Optimal Investment and Insurance

L. C. G. Rogers*
Statistical Laboratory
University of Cambridge

29 June, 2006

1 Introduction

If w_t denotes the assets at time t of an insurance company, then the evolution of w is given by

$$dw_t = r_t(w_t - \theta_t \cdot \mathbf{1})dt + \theta_t \cdot (dS_t/S_{t-}) - \delta_t dt - dC_t + \pi_t dt,$$

where r_t is the riskless rate, θ_t is the portfolio of holdings of the N risky assets, whose prices are $S_t = (S_t^1, \ldots, S_t^N)$ at time t, δ_t is dividend rate at time t, C_t is the cumulative claims by time t, π_t the rate of premium income. Here $\mathbf{1}$ is the N-vector of 1's.

The company controls θ, δ, and in some measure it controls C and π as well (it can refuse, or reinsure, certain risks; it can vary the premia which it charges for cover – higher premiums typically will reduce the volume of business coming in); how should the firm choose these controls?

The answer to this depends on the dynamics of r and S; on the way that π and C respond to pricing decisions, but mainly on the *objective* of the insurance company. Some examples of objectives we might be interested in:

(i) $\mathbb{E}\left[\int_0^\tau e^{-\rho t} U(\delta_t) dt\right]$, where $\tau := \inf\{t : w_t = 0\}$;

(ii) $\mathbb{E}\left[\int_0^\tau e^{-\rho t} U(\delta_t) dt - K e^{-\rho\tau}\right]$

*These notes grew out of a short course given at the Mathematisches Institut, Ludwig-Maximilians Universität, München, 29-30 June 2006. It is a pleasure to thank the organisers, Damir Filipovic and Francesca Biagini, and all the participants for their interest and comments. Special thanks to Yuliya Bregman who prepared the first LaTeXdraft of these notes. The study of examples is far from complete, and if anyone is interested to add others to the collection in Section 3, I would be pleased to hear more.

(iii) $\mathbb{E}\left[\int_0^\tau e^{-\rho t}\{U(\delta_t) - L(\delta_t - \bar{\delta}_t)dt - Ke^{-\rho\tau}\right]$, where $\bar{\delta}_t = \int_{-\infty}^t \lambda e^{\lambda(s-t)}\delta_s ds$

The second can be thought of as the first with a penalty for ruin, which would constrain the firm to behave prudently. The third criterion incorporates a penalty also for dividends which vary too much over time, which would induce the firm to offer a much smoother dividend flow, something which investors appear to like.

Typically in these criteria, we shall think of U as being strictly increasing, strictly concave, though for a profit-maximising firm we shall have U is linear. In the last example, we would have a convex loss function L.

More general objectives (recursive utility, for example) might be considered too, but these will do for the present discussion. The forms assumed are simple enough to allow us to get some quite explicit solutions to these problems, for some quite interesting asset dynamics as well.

Modelling ideas:

(i) The spot rate $(r_t)_{t\geq 0}$ is often taken to be constant, but if we drop this we could allow it to be some simple diffusion process, such as a Vasicek model

$$dr_t = \sigma_s dW_t + \beta(\bar{r} - r_t)dt.$$

Even this simple dynamic must be solved numerically, so we could easily allow more general one-dimensional diffusions for the spot rate, such as the Cox-Ingersoll-Ross model, or the Black-Karasinski model.

(ii) The portfolio process θ may be constrained to take values in a convex set; or may have to be piecewise constant; or changes may incur losses (proportional transaction costs, liquidity costs, for example).

(iii) A common assumption for S is than

$$\frac{dS_t^i}{S_t^i} = \sum_j \sigma_{ij} dW_t^j + \mu_i dt$$

for constants σ_{ij}, μ_i. Sometimes we use log-Lévy dynamics, but this always complicates the analysis considerably, and only occasionally leads to qualitatively novel results. We shall later discuss some interesting variants on these dynamics, where we know σ, but not μ, and have to filter it from the data; and the case where σ, μ are known functions of an underlying Markov chain (Markov-modulated dynamics):

$$dS_t = S_t(\sigma(\xi_t)dW_t + \mu(\xi_t)dt),$$

where ξ is some finite state Markov chain.

(iv) The claims process C_t is frequently modelled as a subordinator, but often it is technically easier to work with

$$C_t = at + bW_t$$

for constants $a > 0$ and $b \in \mathbb{R}$ (the reasoning being that a subordinator less its mean growth rate looks quite like a Brownian motion).
The premium income could b assumed to be constant for a first approximation, but more interestingly is to impose some relationship for the dependence of volume of business on price, so that $\pi = pv(p)$. In this case the Brownian approximation to the claims will be

$$dC_t = \sqrt{v(p_t)} a dW_t + v(p_t) b dt.$$

Types of risk

- Market risk - financial assets do badly;
- Model risk - uncertainty about parameters makes some of your choices bad;
- Regime risk - rates of return on assets may vary with business cycle, mortality may change a lot (AIDS, new treatments, global warming, floods);
- Large loss risk - earthquakes for property, longevity for life.

While the models we look at may be oversimplified, they do give us the possibility of discovering relative magnitudes of the effects studied. This is important, because it helps us to understand what effects matter most; these can then be modelled more realistically as the next stage of our understanding.

2 Optimal investment without insurance

Just to get started, let's look at the optimal investment/consumption in the absence of any insurance component. For simplicity, we will just derive things for the situation with a single risky asset; only the notation gets more complicated with many risky assets. So the wealth dynamics become

$$dw_t = r_t(w_t - \theta_t)dt + \theta_t dS_t / S_{t-} - \delta_t dt.$$

In terms of the bank account numeraire $dB_t = r_t B_t dt$ (so that $B_t = \exp\{\int_0^t r_s ds\}$), we may consider $\tilde{W}_t := W_t / B_t$ and obtain

$$d\tilde{w}_t = \theta_t d\tilde{S}_t / \tilde{S}_{t-} - \tilde{\delta}_t dt,$$

where $\tilde{S}_t = S_t/B_t$, $\tilde{\delta}_t = \delta_t/B_t$. In effect, this reduces to the case where $r \equiv 0$. Suppose now we consider the problem

$$\sup_{\theta,\delta} \mathbb{E}[\int_0^T U(t,\delta_t)dt + u(w_t)]$$

and suppose that this supremum is attained at (θ^*, δ^*). If we perturb to $\theta = \theta^* + \varepsilon\eta$, $\delta = \delta^* + \varepsilon\psi$, then the change in w_T is given by

$$\begin{aligned} \tilde{w_T} &= \int_0^T \theta_t \frac{d\tilde{S}_t}{\tilde{S}_t} - \int_0^T \tilde{\delta}_t dt \\ &= \tilde{w_T^*} + \varepsilon \left\{ \int_0^T \eta_t \frac{d\tilde{S}_t}{\tilde{S}_t} - \int_0^T \tilde{\psi}_t dt \right\}. \end{aligned}$$

The change in the objective to first order in ε is

$$\varepsilon\mathbb{E}\left[\int_0^T U'(t,\delta^*)\psi_t dt + u'(w_T^*)B_T\left(\int_0^T \eta_t \frac{d\tilde{S}_t}{\tilde{S}_t} - \int_0^T \frac{\tilde{\psi}_t}{B_t}dt\right)\right] = 0.$$

Interpreting what this says, since the perturbations ψ, η are (presumably) arbitrary, we must have

$$U'(t,\delta_t^*) = \frac{1}{B_t}\mathbb{E}_t[B_T u'(w_T^*)] \tag{2.1}$$

and that under the measure $\mathbb{Q}$

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} \propto B_t u'(w_T^*) \tag{2.2}$$

the process $\tilde{S}_t$ is a martingale. So we have argued that if we define the state-price density process ζ_t by

$$\zeta_t = U'(t,\delta_t^*) \tag{2.3}$$

then

$$\zeta_t S_t, \ Z_t := \zeta_t B_t \ \text{ are } \mathbb{P}\text{-martingales.}$$

Special case (Merton, complete log-Brownian market): Take the dynamics of S to be

$$dS_t = S_t(\sigma dW_t + \mu dt), \quad \sigma > 0, \mu \in \mathbb{R},$$

with constant $r_t =: r$. In this case, the martingale Z_t has to have the property that $\zeta_t S_t = Z_t B_t^{-1} S_t = Z_t \tilde{S}_t$ is a martingale. But we know that

$$\begin{aligned} d\tilde{S}_t &= \tilde{S}_t(\sigma dW_t + (\mu - r)dt) \\ &= \tilde{S}_t\sigma(dW_t + \kappa dt), \end{aligned}$$

where $\kappa \equiv (\mu - r)/\sigma$, the so-called *Sharpe ratio.* So the change-of-measure martingale Z has to change $dW_t + \kappa dt$ into a martingale. By the Cameron-Martin-Girsanov (CMG) theorem, this gives us

$$dZ_t = Z_t(-\kappa dW_t).$$

So we have explicitly

$$\zeta_t = \exp\{-rt - \kappa W_t - \frac{1}{2}\kappa^2 t\}$$

(or we could equally well work with any positive multiple of this process).

Suppose we now take the objective

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t} U(\delta_t) dt\right]$$

to be maximised with the constraint $w_t \geq 0$ for any t. Our little first-order argument tells us to expect that

$$e^{-\rho t} U'(\delta_t^*) = \lambda \zeta_t$$

for some $\lambda > 0$ fixed, so that

$$\delta_t^* = I(\lambda e^{-\rho t} \zeta_t), \quad I := (U')^{-1}.$$

This gets us a long way, but how do we decide what λ is? We use risk-neutral valuation principle; wealth at time 0 equals the net present value (NPV) of all future consumption, so that

$$w_t = \mathbb{E}_t\left[\int_t^\infty \frac{\zeta_u}{\zeta_t} \delta_u^* du\right] = \zeta_t^{-1} \mathbb{E}_t\left[\int_t^\infty \zeta_u I(\lambda e^{-\rho u} \zeta_u) du\right]$$

which relates λ and w_0. This looks good (and is correct) but there are a couple of outstanding questions:

(i) is there some portfolio that would allow us to consume δ^* while keeping $w \geq 0$?

(ii) Is this optimal?

The answer to the first is "yes": it uses Brownian integral representation result. For the second, we can carry out a direct verification, since we have the solution so explicitly.

A special case: CRRA utility.

This section illustrates the use of the HJB (=Hamilton-Jacobi-Bellman) approach to the Merton problem. Suppose that we have

$$U(x) = \frac{x^{1-R}}{1-R} \tag{2.4}$$

for some $R > 0,\ R \neq 1$[1]. Then $U'(x) = x^{-R},\ \ I(x) = x^{-1/R}$ and we shall have

$$\begin{aligned}
w_0 &= \mathbb{E}\left[\int_0^\infty \zeta_u^{1-1/R} e^{-\rho u/R} \lambda^{-1/R} du\right] \\
&= \lambda^{-1/R}\mathbb{E}\left[\int_0^\infty \exp\{-\frac{\rho u}{R} - (r + \frac{1}{2}\kappa^2)(1-\frac{1}{R})u - (1-\frac{1}{R})\kappa W_u\}\ du\ \right] \\
&= \lambda^{-1/R}\int_0^\infty \exp\{-\frac{\rho u}{R} - (r + \frac{1}{2}\kappa^2)(1-\frac{1}{R})u - \frac{1}{2}(1-\frac{1}{R})^2\kappa^2\}\ du \\
&= \lambda^{-1/R} R\left(\rho + (R-1)\left(r + \kappa^2/2R\right)\right)^{-1} \\
&\equiv \frac{\lambda^{-1/R}}{\gamma},
\end{aligned}$$

where we have defined the constant γ by

$$\gamma = R^{-1}\Big(\ \rho + (R-1)(r + \kappa^2/2R)\ \Big) \tag{2.5}$$

So this links wealth and λ; we have $\lambda = (\gamma w_0)^{-R}$. Doing the similar calculation at time t, we get

$$w_t = \frac{\lambda^{-1/R}(e^{\rho t}\zeta_t)^{-\frac{1}{R}}}{\gamma} = \frac{\delta_t}{\gamma}.$$

The value of the objective is

$$\begin{aligned}
V(w_0) &= \mathbb{E}\left[\int_0^\infty U(\delta_t^*) e^{-\rho t} dt\right] \\
&= \mathbb{E}\left[\int_0^\infty \frac{(\lambda e^{\rho t}\zeta_t))^{1-1/R}}{1-R} e^{-\rho t}\ dt\right] \\
&= \frac{\lambda^{1-1/R}}{1-R}\mathbb{E}\left[\int_0^\infty \exp\{-\rho t/R - (r + \frac{1}{2}\kappa^2)(1-R^{-1})t - \kappa(1-R^{-1})W_t\}\ dt\right] \\
&= \frac{\lambda^{1-1/R}}{1-R}\cdot\frac{1}{\gamma} \\
&= \frac{(\gamma w_0)^{1-R}}{1-R}\cdot\frac{1}{\gamma} = \gamma^{-R} U(w_0).
\end{aligned}$$

Alternative approach: Martingale Principle of Optimal Control

If

$$V(w) = \sup E\left[\int_0^\infty e^{-\rho t} U(\delta_t) dt\ \middle|\ w_0 = w\right],$$

[1]The case $R = 1$ corresponds to log utility, and is handled by similar techniques, though the forms of the answers look quite a bit different.

then $Y_t = \int_0^t e^{-\rho s} U(\delta_s) ds + e^{-\rho t} V(w_t)$ is a supermartingale, and a martingale under optimal control.
Now we have dynamics

$$dw = rwdt + \theta(\sigma dW + (\mu - r)dt) - \delta_t dt,$$

so if we expand Y by Itô's formula,

$$\begin{aligned} dY &= d(\text{loc.mart.}) + \left(U(\delta_t) - \rho V(w_t) + \frac{1}{2}\sigma^2\theta_t^2 V''(w_t) + (rw_t + \theta(\mu - r) - \delta_t)V'(w_t) \right) e^{-\rho t} dt \\ &\equiv d(\text{loc.mart.}) + \Phi dt \end{aligned}$$

Now in order that Y should be a supermartingale, we need that the drift Φ should be non-positive, and for Y to be a martingale under optimal control we shall have to have that the supremum of Φ over choices of the controls should be zero[2]. So

$$\begin{aligned} 0 &= \sup_{\theta,\delta} \left(U(\delta_t) - \rho V(w_t) + \frac{1}{2}\sigma^2\theta_t^2 V''(w_t) + (rw_t + \theta(\mu - r) - \delta_t)V'(w_t) \right) \\ &= \tilde{U}(V') - \rho V + rw_t V' - \frac{1}{2}\kappa^2 \frac{(V')^2}{V''}, \end{aligned} \tag{2.6}$$

where $\tilde{U}$ is the convex dual function $\tilde{U}(y) = \sup\{U(x) - yx\}$. In general, there is no closed-form solution to the HJB equations, but if $U(x) = x^{1-R}/(1-R)$ we have

$$\tilde{U}(y) = -\frac{y^{1-1/R}}{1 - 1/R},$$

and a solution of the form $V(w) = aU(w)$ for some constant a is evident from scaling. Check it out - you get the same.

The Merton problem with CRRA is a baseline example for the whole subject. It is simple, and has a closed-form solution, in which the investor's preference parameters (ρ, R) appear completely explicitly. The form of the solution is simple and intuitive - keep proportion [3] $\pi_M = (\mu - r)/\sigma^2 R$ of your wealth in the risky asset, consume at rate γw. It provides a natural framework for studying the effects of many variants of the basic Merton example. Here's how. If we consider some variant of the Merton problem (for example, with transaction costs) which has a value function v instead of the Merton value function V_M, then we derive the efficiency of the modified problem to be that $\theta > 0$ for which

$$v(w) = V_M(\theta w).$$

This gives us a way to assess the impact of the variation introduced.

[2] These statements are of course only approximately correct - a local martingale can be a supermartingale - but the point is that we are describing a recipe for calculating the solution to the problem. Once we know what that solution is, we have a simple *verification* methodology to confirm that what we believe is optimal actually is.

[3] When there are multiple log-Brownian assets, you keep proportions $\pi_M = R^{-1}(\sigma\sigma^T)^{-1}(\mu - r\mathbf{1})$ in the risky assets.

3 Variations on the basic story.

1. Stochastic interest rates.

Suppose that the riskless rate is not constant, but diffuses as in a Vasicek model:

$$dr = \sigma_r dW' + \beta(\bar{r} - r)dt,$$

where $dW'dW = \eta dt$. What happens? We can derive the HJB equations for

$$V(w, r) = \sup E\left[\int_0^\infty e^{-\rho t} U(\delta_t) dt \;\middle|\; w_0 = w, r_0 = r\right],$$

which must satisfy

$$\begin{aligned} 0 &= \sup_{\delta,\theta}\Big(U(\delta) - \rho V + \frac{1}{2}(\sigma^2\theta^2 V_{ww} + 2\eta\theta\sigma\sigma_r V_{wr} + \sigma_r^2 V_{rr}) \\ &\qquad + (rw + \theta(\mu - r) - \delta)V_w + \beta(\bar{r} - r)V_r \Big) \\ &= \tilde{U}(V_w) - \rho V + rwV_w + \beta(\bar{r} - r)V_r + \frac{1}{2}\sigma_r^2 V_{rr} - \frac{((\mu - r)V_w + \eta\sigma\sigma_r V_{wr})^2}{2\sigma^2 V_{ww}}. \end{aligned}$$

However, by scaling we see we must have

$$V(w, r) = f(r)U(w)$$

for some function f; reworking the HJB equations gives the ODE

$$0 = (1 - R)\tilde{U}(f) - \rho f + r(1 - R)f + \mathcal{L}f + \frac{((\mu - r)f + \eta\sigma\sigma_r f')^2}{2\sigma^2 R f}(1 - R)$$

for f, where $\mathcal{L}$ is the generator of the diffusion for r:

$$\mathcal{L} \equiv \frac{1}{2}\sigma_r^2 \frac{\partial^2}{\partial r^2} + \beta(\bar{r} - r)\frac{\partial}{\partial r}.$$

We can re-express this in the form

$$(\rho + (R - 1)r)f - \mathcal{L}f = \frac{((\mu - r)f + \eta\sigma\sigma_r f')^2(1 - R)}{2\sigma^2 R f} + (1 - R)\tilde{U}(f) \equiv \Psi(f),$$

so iterative solution possible. In more detail, we take some initial guess $f^{(0)}$ for f, and then recursively generate approximations $f^{(n)}$ by the recipe

$$(\rho + (R - 1)r)f^{(n)} - \mathcal{L}f^{(n)} = \Psi(f^{(n-1)}.$$

2. Transaction costs.
Consider the situation where

$$\begin{aligned} dX_t &= rX_t dt + (1-\epsilon)dM_t - (1+\epsilon)dL_t - \delta_t dt \\ dY_t &= Y_t(\sigma dW_t + \mu dt) - dM_t + dL_t, \end{aligned}$$

where X_t is value of holding of cash, Y_t is value of holding of stock at time t. $M_t(L_t)$ the cumulative sales of stock by time t. The investor's goal is to achieve

$$V(x,y) = \sup E\left[\int_0^\infty e^{-\rho t} U(\delta_t) dt \;\middle|\; X_0 = x, Y_0 = y\right],$$

with $U(x) = x^{1-R}/(1-R)$ as in the Merton problem. The HJB equations here give

$$\sup\left[U(\delta) - \rho V + \frac{1}{2}\sigma^2 y^2 V_{yy} + \mu y V_y + (rx - \delta)V_x\right] \le 0, \quad (1-\epsilon)V_x \le V_y \le (1+\epsilon)V_x.$$

We shall once again have scaling, so if we set $V(x,y) = y^{1-R}f(p)$, where $p \equiv x/y$, we can re-express this as

$$\begin{aligned} 0 &= \tilde{U}(f') + \frac{1}{2}\sigma^2 p^2 f"(p) + (\sigma^2 R - \mu + r)pf'(p) + \{\mu(1-R) - \rho - \frac{1}{2}\sigma^2 R(1-R)\}f(p), \\ &\quad (1-\epsilon)f' \le (1-R)f - pf'(p) \le (1+\epsilon)f'. \end{aligned}$$

Alternatively, if we write $f(p) \equiv g(\log(p))$, we simplify the HJB differential operator quite a bit:

$$\begin{aligned} 0 &\ge e^{-t(1-1/R)}\tilde{U}(g'(t)) + a_2 g''(t) + a_1 g'(t) + a_0 g(t) - \rho g(t), \\ 0 &\ge (1 - \varepsilon + e^t)g'(t) - (1-R)e^t g(t), \\ 0 &\ge -(1 + \varepsilon + e^t)g'(t) + (1-R)e^t g(t). \end{aligned}$$

where $t \equiv \log(p)$, and

$$\begin{aligned} a_2 &= \frac{1}{2}\sigma^2, \\ a_1 &= (\sigma^2 R + r - \mu - \frac{1}{2}\sigma^2), \\ a_0 &= (R-1)(\frac{1}{2}\sigma^2 R - \mu). \end{aligned}$$

Constantinides solves a simplified form of this problem, and Davis & Norman analyse it quite completely. The main conclusion is that there is some interval $K = [t_s, t_b]$ for t such that while t remains within $[t_s, t_b]$, you make no change in your portfolio; if ever $t < t_s$ you immediately sell enough stock to move back into the interval K, and if ever

$t > t_b$ you immediately buy sufficient stock to move t back into the interval K. No closed-form solution is known, but Davis & Norman show how the ODE for g may be solved by iteratively solving the ODE with different initial conditions until the solution closes in on one which satisfies the C^2 pasting condition at the ends of K.

The main things we need to note are

(i) the form of the solution;

(ii) the fact that the loss of efficiency is $O(\epsilon^{\frac{2}{3}})$ - see (4), (2).

This last tells us that when we consider typical values for the transaction cost (of the order of 1% or less), the impact on efficiency will be *small*.

3. Parameter uncertainty.

This example and the next are discussed in more detail in (1).

The 20's example. This little example, which requires no more than an understanding of basic statistical concepts, should be remembered by anyone who works in finance.

Suppose we consider a stock, with annualised rate of return $\mu = 0.2$, and annualised volatility $\sigma = 0.2 = 20\%$. We see daily prices for N years, and we want to observe for long enough that our 95% confidence interval for σ (respectively, μ) is of the form $[\hat{\sigma} - 0.01, \hat{\sigma} + 0.01]$ (respectively, $[\hat{\mu} - 0.01, \hat{\mu} + 0.01]$) - so we have a 19 in 20 chance of knowing the true value to one part in 20.

How big must N be to achieve this precision in $\hat{\sigma}$?

Answer: about 11 years;

How big must N be to achieve this precision in $\hat{\mu}$?

Answer: about 1580 years !!

The most important thing to know about the rate of growth of a stock is that *we know almost nothing about it!* When it comes to the Merton problem, we invest a fixed proportion of our wealth in the risky asset, but that proportion depends on the rate of growth parameter μ, which we do not know ... So while the analysis of the Merton problem is correct, the model assumptions do not fit reality well.

Instead, let us suppose we have wealth dynamics

$$dw_t = r(w_t - \theta_t)dt + \theta_t \sigma(dW_t + \alpha dt) - \delta_t dt,$$

where $\alpha = \sigma\mu$ is not known, but σ is. We propose a $N(\hat{\alpha}, \tau_0^{-1})$ prior for α, and try to filter α from the observations. Write

$$X_t = W_t + \alpha t = \sigma^{-1}\{\log(S_t/S_0) + \frac{1}{2}\sigma^2 t\},$$

so that we see X, and must filter α from that. According to CMG, the likelihood of $(X_s)_{0\leq s\leq t}$ if α is the true value is

$$\exp\{\alpha X_t - \frac{1}{2}\alpha^2 t\},$$

so the posterior for α is proportional to

$$\exp\{-\frac{1}{2}\tau_0(\alpha-\hat{\alpha}_0)^2 + \alpha X_t - \frac{1}{2}\alpha^2 t\} \propto \exp\left\{-\frac{1}{2}(\tau_0+t)\left(\alpha - \frac{\hat{\alpha}_0\tau_0 + X_t}{\tau_0+t}\right)^2\right\}.$$

So the posterior for α given $(X_s)_{0\leq s\leq t}$ is $N(\hat{\alpha}_t, \tau_t^{-1})$, where $\tau_t = \tau_0 + t$, $\hat{\alpha}_t = (\hat{\alpha}_0\tau_0 + X_t)/(\tau_0+t)$. Standard results from filtering theory (see, for example, VI.8 in (3)) tell us that

$$\begin{aligned} dX_t &= dW_t + \alpha dt = d\hat{W}_t + \hat{\alpha}_t dt = d\hat{W}_t + \frac{\hat{\alpha}\tau_0 + X_t}{\tau_0+t}dt, \\ d\hat{\alpha}_t &= \frac{d\hat{W}_t}{\tau_0+t}, \end{aligned}$$

so the dynamics of X_t are different in the observation filtration. For this situation, probably the best thing to do is to work with state-price density approach. As we've seen, we relate the state-price density to the optimal consumption rate process by

$$e^{-\rho t}U'(\delta_t) = \lambda\zeta_t,$$

and this then has the property that $\zeta_t S_t$ is a martingale. Now

$$dS_t = \sigma S_t dX_t = \sigma S_t\left(d\hat{W}_t + \hat{\alpha}_t dt\right),$$

so the change-of-measure martingale is

$$\exp\left[\int_0^t (\sigma^{-1}r - \hat{\alpha}_s)^2 d\hat{W}_s - \int_0^t (\sigma^{-1}r - \hat{\alpha}_s)^2 ds\right]$$

and we are able to derive after some calculations the expression

$$\zeta_t = \left(\frac{\tau_t}{\tau_0}\right)^{\frac{1}{2}} e^{-rt}\exp\left\{-\frac{1}{2}m_t^2\tau_t + \frac{1}{2}m_0^2\tau_0\right\}, \quad m_t = \hat{\alpha}_t - \sigma^{-1}r.$$

We therefore conclude that

$$\delta_t^* = I(\lambda\zeta_t e^{\rho t}),$$

where λ is fixed by the budget constraint

$$w_0 = \mathbb{E}\left[\int_0^\infty \zeta_t \delta_t^* dt\right] = \lambda^{-1/R}\mathbb{E}\left[\int_0^\infty e^{-\rho t/R}\zeta_t^{1-1/R} dt\right] = \lambda^{-1/R}\phi(\hat{\alpha}_0, \tau_0)$$

and the optimised objective is

$$\begin{aligned}\mathbb{E}\left[\int_0^\infty e^{-\rho t} U(\delta_t^*) dt\right] &= \mathbb{E}\left[\int_0^\infty \frac{1}{1-R} e^{-\rho t/R}\lambda^{1-1/R}\zeta_t^{1-1/R} dt\right] \\ &= \frac{\lambda^{1-1/R}}{1-R}\,\phi(\hat{\alpha}_0, \tau_0) \\ &= u(w_0)\phi(\hat{\alpha}_0, \tau_0)^R.\end{aligned}$$

Numerical examples:. . .

4. Infrequent portfolio rebalancing.

Suppose we have a standard Merton investor, maximizing

$$\mathbb{E}[\int_0^\infty e^{-\rho t} U(\delta_t) dt], \quad U'(x) = x^{-R}$$

and now instead of rebalancing continuously, we only allow the agent to rebalance the portfolio at times $t = 0, h, 2h, \ldots$. Obviously the agent does less well, but does this actually matter? Let's also suppose the rate δ is held constant in each interval, so the value solves

$$V(w) = \sup_{\delta, p}\left[\, U(\delta)\, \frac{1-e^{-\rho h}}{\rho} + e^{-\rho h}\mathbb{E}V\left((w - \delta h)(pS + (1-p)e^{rh})\right)\right],$$

where $S = \exp\{\sigma\sqrt{h}Z + (\mu - \frac{1}{2}\sigma^2)h\}$ with $Z \sim N(0,1)$. Now scaling tells us that for some constant a, $V(w) = aU(w)$, so we get

$$\begin{aligned}\frac{a}{1-R} &= \sup_{t,p}\left[\frac{t^{1-R}}{1-R}\frac{1-e^{-\rho h}}{\rho} + e^{-\rho h} a\mathbb{E}\frac{(1-th)^{1-R}}{1-R}(pZ + (1-p)e^{rh})^{1-R}\right] \\ &= \sup_t \left[\tilde{h}\, \frac{t^{1-R}}{1-R} + ae^{-\rho h}(1-th)^{1-R} K\right],\end{aligned}$$

where $K = (1-R)^{-1}\sup_p \mathbb{E}(pS + (1-p)e^{rh})^{1-R}$, $\tilde{h} = \frac{1-e^{-\rho h}}{\rho}$, and therefore we can maximize explicitly. Routine calculations lead us to

$$a^{1/R} = \frac{h(\tilde{h}/h)^{1/R}}{1 - (K(1-R)e^{-\rho h})^{1/R}}.$$

Some numerical values. . .

5. Optimisation under drawdown constraints.

In this problem, we assume the (by now) standard dynamics

$$dw_t = r(w_t - \theta_t)dt + \theta_t(\sigma dW_t + \mu dt) - \delta_t dt$$

for the wealth and objective

$$\sup \mathbb{E}[\int_0^\infty e^{-\rho t} U(\delta_t) dt], \quad U'(x) = x^{-R},$$

but now we shall impose the constraint

$$w_t \geq b\bar{w}_t = b \sup_{s \leq t} w_s, \quad \forall t, \tag{3.7}$$

where $b \in (0,1)$ is fixed. This is called a *drawdown constraint*, in a natural terminology. Drawdown constraints are of practical importance for fund managers, because if their portfolio loses too much of its value, the investors are likely to take their money out and that is the end of the story, however clever (or even optimal!) the rule being used by the fund manager. For this problem, the value function

$$V(w, \bar{w}) = \sup \mathbb{E}[\int_0^\infty e^{-\rho t} U(\delta_t)\, dt \,\Big|\, w_0 = w, \bar{w}_0 = \bar{w}]$$

evidently scales like

$$V(w, \bar{w}) = \bar{w}^{1-R} V(w/\bar{w}, 1) = \bar{w}^{1-R} v(w/\bar{w}) = \bar{w}^{1-R} v(x), \quad x = w/\bar{w} \in [b, 1].$$

So the HJB equation here is

$$\sup_{\delta,\theta} \Big[U(\delta) - \rho V + \frac{1}{2}\sigma^2\theta^2 V_{ww} + (r(w-\theta) + \mu\theta - \delta)V_w \Big] = 0$$

with the boundary condition that $V_w = 0$ at $w = \bar{w}$. So HJB is

$$\tilde{U}(V_w) - \rho V + rwV_w - \frac{1}{2}\kappa^2 \frac{V_w^2}{V_{ww}} = 0,$$

where as before $\kappa = (\mu - r)/\sigma$, and in terms of v this gives

$$\tilde{U}(v') - \rho v + rxv' - \frac{1}{2}\kappa^2 \frac{(v')^2}{v''} = 0, \tag{3.8}$$

$$(1-R)v(1) = v'(1) \tag{3.9}$$

(indeed, $(1-R)v(x) - xv'(x) \leq 0$ always, with equality when $x \geq 1$). The boundary condition at 1 can be understood as saying that we extend v to $(1,\infty)$ by $v(x) = x^{1-R}v(1)$ $(x \geq 1)$, and this extension is C^1.

Now the ODE (3.8) is highly non-linear, and yet *it is possible to linearize it, and solve explicitly!*

The trick is to make

$$z \equiv v'(w)$$

the new variable, and

$$J(z) = v(w) - wz$$

the new function. Then as a little calculus confirms, we have

$$J'(z) = -w, \quad J''(z) = -1/v''(w),$$

and now (3.8) becomes simply

$$\tilde{U}(z) + \frac{1}{2}\kappa^2 z^2 J'' + (\rho - r)zJ' - \rho J = 0, \tag{3.10}$$

$$-(1 - \frac{1}{R})J(z) + zJ'(z) \le 0, \tag{3.11}$$

with equality in (3.11) when $J'(z) \le -1$.

One other observation is required: as $w \downarrow b\bar{w}$, the portfolio weight $\theta \to 0$, because otherwise at the boundary the constraint (3.7) would get violated. But recall that the optimal portfolio is

$$\theta = \frac{(\mu - r)V_w}{\sigma^2 V_{ww}};$$

this implies that $v''(b) = +\infty$, $J''(v'(b)) = 0$. Thus there exist $z_b = v'(b) > z_1 = v'(1)$ such that the solution J has the form

$$J(z) = \begin{cases} A_0 \tilde{U}(z) & \text{for } z \le z_1; \\ A_1 (z/z_b)^{-\alpha} + B_1 (z/z_b)^{\beta} + q\tilde{U}(z) & \text{for } z_1 \le z \le z_b; \\ q\tilde{U}(z_b) + A_1 + B_1 + b(z_b - z) & \text{for } z \ge z_b \end{cases}$$

where $q = -1/Q(1 - R^{-1})$, and $Q(t) \equiv \frac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho$ is the quadratic whose roots are $-\alpha < 0 < \beta$. In order that the problem is well posed, it is necessary and sufficient that $q > 0$. The constants A_0, A_1, B_1, z_1, and z_b are to be determined from the conditions

(i) J is C^2 at z_b;

(ii) J is C^1 at z_1.

Thus if we pick z_b, we know that $J'(z_b) = -b$, $J''(z_b) = 0$, so the ODE (3.10) gives us

$$\rho J(z_b) = -(\rho - r)z_b b + \tilde{U}(z_b).$$

We also have the condition that $J'(z_1) = -1 = -A_0 z_1^{-1/R}$, giving us the relation $z_1 = A_0^R$. Using these conditions it is not too hard to find (numerically) the solution J, and hence the original value function v.

6. Optimisation under Markov-modulated dynamics.

This time, the wealth dynamics are given by

$$dw_t = r(\xi_t) w_t dt + \theta_t(\sigma(\xi_t) dW_t + (\mu(\xi_t) - r(\xi_t))dt) - \delta_t dt,$$

where ξ is a finite state irreducible Markov chain with generator Q. The value function

$$V(w, \xi) = \sup \mathbb{E}[\int_0^\infty e^{-\rho t} U(\delta_t) dt \;\Big|\; w_0 = w, \xi_0 = \xi]$$

is of the form $V(w, \xi) = f(\xi) U(w)$, by the scaling properties again, and the HJB equations are

$$\begin{aligned} 0 &= \sup_{\delta,\theta} \left[U(\delta) - \rho V + \frac{1}{2}\sigma^2\theta^2 V_{ww} + (rw + \theta(\mu - r) - \delta)V_w + QV \right] \\ &= \tilde{U}(V_w) - \rho V + rwV_w - \frac{1}{2}\kappa^2 \frac{V_w^2}{V_{ww}} + U(w)Qf \end{aligned}$$

from we which we deduce the equations for f:

$$\begin{aligned} 0 &= Rf^{1-1/R} - \rho f + r(1-R)f + \frac{\kappa^2}{2R}(1-R) + Qf \\ &= Rf^{1-1/R} - R\Gamma f + Qf, \end{aligned}$$

where $\Gamma(\xi) = R^{-1}(\rho + (R-1)(r(\xi) + \frac{\kappa(\xi)^2}{2R}))$. We therefore have to solve

$$(R\Gamma - Q)f = Rf^{1-1/R}.$$

This can again be done recursively, by setting $f^{(0)} = 1$,

$$(R\Gamma - Q)f^{n+1} = R(f^{(n)})^{1-1/R}.$$

The efficiency is then

$$(\gamma^R f)^{1/(1-R)}.$$

7. An example related to insurance.

At last we reintroduce the premium terms in the wealth dynamics which we have so far omitted, and consider a very simple model for the evolution of the wealth of an insurance company. We suppose that the wealth dynamics are

$$dw_t = rw_t dt + \theta(\sigma dW_t + (\mu - r)dt) - \delta_t dt - kdt, \tag{3.12}$$

where the outflow $k > 0$, constant, represents payments to policyholders less premium income – too simple, but at least a place to start. The objective is

$$\max \mathbb{E}^w \left[\int_0^\tau e^{-\rho t} U(\delta_t) dt - Ke^{-\rho\tau} \right] = V(w),$$

where $K > 0$ is a penalty for the firm going broke, at time $\tau = \inf\{t : \ w_t = 0\}$. This may be infinite of course. Varying K allows us to impose different degrees of security on the firm. The HJB equation for this problem is

$$\begin{aligned} \max_{\theta,\delta} \left[U(\delta) - \rho V + \frac{1}{2}\sigma^2\theta^2 V_{ww} + \{rw + (\mu - r)\theta - \delta - k\} V_w \right] &= 0, \\ V(0) &= -K \end{aligned}$$

and as before the optimal dividend and investment in the risky asset are given by

$$U'(\delta) = V_w, \quad \theta = -\frac{\mu - r}{\sigma^2}\frac{V_w}{V_{ww}} = -\frac{\kappa}{\sigma}\frac{V_w}{V_{ww}},$$

leading to the HJB differential equation

$$\tilde{U}(V_w) - \rho V + (rw - k)V_w - \frac{1}{2}\frac{\kappa^2 V_w^2}{V_{ww}} = 0.$$

Now we use the dual variables trick: $z = V'(w)$, $J(z) = V(w) - wz$, giving

$$J' = -w, \quad V(w) = J(z) - zJ'(z), \quad V''(w) = -1/J''(z),$$

which linearises the HJB equation to

$$0 = \tilde{U}(z) + \frac{1}{2}\kappa^2 z^2 J''(z) + (\rho - r)zJ' - \rho J - kz$$

which we can solve explicitly! The general solution of the homogeneous equation is of the form $Az^{-\alpha} + Bz^{\beta}$, where $-\alpha < 0 < \beta$ are the roots of $Q(t) = \frac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho$, and the particular solution is $-kz/r + q\tilde{U}(z)$, where $q = -1/Q(1 - R^{-1}) > 0$ (this is the condition for the problem to be well posed, and in fact $Q(1 - R^{-1}) = -\gamma$), so we have a solution of the form

$$J(z) = -\frac{kz}{r} + \frac{1}{\gamma}\tilde{U}(z) + Az^{-\alpha} + Bz^{\beta}, \tag{3.13}$$

at least for $z \le z_* = V'(0)$; for $z \ge V'(0)$, we have $J(z) = -K$. There will be C^1 contact of J to $-K$ at $z = z_*$. Now the problem is well posed if and only if $Q(1 - 1/R) < 0$, which is easily seen on consideration of the form of the quadratic Q to be equivalent to $-\alpha < 1 - 1/R$. Hence the dominant term in (3.13) is the term $Az^{-\alpha}$, where $A \ge 0$ to ensure the convexity of J. But if $A > 0$, we would have that V grows faster at infinity than w^{1-R}, and this is impossible, because the value for this insurance problem cannot be greater than the value for the standard Merton problem where there is no downward outflow k. The only possibility is therefore that $A = 0$, and J has the form

$$J(z) = \begin{cases} -\frac{kz}{r} + \frac{1}{\gamma}\tilde{U}(z) + B(\frac{z}{z_*})^{\beta} & \text{if } z \le z_* \\ -K & \text{if } z \ge z_*. \end{cases}$$

So matching the slope at ξ_* gives us

$$0 = -\frac{k}{r} - \frac{1}{\gamma} z_*^{-1/R} + \frac{B^\beta}{z_*} \quad \Rightarrow \quad B = \frac{1}{\beta}\left[\frac{k z_*}{r} + \frac{1}{\gamma} z_*^{1-1/R}\right].$$

If we select z_*, this tells us B, hence $J(z_*)$, so we now just need to find the z_* to make $J(z_*) = -K$.

8. Annual tax accounting.

What is the effect on the Merton problem of an annual tax on capital gains? Suppose that U is again CRRA, and at each time $t = nh$ we have to pay tax on wealth gain over the last time period of length h. Thus $w_{nh} = w_{nh-} - \tau(w_{nh-} - w_{nh-h}) = (1-\tau)w_{nh-} + \tau w_{nh-h}$. If we do this, then the problem becomes a finite-horizon problem,

$$V(w) = \sup \mathbb{E}\left[\int_0^h e^{-\rho s} U(\delta_s) ds + e^{-\rho h} U(\tau w + (1-\tau) w_h)\right].$$

Clearly by scaling again, there is some positive constant A such that $V(w) = AU(w)$, so we have to consider

$$\sup \mathbb{E}\left[\int_0^h e^{-\rho s} U(\delta_s) ds + A e^{-\rho h} U(\tau w + (1-\tau) w_h)\right].$$

As we saw in Section 2, by (2.1) the optimal terminal wealth w_h^* and running consumption δ^* are related to the state-price density process ζ by

$$e^{-\rho t} U'(\delta_t^*) = e^{-rt} Z_t = \lambda \zeta_t, \quad Z_t = \mathbb{E}_t[e^{rh} A e^{-\rho h}(1-\tau) U'(\tau w + (1-\tau) w_h^*)],$$

where $\zeta_t = \exp\{-rt - \kappa W_t - \frac{1}{2}\kappa^2 t\}$ is the SPD, $\zeta_0 = 1$. We deduce that $\delta_t^* = I(\lambda e^{\rho t} \zeta_t)$ and

$$\lambda \zeta_h = e^{-rh} Z_h = A e^{-\rho h}(1-\tau) U'(\tau w + (1-\tau) w_h^*);$$

rearranging to make w_h^* the subject of the equation gives us

$$w_h^* = \frac{1}{1-\tau}\left\{-\tau w + I\left(\frac{\lambda e^{\rho h} \zeta_h}{A(1-\tau)}\right)\right\}.$$

We now need to relate λ to initial wealth w:

$$\begin{aligned} w &= \mathbb{E}\left[\int_0^h \zeta_u \delta_u^* du + \zeta_h w_h^*\right] \\ &= \mathbb{E}\left[\int_0^h \zeta_t^{1-1/R} \lambda^{-1/R} e^{-\rho t/R} dt - \zeta_h \frac{\tau w}{1-\tau} + \frac{\zeta_h^{1-1/R}}{1-\tau} \lambda^{-1/R} e^{-\rho h/R} A^{1/R} (1-\tau)^{1/R}\right] \\ &= -\frac{\tau w e^{-rh}}{1-r} + \lambda^{-1/R} \frac{1-e^{-\gamma h}}{\gamma} + \lambda^{-1/R} A^{1/R} (1-\tau)^{1/R-1} e^{-\gamma h}. \end{aligned}$$

Thus

$$w\left(1+\frac{\tau e^{-\tau h}}{1-\tau}\right)=\lambda^{-1/R}\left(\frac{1-e^{-\gamma h}}{\gamma}+A^{1/R}(1-\tau)^{1/R-1}e^{-\gamma h}\right). \tag{3.14}$$

Now we need to compute the value,

$$\begin{aligned}
V(w) &= \mathbb{E}\left[\int_0^h e^{-\rho t}U(\delta_t^*)dt + Ae^{-\rho h}U(\tau w+(1-\tau)w_h^*)\right] \\
&= \mathbb{E}\left[\int_0^h e^{-\rho t}\frac{(\lambda e^{\rho t}\zeta_t)^{1-1/R}}{1-R}dt + \frac{Ae^{-\rho h}}{1-R}\left(\frac{\lambda e^{\rho h}\zeta_h}{A(1-\tau)}\right)^{1-1/R}\right] \\
&= \frac{\lambda^{1-1/R}}{1-R}\mathbb{E}\left[\int_0^h e^{-\rho t/R}\zeta_t^{1-1/R}dt + A^{1/R}e^{-\rho h/R}(1-\tau)^{1/R-1}\zeta_h^{1-1/R}\right] \\
&= \frac{\lambda^{1-1/R}}{1-R}\left(\frac{1-e^{-\gamma h}}{\gamma}+A^{1/R}(1-\tau)^{1/R-1}e^{-\gamma h}\right) \qquad (3.15)
\end{aligned}$$

Now from the equation (3.14), $\lambda^{-1/R}=Bw/K$, where $B=1+\tau e^{-rh}/(1-\tau)$ and $K=\gamma^{-1}(1-e^{-\gamma h})+A^{1/R}(1-\tau)^{1/R-1}e^{-\gamma h}$, so we have that $\lambda=(Bw/K)^{-R}$, and from (3.15) we deduce that

$$V(w)=U(w)\left(\frac{B}{K}\right)^{1-R}K=U(w)B^{1-R}K^R=AU(w).$$

This implies that

$$A^{1/R}=KB^{1/R-1}=B^{(1-R)/R}\left(\frac{1-e^{-\gamma h}}{\gamma}+A^{1/R}(1-\tau)^{\frac{1-R}{R}}e^{-\gamma h}\right).$$

We can now make $A^{1/R}$ the subject of this equation:

$$A^{1/R}=\frac{\gamma^{-1}(1-e^{-\gamma h})B^{1/R-1}}{1-e^{-\gamma h}(((1-\tau)B)^{1/R-1})},$$

expressing A (and hence the value) explicitly in terms of the variables of the problem. The efficiency can now be expressed explicitly as

$$\theta=(A\gamma^R)^{1/(1-R)}.$$

We are now able compute numerical values quite explicitly.

References

[1] L. C. G. Rogers. The relaxed investor and parameter uncertainty. *Finance & Stochastics*, 5:131–154, 2001.

[2] L. C. G. Rogers. Why is the effect of proportional transactions costs $O(\delta^{2/3})$? In G. Yin & Q. Zhang, editor, *Mathematics of Finance*, volume 351, pages 303–308. American Mathematical Society Contemporary Mathematics Series, 2004.

[3] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales, Vol. 2*. Cambridge University Press, Cambridge, 2000.

[4] S. E. Shreve. Liquidity premium for capital asset pricing with transaction costs. In M. H. A. Davis, D. Duffie, W. H. Fleming, and S. E. Shreve, editors, *Mathematical Finance*, number 65 in IMA Volumes in Mathematics and its Applications. Springer, 1995.