Optimal Investment Michael Tehranchi

Lecture 3 - Lent 2012

The optimal investment problem

We are ready to consider the central problem of this course. We work for now in a one-period setting to highlight the central issues.

Consider an investor who has initial wealth X_0 who wants to invest in such a way that his terminal wealth X_1 (for instance, the amount in his retirement account) is as large as possible at some fixed future date. We model X_0 as fixed real number since he knows his initial wealth now, and model X_1 as a random variable defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ since he does not know what the future will bring.

What does it mean to invest optimally? We endow the investor with a preference relation $\succ$ on the set of probability measures on $(\mathbb{R}, \mathcal{B})$ (i.e. the distribution $\mathbb{P} \circ X_1^{-1}$ of his terminal wealth) and define optimality in terms of this preference order.

We assume the numerical representation of the investor's preferences are given by the von Neumann–Morgenstern representation:

$$U_0(\mathbb{P} \circ X_1^{-1}) = \mathbb{E}\, U(X_1)$$

for a utility function U. Remember that this means we are assuming that his preference relation $\succ$ satisfy the transitivity, completeness, independence and continuity axioms.

So, the investor's problem is to

$$\text{maximise } \mathbb{E}\, U(X_1) \text{ given } X_0$$

Next, introduce a financial market model. Suppose that there are exactly n assets in the economy, and let P_t^i denote the price of asset $i \in \{1, \ldots, n\}$ at time $t \in \{0, 1\}$. We will model the time-0 price P_0^i as a fixed number, and the time-1 price P_1^i as a random variable. We will let H^i denote the investor's holdings of asset i during the time interval $(0, 1]$. We will allow each H^i to take any real value, with positive values corresponding to 'long' positions and negative values to 'short' positions.[1]

Since there are exactly n assets in the economy, we assume that the investor's initial wealth can be written as

$$\begin{aligned} X_0 &= H \cdot P_0 \quad \text{(budget constraint)} \\ &= \sum_{i=1}^{n} H^i P_0^i \end{aligned}$$

That is to say, the investor has no choice but to invest all of his money.[2]

For this lecture, we assume that the investor has no external source of income, and that he does not consume any part of his wealth. Hence, his time-1 wealth is given by

$$X_1 = H \cdot P_1 \quad \text{(self-financing constraint)}$$

In particular, all changes of the investors wealth arise from changes in the asset prices.

[1]Reality, of course, is more complicated. In real life, to short a share involves borrowing the asset from a broker and paying the broker a fee. And the broker may call for the return of the share at anytime, for instance if the price increases and the broker wants to sell it.

[2]Note that we can always assume that one of the assets, say asset 1, is cash so that $P_0^1 = P_1^1 = 1$. However, we do not need this assumption now, so we will not make it.

We can now rewrite the investor's problem is to choose the portfolio $H \in \mathbb{R}^n$ to

$$\text{maximise } \mathbb{E}\, U(H \cdot P_1) \quad \text{subject to } H \cdot P_0 = X_0.$$

Before we analyse this problem, we first recall basic ideas from Lagrangian optimisation.

Crashcourse in Lagrangian optimisation

Consider the contrained optimisation problem

$$P: \text{ maximise } f(x) \text{ subject to } g(x) = b, x \in \mathbb{R}^n$$

for given functions $f: \mathbb{R}^n \to \mathbb{R}$ and $g: \mathbb{R}^n \to \mathbb{R}^m$ and a constant $b \in \mathbb{R}^m$. The function f is called a the *objective function*, any point $x \in \mathbb{R}^n$ such that $g(x) = b$ is called *feasible solution*, and a feasible x^* is called *optimal* iff

$$f(x^*) \geq f(x)$$

for all feasible x.

To the problem P, we assign a function $L: \mathbb{R}^n \times \mathbb{R}^m$, called the *Lagrangian*, defined by

$$L(x, \lambda) = f(x) + \lambda \cdot (b - g(x)).$$

The basis theorem is this:

Theorem. (Lagrangian sufficiency) Suppose x^* is feasible and that there exists a λ^* such that

$$L(x^*, \lambda^*) \geq L(x, \lambda^*)$$

for all $x \in \mathbb{R}^n$. Then x^* is optimal for problem P.

Proof. If x is feasible then

$$\begin{aligned} L(x, \lambda) &= f(x) + \lambda \cdot (b - g(x)) \\ &= f(x). \end{aligned}$$

for any λ. Hence

$$\begin{aligned} f(x^*) &= L(x^*, \lambda^*) \\ &\geq L(x, \lambda^*) \text{ for all } x \text{ by assumption} \\ &= f(x) \text{ for all feasible } x \end{aligned}$$

□

The λ^* appearing in the hypothesis of the Lagrangian sufficiency theorem is called a *Lagrange multiplier* for the problem P. The above theorem says that *given* a feasible x^* and Lagrange multiplier λ^*, we can verify that x^* is optimal.

Do all problems have Lagrange multipliers? In general, the answer is no. Fortunately for us, though, there is an easy to check condition that a given problem has a Lagrange multiplier. To state this condition, we adopt notation to indicated the problem's dependence on the contraint b. We write consider the family of problems

$$P_b: \text{ maximise } f(x) \text{ subject to } g(x) = b, x \in \mathbb{R}^n$$

and associated Lagrangians

$$L_b(x, \lambda) = f(x) + \lambda \cdot (b - g(x)).$$

We define the *value function* of the family of problems by

$$\phi(b) = \sup\{f(x) : x \in \mathbb{R}^n, g(x) = b\}$$

Theorem. (Lagrangian necessity) Suppose for all b there exists an optimal solution x_b^* to problem P_b. If the value function ϕ is concave, then for all b there exists a Lagrange multiplier λ_b^* to problem P_b, that is

$$L_b(x_b^*, \lambda_b^*) \geq L_b(x, \lambda_b^*)$$

for all $x \in \mathbb{R}^n$.

Proof. Suppose that ϕ is concave. Recall that concave functions have the supporting hyperplane property: for fixed b, there exists λ_b^* such that

$$\phi(c) \leq \phi(b) + \lambda_b^* \cdot (c - b)$$

for all c. In other words, the graph of the function ϕ lies beneath the hyperplane tangent to the graph at the point $(b, \phi(b))$. Note[3] that if ϕ happens to be differentiable at b, then we have

$$\lambda_b^* = \nabla \phi(b).$$

Fix b and let x_b^* be the optimal solution of problem P_b. Pick an arbitrary $x \in \mathbb{R}^n$ define c to be $c = g(x)$. In particular, the point x is feasible for problem P_c. Therefore, we have

$$\begin{aligned}
L_b(x_b^*, \lambda_b^*) &= f(x_b^*) \\
&= \phi(b) \\
&\geq \phi(c) + \lambda_b^* \cdot (b - c) \\
&\geq f(x) + \lambda_b^* \cdot (b - c) \\
&= L_c(x, \lambda_b^*) + \lambda_b^* \cdot (b - c) \\
&= L_b(x, \lambda_b^*).
\end{aligned}$$

□

And, finally, here is an easy-to-check condition that the value function ϕ is concave:

Theorem. Suppose f is concave and g is linear. Then then ϕ is concave.

Proof. This is an exercise. □

Back to optimal investment.

Again, consider the problem to

$$\text{maximise } \mathbb{E}\, U(H \cdot P_1) \quad \text{subject to } H \cdot P_0 = X_0.$$

The following gives the basic structural theorem for this one-period problem:

[3] We do not need this fact now, but will will come back to it shortly.

Theorem. Suppose U is concave, everywhere differentiable and that $\mathbb{E}|U(H \cdot P_1)| < \infty$ for all H in a neighbourhood of a point H^*. If H^* is the optimal portfolio, then there exists a real λ such that

$$\mathbb{E}[U'(H^* P_1) P_1] = \lambda P_0.$$

Proof. Consider the value function

$$V(X_0) = \sup\{\mathbb{E}\, U(H \cdot P_1) : H \in \mathbb{R}^n, H \cdot P_0 = X_0\}$$

It is an exercise to show that V is concave. In particular, the Lagrangian necessity theorem applies.

Let

$$L(H, \lambda) = \mathbb{E}\, U(H \cdot P_1) + \lambda(X_0 - H \cdot P_0).$$

The Lagrangian necessity theorem says that there is a λ^* such that

$$L(H^*, \lambda^*) \geq L(H, \lambda^*) \text{ for all } H \in \mathbb{R}^n$$

We now show that $H \mapsto L(H, \lambda^*)$ is differentiable at H^*. Note that for any direction $h \in \mathbb{R}^n$ and $\epsilon > 0$

$$\begin{aligned} 0 &\geq \frac{L(H^* + \epsilon h, \lambda^*) - L(H^*, \lambda^*)}{\epsilon} \\ &= \mathbb{E}\left[\frac{U[(H^* + \epsilon h) \cdot P_1] - U(H^* \cdot P_1)}{\epsilon}\right] - \lambda^* h \cdot P_0 \end{aligned}$$

Notice that the expression inside the square brackets increases pointwise to $h \cdot P_1 U'(H^* \cdot P_1)$ since U is concave[4]. Hence, we may take the limit by the monotone convergence theorem to conclude

$$\nabla_H L(H^*, \lambda^*) = \mathbb{E}[U'(H^* \cdot P_1) P_1] - \lambda^* P_0.$$

However, by Fermat's theorem (i.e. the first order condition for a maximum) we must have

$$\nabla_H L(H^*, \lambda^*) = 0$$

proving the claim. $\square$.

[4]...again, using the fact that the graph of a concave function lies beneath its tangent