Fundamentals of Optimization and Application to Portfolio Selection

CQF

In this lecture...

Fundamentals of Optimization:

- how to formulate an optimization problem;
- how to use calculus to solve unconstrained optimization problems;
- the method of Lagrange;
- the Kuhn-Tucker conditions;

Application to portfolio selection:

- how to select a portfolio in the Markowitz world;
- the minimum variance portfolio;
- portfolio selection with a risk-free asset;
- how to derive the CAPM.

Part I: Fundamentals of Optimization

1 - Introduction

An optimization problem is one in which you are trying to find the "best possible value" that a function, say f, can take subject to a number of constraints. This generally involves finding the minimum or maximum of f or of a function built around f.

Optimization techniques are very often used in finance to solve a wide range of problems ranging from calculating the value of a bond yield, to solving a portfolio selection problem in discrete and continuous time and to valuing derivatives such as American and passport options.

As optimization problems come in all shape and forms, you should expect some to be incredibly easy to solve and some other to be incredibly hard. In any case, formulating the problem well will save you many headaches.

2 - Formulating an optimization problem

Optimization problems are most often defined as

$$\min_{x_1,\ldots,x_n} f(x_1,\ldots,x_n) \tag{1}$$

subject to:

$$\left.\begin{array}{c} g_1(x_1,\ldots,x_n) \\ \vdots \\ g_m(x_1,\ldots,x_n) \end{array}\right\} \begin{array}{c} \leq \\ = \\ \geq \end{array} \left\{\begin{array}{c} b_1 \\ \vdots \\ b_m \end{array}\right.$$

The function f is called the objective function. It is the function we want to optimize (here minimize).

The variables $x_1, \ldots, x_n$ are the decision variables with respect to which we want to optimize the function.

The functions $g_1, \ldots, g_m$ are the m constraints faced in our optimization. These constraints can be equalities or inequalities.

3 - Elementary rules & tips

The following set of rules will help you manipulate your optimization problem to put it in a more convenient formulation.

- some problems are more easily dealt with as a minimization, some others as a maximization. To change between the two, remember that

$$\max f(x) = -\min(-f(x))$$

$$\min f(x) = -\max(-f(x))$$

- multiplying your objective function by a positive constant k leaves the optimal point unchanged, but remember to scale back the value of the objective function at the optimal point by $1/k$

$$\max(kf(x)) = k\max(f(x)) \qquad k > 0$$

  The same is true with the min function.
- if instead of minimizing or maximizing a function, you want to set it to a value c, then you need to minimize the distance (or norm) of $f - c$. This is, however, beyond the scope of this presentation.

4 - Unconstrained optimization

In the absence of any constraint, our system (1) reduces to:

$$\min_{x_1,\ldots,x_n} f(x_1, \ldots, x_n)$$

Now, this is just standard (albeit multivariate) calculus!

So we know that the minimal value of f is an extremum of the function.

How can we pick a vector $x^* = (x_1^*, \ldots, x_n^*)$ so that f reaches a global minimum?

From standard calculus, we know that:

- the gradient (vector of derivatives) of f at x^* must be zero. This is a necessary condition, but it is not sufficient as minima, maxima and inflection points all have a derivative reaching 0!
- the Hessian (i.e. matrix of second derivatives) of f at x^* must be positive definite (negative definite for a maximization). This is a sufficient condition.

These two conditions are fundamental in optimization. The first condition is used to find a set of potential solutions and the second is used to check which of these answers satisfy(ies) the problem.

They are referred to as first order (necessary) condition and second order (sufficient) condition.

5 - Optimization with equality constraints

With linear constraints, the optimization problem becomes:

$$\min_{x_1,\ldots,x_n} f(x_1,\ldots,x_n)$$

subject to:

$$\begin{aligned} g_1(x_1,\ldots,x_n) &= b_1 \\ &\vdots \\ g_m(x_1,\ldots,x_n) &= b_m \end{aligned}$$

Obviously, now we cannot use straight standard calculus to solve our minimization.

However, wouldn't it be nice if we could transform our objective function f to somehow integrate all the constraints so that we can just apply standard calculus on the new function to solve the problem?

This can be done. It is called the Lagrange method.

The method of Lagrange

First form the Lagrangian function $L(x, \lambda)$. L is our objective function f augmented by the addition of the constraint functions. Each constraint function is multiplied by a variable, called a lagrange multiplier.

$$L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j (g_j(x) - b_j)$$

Note that to alleviate the notation we have resorted to using the vectors $x = (x_1, \ldots, x_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ where appropriate.

The Lagrange function effectively transforms a problem in n variables $(x_1, \ldots, x_n)$ and m constraint into an <u>unconstrained optimization</u> with $n + m$ variables $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m)$.

We now reformulate our problem in terms of the Lagrange function:

$$\min_{x,\lambda} L(x, \lambda)$$

and use standard calculus to solve the problem.

By the first order condition, we get a system of $n + m$ equations:

$$\frac{\partial L}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x) + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(x) = 0$$

$$i = 1, \ldots, n$$

$$\frac{\partial L}{\partial \lambda_j}(x) = g_j(x) - b_j = 0 \qquad j = 1, \ldots, m$$

Note that:

- the first n equations correspond to an unconstrained optimization over f penalized by a sum of functions parameterized by λ;
- we have recovered our constraints in the last m equations.

Once the system of equations is solved, all we need to do is check the Hessian of f before concluding [1].

[1] Strictly speaking, checking the Hessian is not enough in general optimization problems, but it will prove sufficient for the types of applications we consider in this lecture.

<u>Application:</u> Maximizing the Area of a Rectangle

We would like to find the dimensions x and y of the sides of a rectangle so that the area of the rectangle is maximized but its perimeter remains equal to p.

The area of the rectangle $A(x, y)$ is given by $A(x, y) = xy$ and the perimeter $P(x, y)$ is $P(x, y) = 2(x + y) = p$

The optimization problem we must solve is formulated as

$$\max_{x,y} A(x, y) = xy$$

subject to

$$P(x, y) - p = 2(x + y) - p = 0$$

Since we have one constraint, we introduce one lagrange multiplier λ and define the lagrange function $L(x, y, \lambda)$ as

$$L(x, y, \lambda) = xy - \lambda(2(x + y) - p)$$

taking the derivatives with respect to each decision variables, we obtain the system:

$$\begin{aligned} 0 &= \frac{\partial L}{\partial x} = y - 2\lambda \\ 0 &= \frac{\partial L}{\partial y} = x - 2\lambda \\ 0 &= \frac{\partial L}{\partial \lambda} = 2x + 2y - p \end{aligned}$$

Solving the system yields:

$$\left\{ \begin{array}{lcl} x & = & p/4 \\ y & = & p/4 \\ \lambda & = & p/8 \end{array} \right.$$

With this choice of x, y and λ, the area of the rectangle is $A(x, y) = p^2/16$

This satisfies the first order condition. What about the second order condition? The Hessian of A,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is neither positive definite nor negative definite. We must therefore find another way to check if we have reached a maximum, minimum or neither.

One way of doing that is by parametrizing the sides of a rectangle as follows. Call x the shorter side and y the longest one. Let $\epsilon \in (0, p/4)$, then we can parametrize any rectangle satisfying the perimeter constraint as:

$$\begin{cases} x &= p/4 - \epsilon \\ y &= p/4 + \epsilon \end{cases}$$

The area of such rectangle is

$$A(x, y) = (p/4 - \epsilon)(p/4 + \epsilon) = p^2/16 - \epsilon^2 < p^2/16$$

The choice of x, y and λ obtained through the Lagrange method is therefore maximizing and solves our problem.

6 - Optimization with inequality constraints

We can always write an optimization problem with inequality constraints as:

$$\min_{x_1,\ldots,x_n} f(x_1,\ldots,x_n)$$

subject to:

$$\begin{gathered} g_1(x_1,\ldots,x_n) \leq b_1 \\ \vdots \\ g_m(x_1,\ldots,x_n) \leq b_m \end{gathered}$$

The Kuhn-Tucker theorem provides a set of <u>necessary</u> conditions for the existence of an optimal solution.

One formulation of the Kuhn-Tucker conditions is:

$$\lambda_0 \frac{\partial f(x)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial (g_j(x) - b_j)}{\partial x_i} \qquad i = 1, \ldots, n$$

$$\lambda_j (g_j(x) - b_j) = 0 \qquad j = 1, \ldots, m$$

$$\lambda_j \geq 0 \qquad j = 0, \ldots, m$$

Underneath the first condition, we recognize the familiar form of the Lagrange function, except that there is an extra Lagrange multiplier. The second set of conditions relate to the constraints, and the third set of conditions relate to the Lagrange multipliers themselves.

Unfortunately, Kuhn-Tucker does not guarantee that a vector x satisfying the condition will be optimal in the general case. The Kuhn-Tucker condition acts as necessary and sufficient condition only when the objective function is convex and the constraints are linear.

Part II: Application to Portfolio Selection

Now comes the time to apply what we have seen to a concrete example: selecting a portfolio of assets.

Throughout this example, we will use vectors and matrices rather than scalars in order to make the exposition as general as possible.

1 - The setting

We are in an economy with n different assets. Each asset i is entirely characterized by its expected return μ_i and expected standard deviation σ_i. In addition, assets i and j are correlated with correlation ρ_{ij}. The proportion of the portfolio invested in asset i is w_i.

The vector of asset expected returns μ is defined as:

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_i \\ \vdots \\ \mu_n \end{pmatrix}$$

The covariance matrix Σ is given by:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & & & \\ \rho_{n1}\sigma_n\sigma_1 & \cdots & \cdots & \sigma_n^2 \end{pmatrix}$$

The vector of weights is:

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{pmatrix}$$

Therefore, the portfolio return μ_π can be written as

$$\mu_\pi = \mu^T w$$

and the portfolio variance, σ_π^2, is given by

$$\sigma_\pi^2 = w^T \Sigma w$$

where the superscript ^{T} denotes the transpose of the vector.

Aside 1: A Useful Covariance Matrix Decomposition

If you have the correlation matrix and the standard deviation vector, you can calculate the covariance matrix in just two steps.
Step 1: Form the diagonal standard deviation matrix S

Consider the standard deviation vector σ:

$$\sigma = \begin{pmatrix} \sigma_1 \\ \vdots \\ \sigma_i \\ \vdots \\ \sigma_n \end{pmatrix}$$

As its name indicates, the diagonal standard deviation matrix is a matrix with standard deviations on its diagonal and 0 everywhere else:

$$S = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \cdots & \sigma_n \end{pmatrix}$$

Step 2: compute the covariance matrix Σ

Consider the Correlation matrix R defined as:

$$R = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & & & \\ \rho_{n1} & \cdots & \cdots & 1 \end{pmatrix}$$

The covariance matrix Σ can be obtained by pre and post multiplying the correlation matrix by the diagonal standard deviation matrix:

$$\Sigma = SRS = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \cdots & \rho_{2n}\sigma_2\sigma_n \\ \vdots & & & \\ \rho_{n1}\sigma_n\sigma_1 & \cdots & \cdots & \sigma_n^2 \end{pmatrix}$$

Note that we should normally have $S^T RS$, but since S is diagonal, then $S^T = S$.

This decomposition of the covariance matrix is not only useful for portfolio selection for risk measurement. It is actually used by RiskMetrics in the calculation of parametric Value at Risk.

(End of Aside).

2 - Problem formulation

The portfolio selection problem is generally defined as a minimization of risk subject to a return constraint. Two reasons for this convention are:

- a return objective seems intuitively easier to formulate than a risk objective;
- risks are easier to control than returns;

If we adopt this convention, our objective function is the portfolio variance, and we will minimize it with respects to the portfolio weights. Actually, instead of using the portfolio variance, we will use a little trick and scale it down by a factor of $1/2$ to ease our calculations. Since the factor is positive, it does not affect the value of the optimal vector of weights w^*.

$$\min_w \frac{1}{2}\sigma_\pi^2 = \frac{1}{2}w^T\Sigma w$$

Now for our constraints. We already have one constraint: the portfolio return must be equal to a prespecified level m. In mathematical terms:

$$\mu_\pi = \mu^T w = m$$

We also have a second constraint on the weights called the 'budget equation'. The sum of all the weights must necessarily equal 1. Since there is no risk-free assets, our wealth must be entirely invested in a combination of the n assets.

$$\mathbf{1}^T w = 1$$

where $\mathbf{1}$ is a n-element unit vector:

$$\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Summing it all, we formulate the portfolio selection problem as

$$\min_{w} \frac{1}{2} w^T \Sigma w$$

Subject to:

$$\begin{aligned} \mu^T w &= m \\ \mathbf{1}^T w &= 1 \end{aligned}$$

3 - Solving the portfolio selection problem

This problem is an optimization with equality constraints. Therefore it can be solved using the method of Lagrange.

We form the Lagrange function with two Lagrange multipliers λ and γ:

$$L(w, \lambda, \gamma) = \frac{1}{2} w^T \Sigma w + \lambda(m - \mu^T w) + \gamma(1 - \mathbf{1}^T w)$$

We now solve for the first order condition by taking the derivative with respect to the vector w:

$$\frac{\partial L}{\partial w}(w, \lambda, \gamma) = w^T \Sigma - \lambda \mu^T - \gamma \mathbf{1}^T = 0$$

Checking the second order condition, the Hessian of the objective function is none else than Σ, the covariance matrix, which is positive definite. Therefore, we have reached the optimal weight vector w^*:

$$w^* = \Sigma^{-1}(\lambda \mu + \gamma \mathbf{1}) \tag{2}$$

To get to this relationship, we have taken the transpose of the previous equation and multiply it by the inverse matrix Σ^{-1}.

All we have to do now is to find the values for λ and γ and then substitute them into (2). Substituting w^* into the two constraint equations yields:

$$\mu^T \Sigma^{-1}(\lambda\mu + \gamma\mathbf{1}) = \lambda\mu^T\Sigma^{-1}\mu + \gamma\mu^T\Sigma^{-1}\mathbf{1} = m$$

$$\mathbf{1}^T \Sigma^{-1}(\lambda\mu + \gamma\mathbf{1}) = \lambda\mathbf{1}^T\Sigma^{-1}\mu + \gamma\mathbf{1}^T\Sigma^{-1}\mathbf{1} = 1$$

Define:

$$\begin{cases} A = \mathbf{1}^T \Sigma^{-1} \mathbf{1} \\ B = \mu^T \Sigma^{-1} \mathbf{1} \\ C = \mu^T \Sigma^{-1} \mu \end{cases}$$

Note also that $AC - B^2 > 0$.

Then

$$\begin{cases} \lambda = \frac{Am-B}{AC-B^2} \\ \gamma = \frac{C-Bm}{AC-B^2} \end{cases}$$

Now all we need is to substitute these values back into (2) to obtain w^*.

Application

Consider a market with 4 assets X_1, X_2, X_3 and X_4. Their return vector and standard deviation vector are respectively given by

$$\mu = \begin{pmatrix} 0.05 \\ 0.07 \\ 0.15 \\ 0.27 \end{pmatrix}$$

and

$$\sigma = \begin{pmatrix} 0.07 \\ 0.12 \\ 0.30 \\ 0.60 \end{pmatrix}$$

The correlation between asset returns is given by

$$R = \begin{pmatrix} 1 & 0.8 & 0.5 & 0.4 \\ 0.8 & 1 & 0.7 & 0.5 \\ 0.5 & 0.7 & 1 & 0.8 \\ 0.4 & 0.5 & 0.8 & 1 \end{pmatrix}$$

What is the optimal asset allocation to obtain a return $m = 10\%$

Answer:

The first step is to create the covariance matrix Σ:

$$\Sigma = SRS = \begin{pmatrix} 0.0049 & 0.00672 & 0.0105 & 0.0168 \\ 0.00672 & 0.0144 & 0.0252 & 0.036 \\ 0.0105 & 0.0252 & 0.09 & 0.144 \\ 0.0168 & 0.036 & 0.144 & 0.36 \end{pmatrix}$$

Next, we compute A, B and C...

$$\begin{cases} A = \mathbf{1}^T \Sigma^{-1} \mathbf{1} = 239.3440468 \\ B = \mu^T \Sigma^{-1} \mathbf{1} = 9.618456078 \\ C = \mu^T \Sigma^{-1} \mu = 0.550280982 \end{cases}$$

... as well as λ well as γ

$$\begin{cases} \lambda = 0.3652794 \\ \gamma = -0.0105013 \end{cases}$$

Now, we can answer our question. The optimal asset allocation to obtain a return $m = 10\%$ is given by

$$\begin{aligned} w^* &= \Sigma^{-1}(\lambda\mu + \gamma\mathbf{1}) \\ &\approx \Sigma^{-1}(0.3652794\mu - 0.01050131\mathbf{1}) \\ &\approx \begin{pmatrix} 0.528412108 \\ 0.172888075 \\ 0.159764343 \\ 0.138935474 \end{pmatrix} \end{aligned}$$

If we were to double check with Excel Solver, we would find that

$$w^* \approx \begin{pmatrix} 0.528412169 \\ 0.172888796 \\ 0.159764425 \\ 0.138935450 \end{pmatrix}$$

As expected, the theory works!

4 - The minimum-variance portfolio

The weights of a portfolio depend on the return objective m specified in the constraints. By letting m span the real line, we will obtain the minimum-variance frontier.

The variance of a portfolio can be expressed as a function of m. A bit of algebra shows that:

$$\sigma_\pi^2(m) = \frac{Am^2 - 2Bm + C}{AC - B^2} \tag{3}$$

The minimum variance portfolio is the portfolio solving the unconstrained problem:

$$\min_{m} \sigma_{\pi}^{2}(m) \tag{4}$$

Solving for the first order condition, we would get:

$$\frac{\partial \sigma_{\pi}^{2}(m)}{\partial m} = 2Am - 2B = 0$$

After checking the second order condition, we get:

$$m_g = \frac{B}{A}$$

Now that we have m_g, we can easily get the vector of weights w_g:

$$w_g = \frac{\Sigma^{-1}\mathbf{1}}{A}$$

and finally the variance:

$$\sigma_g^2 = \frac{1}{A}$$

Application... continued...

What is the global minimum variance portfolio's asset allocation?
What are its return and standard deviation?

Answer:

The global minimum variance portfolio's asset allocation is given by

$$w_g = \frac{\Sigma^{-1}\mathbf{1}}{A} \approx \begin{pmatrix} 1.274886723 \\ -0.263112728 \\ 0.016339421 \\ -0.028113415 \end{pmatrix}$$

its return is

$$m_g = \frac{B}{A} \approx 0.040186736$$

and its standard deviation is equal to

$$\sigma_g = \sqrt{w_g^T \Sigma w_g} \approx 0.064638115$$

5 - Efficient Portfolios Revisited

5.1 - A New Parametrization

So far, we have seen how to compute the vector of weight of an efficient portfolio directly, through the explicit resolution of an optimization problem.

Now we will see an alternate parametrization of efficient portfolios resting implicitly on the definition of the optimization problem.

From equation (2),

$$w^* = \lambda \Sigma^{-1} \mu + \gamma \Sigma^{-1} \mathbf{1}$$

Notice that the second term of the right-hand side is γA times the weight of the minimum-variance portfolio.

Pushing the analysis a little further we see that any efficient portfolio can be expressed as a linear combination of exactly two efficient portfolios

- the global minimum variance portfolio Π_g;
- an efficient portfolio Π_d defined by the investment strategy $w_d := \frac{\Sigma^{-1}\mu}{B}$.

with respective "weights" γA and λB.

To conclude, we have defined the following parametrization:

$$w^* = \gamma A w_g + \lambda B w_d$$

with the obvious budget constraint:

$$\gamma A + \lambda B = 1$$

5.2 - Covariance of Efficient Portfolios

The covariance of the global minimum variance portfolio has the interesting property that its covariance with any other portfolio p is $\frac{1}{A}$.

To see this, consider an arbitrary portfolio p, then

$$\begin{aligned} \sigma^2_{g,p} &= w_g^T \Sigma w_p \\ &= \frac{1^T \Sigma^{-1} \Sigma w_p}{A} \\ &= \frac{1}{A} \end{aligned}$$

where the last step follows from the fact that since p is a portfolio, its asset allocation w_p satisfies the budget equation

$$\mathbf{1}^T w_p = 1$$

Consider two arbitrary portfolios Π_a and Π_b. Using the characterization previously developed, we state that there exists two numbers a and b such that

$$w_a = (1 - a)w_g + aw_d$$

and

$$w_b = (1 - b)w_g + bw_d$$

The covariance of the two portfolios is:

$$\begin{aligned}
\sigma_{a,b} &= (1-a)(1-b)\sigma_g^2 + ab\sigma_d^2 \\
&\quad +[a(1-b)+b(1-a)]\sigma_{g,d} \\
&= \frac{(1-a)(1-b)}{A} + \frac{abC}{B^2} + \frac{a+b-2ab}{A} \\
&= \frac{1}{A} + \frac{ab(AC-B^2)}{AB^2}
\end{aligned}$$

6 - Introducing the risk-free asset

Now, we introduce a risk-free asset, such as a bank account, with certain return r, 0 variance and 0 correlation with the other (risky) assets. The weight w_0 of the risk-free asset is defined as the weight of the portfolio not invested in risky assets. Mathematically,

$$w_0 = 1 - \mathbf{1}^T w$$

The return of the portfolio is given by:

$$\mu_\pi = \mu^T w + r(1 - \mathbf{1}^T w) = r + (\mu - r\mathbf{1})^T w$$

So, the introduction of a risk-free asset split returns between the risk-free rate r and the vector of risk premia $(\mu - r\mathbf{1})$.

Also, the budget constraint has now vanished since the balance of our wealth not invested in the risky assets is now in the risk-free asset.

The variance of the portfolio is unaffected by the introduction of the risk-free asset since the risk-free asset does not contribute to the risk of the portfolio.

6.1 - Problem formulation with a risk-free asset

Our new optimization problem is:

$$\min_{w} \frac{1}{2} w^T \Sigma w$$

Subject to:

$$r + (\mu - r\mathbf{1})^T w = m$$

We form the Lagrange function:

$$L(x, \lambda) = \frac{1}{2} w^T \Sigma w + \lambda(m - r - (\mu - r\mathbf{1})^T w)$$

We now solve for the first order condition by taking the derivative with respect to the vector w:

$$\frac{\partial L}{\partial w} = w^T \Sigma - \lambda(\mu - r\mathbf{1})^T = 0$$

Checking the second order condition, the Hessian of the objective function is still the covariance matrix, which is positive definite. Therefore, we have reached the optimal weight vector w^*:

$$w^* = \lambda \Sigma^{-1}(\mu - r\mathbf{1})$$

Here again, we have transposed our equation and multiplied both sides by the inverse matrix Σ^{-1}.

Substituting the value of w^* into the constraint, we solve for λ:

$$\lambda = \frac{m - r}{(\mu - r\mathbf{1})^T \Sigma^{-1} (\mu - r\mathbf{1})}$$

Finally, we get w^*:

$$w^* = \frac{(m - r)\Sigma^{-1}(\mu - r\mathbf{1})}{(\mu - r\mathbf{1})^T \Sigma^{-1} (\mu - r\mathbf{1})}$$

Application ...continued...

Assume the risk-free rate is 2.5%. What is the allocation of the portfolio returning 10%?

Answer:

The allocation of the portfolio returning 10%? is given by

$$w^* = \frac{(m-r)\Sigma^{-1}(\mu - r\mathbf{1})}{(\mu - r\mathbf{1})^T\Sigma^{-1}(\mu - r\mathbf{1})} \approx \begin{pmatrix} 0.887352483 \\ 0.08126325 \\ 0.15484343 \\ 0.121648624 \end{pmatrix}$$

6.2 - The tangency portfolio

The tangency portfolio is the portfolio that is entirely invested in risky assets.

The vector of weights for the tangency portfolio, w_t is given by:

$$w_t = \frac{\Sigma^{-1}(\mu - r\mathbf{1})}{B - Ar}$$

The mean and standard deviation of this portfolio are given by

$$m_t = w_t^T \mu = \frac{C - Br}{B - Ar}$$

and

$$\sigma_t = \sqrt{w_t^T \Sigma w_t} = \sqrt{\frac{C - 2rB + R^2 A}{(B - Ar)^2}}$$

Application ...concluded

What is the tangency portfolio's asset allocation? What are its return and standard deviation?

Answer:

The asset allocation of the tangency portfolio is given by

$$w_t = \frac{\Sigma^{-1}(\mu - r\mathbf{1})}{B - Ar} \approx \begin{pmatrix} 0.712671217 \\ 0.065266037 \\ 0.124361466 \\ 0.097701279 \end{pmatrix}$$

its return is

$$m_t = \frac{C - Br}{B - Ar} \approx 0.085235749$$

and its standard deviation is equal to

$$\sigma_t = \sqrt{\frac{C - 2Br + Ar^2}{(B - Ar)^2}} \approx 0.128731141$$

6.3 - Expected return

We define the covariance vector of the tangency portfolio as

$$\boldsymbol{\sigma}_t := \Sigma \boldsymbol{w}_t = \frac{\boldsymbol{\mu} - r\mathbf{1}}{B - Ar} \tag{5}$$

where we have used bold fonts to emphasize that $\boldsymbol{\sigma}_t$ is a vector.

Premultiplying by w_t, we revert to the variance of the tangency portfolio

$$\sigma_t^2 := w_t \sigma_t = \frac{\mu_t - r}{B - Ar} \tag{6}$$

From (5) and (6)we get the following relation for asset returns

$$\mu - R\mathbf{1} = \sigma_t \frac{\mu_t - r}{\sigma_t^2} := \beta_t(\mu_t - r) \tag{7}$$

This relationship holds not only for the tangency portfolio but also for any other portfolio on the efficient frontier.

For portfolio Π_d, the relation can be significantly simplified. As a matter of fact, we have:

$$\sigma_d := \Sigma w_d = \frac{\mu}{B}$$

and therefore

$$\mu = \frac{\sigma_d}{\sigma_d^2}\mu_d =:= \beta_d \mu_d$$

7 - The CAPM

7.1 - Assumptions

Now, we have all the ingredients necessary to setup the CAPM.

So far, the results we have derived do not constitute an economic equilibrium. In fact they only represent a set of predictions given a set of fairly limited assumptions. Precisely, the CAPM achieves this transition from a predictive model (i.e. detailing what a set of relationships should (could) become) to an equilibrium model that is a model which is both explanatory (i.e. explaining why a set of relationships are the way they are) and normative (i.e. explaining how a set of relationships should (ought to) behave).

Deriving the CAPM requires the following assumptions:

1. All investors are utility maximizer;
2. All investors have a common time horizon and common predictions relating to μ and Σ;
3. All assets (including the risk-free asset) can be bought or sold in unlimited amounts;
4. Any investment is infinitely divisible;

The first assumption is economic in nature and states that rather than maximizing their returns or minimizing their risks, investors are more subtle and try to maximize the utility they derive from their wealth.

The second assumption is the most restrictive and the most difficult to satisfy as it assumes homogeneity of beliefs among market participants.

The third assumption is not realistic but is necessary if we want to obtain a tractable analytical solution. the more constraints, the larger the system and the more inequality constraints, the more difficult to find an analytical solution.

The fourth assumption is here for mathematical convenience and is relatively harmless.

Let $V(\mu_p, \sigma_p^2)$ be the utility function reflecting the utility gained from holding portfolio Π_p. We will not need to give a functional form to V at this time.

7.2 - Optimization Problem

Using the characterization of expected returns developed earlier, the optimization problem can be expressed as

$$\max_{w} V(r + w^T(\mu - r\mathbf{1}), w^T \Sigma w) \tag{8}$$

with the usual substitution in the budget equation.

Solving, we get

$$0 = \frac{\partial V}{\partial w} = \frac{\partial V}{\partial \mu_P}(\mu - r\mathbf{1}) + 2\frac{\partial V}{\partial \sigma_p^2}\Sigma w$$

$$w = \frac{-\frac{\partial V}{\partial \mu_P}}{2\frac{\partial V}{\partial \sigma_p^2}}\Sigma^{-1}(\mu - r\mathbf{1})$$

The second-order condition is met. We could check that $\frac{\partial^2 V}{\partial w \partial w^T}$ is indeed negative definite.

Equation (9) reveals that the optimal investment strategy is proportional to the tangency portfolio.

7.3 - The market portfolio

Since the optimal policy is always to invest a fraction of wealth into the tangency portfolio, every investor will hold the tangency portfolio.

As a consequence of the trading, in equilibrium, the market value of each risky asset will become proportional to their weights in the tangency portfolio.

The result is that the tangency portfolio becomes an accurate reflection of the structure of financial markets. In equilibrium, the tangency portfolio becomes the market portfolio.

Remark: We have come back full circle. All along, we were looking for an elusive characterization of the optimal portfolio of risky assets. In fact, the optimal portfolio of risky assets is "directly observable": it is just the market of all risky asset with weights determined by the relative market values of each asset!

In equilibrium, the tangency portfolio has become the market portfolio. Equation (7) becomes the Sharpe-Lintner CAPM:

$$\mu - R\mathbf{1} = \beta_M(\mu_M - r) \tag{9}$$

where the subscript M refers to the market portfolio Π_M.

Aside: Utility Maximization

Utility Functions

From an economic standpoint, wealth or return on investments are not very satisfactory measures. One of the reproaches made to wealth is that it is not representative of economic well-being. Would you be twice as happy if your salary was twice as high? Probably less than twice. Would you be half as happy if your lost half of your wealth? Probably more unhappy than that... The same reproach can be made to return.

Enters the relatively abstract, but always useful concept of utility. Utility was built to address the shortcomings of wealth. Broadly speaking, utility measures the economic well-being derived from a given level of wealth.

Utility is a general concept that can be quantified through a large number of functions. In fact, to be a utility function of wealth W, a function $U(W)$ must satisfy the following properties for $W \in \mathbb{R}^{+*}$:

Properties of Utility Functions

1. the utility function is positive: $U(W) > 0$;
2. the utility is function increasing: $U'(W) > 0$;
3. the utility is function concave: $U''(W) < 0$;

This is conform to our observation that a gain in wealth of one unit should result in a gain in utility of less than one unit and a loss of unit of wealth should result in a loss of more than one unit of utility.

Utility and Portfolio Selection

We will now revisit the optimization problem that we setup to uncover the CAPM.

The utility function traditionally associated with the portfolio selection problem is the quadratic utility function:

$$U(W) = W - \frac{kW^2}{2}$$

Taking the expectation, and expressing the result per unit of wealth, we get

$$V(\mu_p, \sigma_p^2) = \mathbf{E}[U(W)] = \mu_p - \frac{k}{2}(\mu_p + \sigma^2)$$

We now have an expression we can use in the portfolio optimization problem.