positions as they are very risky.

Example 14.9 Revisited.

Suppose we get greedy and we want a return of 0.1, a seriously large 10% return! Let's see what happens:

$$
\begin{bmatrix} \bar{x} \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 2\Sigma & \bar{\mu} & \bar{1} \\ \bar{\mu}^T & 0 & 0 \\ \bar{1}^T & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{0} \\ 0.1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.5792 \\ -1.8011 \\ 0.2219 \\ -2.3855 \\ 0.0401 \end{bmatrix}
$$

This suggests that we should allocate our portfolio according to:

$$\bar{x} = \begin{bmatrix} 2.5792 \\ -1.8011 \\ 0.2219 \end{bmatrix}$$

So it seems possible even with the strange numbers, so why wouldn't we do it? The answer is in the risk (the variance). The variance corresponding to this return is:

$$Var(R) = \bar{x}^T \Sigma \bar{x} = 0.0992$$

which is quite high.

This becomes even more pronounced if we get even greedier and demand, for example 100% return.

We find:

$$\begin{bmatrix} \bar{x} \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 2\Sigma & \bar{\mu} & \bar{1} \\ \bar{\mu}^T & 0 & 0 \\ \bar{1}^T & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{0} \\ 1.00 \\ 1 \end{bmatrix} = \begin{bmatrix} 27.2121 \\ -26.6198 \\ 0.4077 \\ -29.6900 \\ 0.6236 \end{bmatrix}$$

This suggests that we should allocate our portfolio according to:

$$\bar{x} = \begin{bmatrix} 27.2121 \\ -26.6198 \\ 0.4077 \end{bmatrix}$$

So we'd have to massively short Company 2 and invest the money from this short primarily in Company 1.

However the variance corresponding to this return is:

$$Var(R) = \bar{x}^T \Sigma \bar{x} = 14.5332$$

which implies that this would be an extremely risky (stupidly so) portfolio.

14.4.5 Questions to Answer

A few closing notes:

(I) Why does the Method of Lagrange Multipliers find a minimum?

Since variance $Var(R) = \bar{x}^T \Sigma \bar{x}$ is always nonnegative when we are looking at the global minimum variance portfolio if we examine the set:

18

$$\left\{ \bar{x}^T \Sigma \bar{x} \,\middle|\, \bar{x}^T \bar{1} = 1 \right\}$$

or when we are looking the minimum variance portfolio if we examine the set:

$$\left\{ \bar{x}^T \Sigma \bar{x} \,\middle|\, \bar{x}^T \bar{1} = 1, \bar{x}^T \bar{\mu} = \mu_0 \right\}$$

we know the set has an infimum which is in fact a minimum by the continuity of all functions involved. This guarantees that the method of Lagrange Multipliers will in fact find that minimum.

(II) The second such set actually leads to a nice graph which shows what's going on.

If we plot a set of desired returns μ_0 along with their corresponding variance values we see the pattern:

Example 14.9 Revisited.

Here is a plot of the minimum variance values coresponding to desired return values between 0 and 0.1:
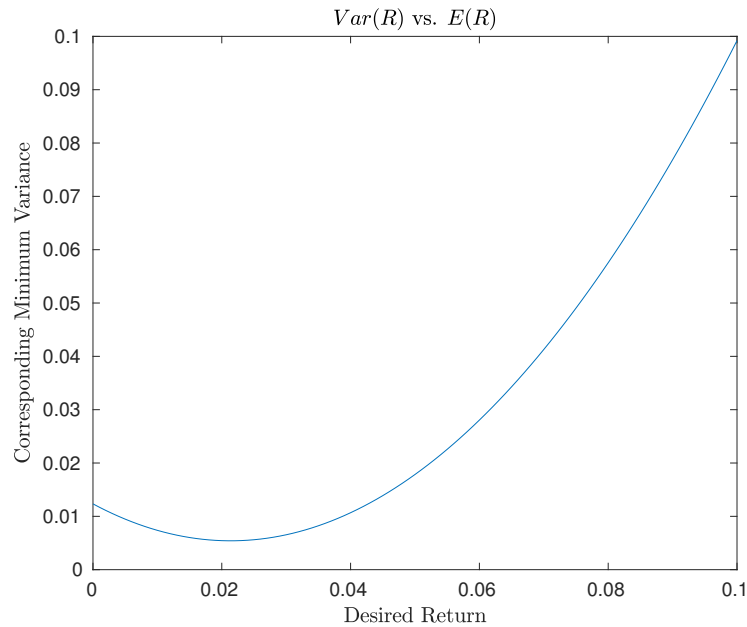
From this graph we see that there is a minimum variance of about 0.005 corresponding to a return of slightly more than 0.02. The exact values are 0.0054 and 0.0214 which we found as our global minimum variance portfolio.

(III) Why does the matrix inverse exist when we used it?

Omitted for now since a simple explanation isn't clear to me.

14.5   Matlab

If we have the covariance matrix we can create the appropriate necessary matrix very easily for the global minimum variance portfolio easily.

Here is an example where we start with the covariance matrix Σ, construct the matrix:

$$M = \left[\begin{array}{cc} 2\Sigma & \bar{1} \\ \bar{1}^T & 0 \end{array} \right]$$

and use it to find the global minimum variance portfolio using the inverse as per the text:

```
>> S = [
0.0110    0.0015    0.0008
0.0015    0.0121    0.0016
0.0008    0.0016    0.0218];
>> M = [
2*S ones(3,1)
ones(3,1)' 0
]
M =
    0.0220    0.0030    0.0016    1.0000
    0.0030    0.0242    0.0032    1.0000
    0.0016    0.0032    0.0436    1.0000
    1.0000    1.0000    1.0000         0
>> a = inv(M)*[zeros(3,1);1]
a =
    0.4271
    0.3672
    0.2057
   -0.0108
```

Note: We could do `ones(1,3)` instead of `ones(3,1)'` but it's been left this way to be consistent with the mathematical notation in the text.

Then we can also find the associated variance:

```
>> x = a(1:3);
>> x'*S*x
ans =
    0.0054
```

If in addition we have the expected value vector and a specified desired return then we can find the minimum variance portfolio.

Here is how we construct and use the matrix:

$$M = \begin{bmatrix} 2\Sigma & \bar{\mu} & \bar{1} \\ \bar{\mu}^T & 0 & 0 \\ \bar{1}^T & 0 & 0 \end{bmatrix}$$

```
>> mu = [0.0385;0.0021;0.0202]
mu =
    0.0385
    0.0021
    0.0202
>> M = [
2*S mu ones(3,1)
mu' 0 0
ones(3,1)' 0 0
]
M =
    0.0220    0.0030    0.0016    0.0385    1.0000
    0.0030    0.0242    0.0032    0.0021    1.0000
    0.0016    0.0032    0.0436    0.0202    1.0000
    0.0385    0.0021    0.0202         0         0
    1.0000    1.0000    1.0000         0         0
>> inv(M)*[zeros(3,1);0.0305;1]
ans =
    0.6770
    0.1154
    0.2076
   -0.2770
   -0.0049
```

If we'd like to check the variance that's easy:

```
>> a = inv(M)*[zeros(3,1);0.0305;1];
>> x = a(1:3);
>> x'*S*x
ans =
    0.0067
```

The graph of variance versus return can be drawn easily. Here is the example which appears in the text. We plot all return values from 0 to 0.01 in steps of 0.001. The titles and labels are just there to be pretty. The picture is omitted because it's included in the text.

```
rets = [0:0.001:0.10];
risks = [];
for mu0 = rets
    M = [2*S mu ones(3,1) ; mu' 0 0 ; ones(3,1)' 0 0];
    soln = inv(M)*[ones(3,1);mu0;1];
    xbar = soln(1:3);
    risks = [risks,xbar'*S*xbar];
end
plot(rets,risks)
title('$Var(R)$ vs. $E(R)$','interpreter','latex');
xlabel('$\mu_0$','interpreter','latex');
ylabel('$\bar x^T\Sigma\bar x$','interpreter','latex');
```

14.6   Exercises

Exercise 14.1. Suppose the covariance matrix for the stocks of three companies 1, 2 and 3 is given by:

$$\Sigma = \begin{bmatrix} 0.0250 & 0.0018 & -0.0021 \\ 0.0018 & 0.0301 & 0.0010 \\ -0.0021 & 0.0010 & 0.0111 \end{bmatrix}$$

(a) How should a portfolio be allocated in order to minimize the variance and what would the associated variance be?

(b) If the average returns of the three stocks are given by $\mu = [0.0100; 0.0200; 0.0182]$ and the desired return is 0.0150 how should a portfolio be allocated in order to minimize the variance and what would the associated variance be?

(c) Answer the previous questions in the context of a portfolio with total value $84230.

Exercise 14.2. Suppose the covariance matrix for the stocks of three companies 1, 2 and 3 is given by:

$$\Sigma = \begin{bmatrix} 0.0200 & -0.0005 & -0.0009 \\ -0.0005 & 0.0210 & -0.0007 \\ -0.0009 & -0.0007 & 0.0180 \end{bmatrix}$$

(a) How should a portfolio be allocated in order to minimize the variance and what would the associated variance be?

(b) If the average returns of the three stocks are given by $\mu = [0.0200; 0.0210; 0.0112]$ and the desired return is 0.0250 how should a portfolio be allocated in order to minimize the variance and what would the associated variance be?

(c) Answer the previous questions in the context of a portfolio with total value $8.4M.

Exercise 14.3. At the instant this problem is written the 30-day historical data for Bitcoin and Ethereum has approximately the following covariance matrix where index 1 corresponds to Bitcoin and index 2 corresponds to Ethereum:

$$\Sigma = \begin{bmatrix} 0.024 & 0.012 \\ 0.012 & 0.019 \end{bmatrix}$$

How should a portfolio be allocated in order to minimize the variance and what would the associated variance be?

Exercise 14.4. Suppose the covariance matrix for the stocks of three companies 1, 2 and 3 is given by:

$$\Sigma = \begin{bmatrix} 0.0250 & 0.0018 & -0.0021 \\ 0.0018 & 0.0301 & 0.0010 \\ -0.0021 & 0.0010 & 0.0111 \end{bmatrix}$$

Suppose the average returns of the three stocks are given by $\mu = [0.0100; 0.0200; 0.0182]$.

(a) If the desired return is r how should a portfolio be allocated in order to minimize the variance? Note that your answer will have r in it.

(b) What would the associated variance be?

(c) Use the answer to the previous question to calculate the return r which would minimize the variance and calculate that minimum variance. This latter answer should agree with the answer to Exercise 15.1(a).

(d) For which values of r can you avoid shorting any of the stocks?