Non-expected utility preferences (15/1/11)

(1) Machina presents various instances and examples where the standard expected utility paradigm appears to be violated. He proposes that preferences over distributions of a real-valued random outcome should be ordered according to a function $V(F)$ which is suitably continuous. He shows that if the measures are concentrated on a fixed compact interval, then weak convergence is equivalent to $\int |F_n(x) - F(x)|\, dx \to 0$, so there is a Banach space structure, and he requires $V(\cdot)$ to be Fréchet differentiable. He also proposes a simple form of $V(\cdot)$ which is quadratic.

(2) Continuity of preferences and the independence axiom [that is, $F_1 \succsim F_2$ iff $pF_1 + (1-p)G \succsim pF_2 + (1-p)G$ for any $0 \le p \le 1$, any G] give an expected utility form. I was thinking that the following is a pretty natural property:

Axiom A: If $F \simeq G$ (neither F nor G is preferred, they are equally good) then $pF + (1-p)G \simeq F$.

This seems reasonable: you tell me you don't mind which of F, G you receive, so why should you care if I use some randomization to decide which to give you?!

Together with weak continuity of $V(\cdot)$ this implies (and is implied by)

Axiom B: If $F \succsim G$, then for all $0 \le t \le 1$, $tF + (1-t)G \succsim G$, $F \succsim tF + (1-t)G$

(Starmer calls this "betweenness", an ugly term).

(3) It appears that Axiom A does in effect force EU preferences. Let's look at the case where there are only $N+1$ possible outcomes, and we characterise dist^{ns} by a point $p \in \{(p_1, \dots, p_N) : p_j \ge 0, \sum p_j \le 1\} \equiv \mathcal{D}$ in $\mathbb{R}^N$. Suppose that $\delta_1 \succsim \delta_2 \succsim \dots \succsim \delta_{N+1}$. Then for each $j = 2, \dots, N$ there is some $\theta_j \in [0,1]$ such that $\delta_j \simeq \theta_j \delta_1 + (1-\theta_j)\delta_{N+1}$, by continuity.

Suppose that $\bar{p} \in \mathrm{int}(\mathcal{D})$, and consider the subspace in $\mathbb{R}^N$, Y, say, orthogonal to $\nabla V(\bar{p}) \equiv b$. Now consider

$$H_\pm = \{ p \in \mathcal{D} : b \cdot (p - \bar{p}) \gtrless 0 \},$$

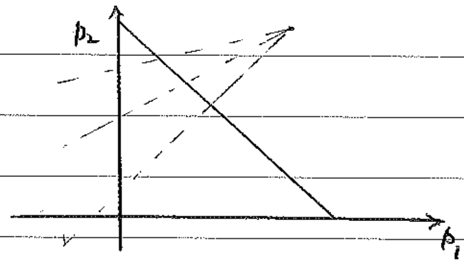
nonempty convex sets. Suppose there were some $\tilde{p} \in H_+$ where $V(\tilde{p}) = V(\bar{p})$. By considering the line segment $t\tilde{p} + (1-t)\bar{p}$, we see that all points in that line segment have the same desirability, so

$$0 = \frac{d}{dt} V(t\tilde{p} + (1-t)\bar{p}) = (\tilde{p} - \bar{p}) \cdot \nabla V(t\tilde{p} + (1-t)\bar{p}).$$

By taking $t=0$ we get $(\tilde{p} - \bar{p}) \cdot b = 0$, contradicting $\tilde{p} \in H_+$. Hence we must have $V(\tilde{p}) > V(\bar{p})$ for $\tilde{p} \in H_+$, similarly $V(\tilde{p}) < V(\bar{p})$ for $\tilde{p} \in H_-$, and by continuity $V(\tilde{p}) = V(\bar{p})$ if $(\tilde{p} - \bar{p}) \cdot b = 0$.

So axiom A forces the level sets of V to be hyperplanes. However this isn't quite EU. Think of the case where $N=2$ and $\mathcal{D}$ looks like:

We could have indifference curves which fan out:



$$Q_t \equiv \exp\left(-\frac{1}{2\sigma^2}\sum_{i=0}^{t} x_i^2\right) \sigma^{-2(t+1)/2}.$$

Equilibria involving different memory (19/1/11)

1) Suppose that $X_t \sim N(\mu_t, \sigma^2)$ where in the reference measure $\mathbb{P}^0$, $\mu_t = 0$ for all t, whereas an agent thinks that in fact the μ_t follow a Gaussian zero-mean random walk with variance σ_μ^2: $\mu_t - \mu_{t-1}$ are IID $N(0, \sigma_\mu^2)$. In this story, we have that if

$$(\mu_t | \mathcal{F}_t) \sim N(m_t, v_t)$$

then

$$\begin{pmatrix} \mu_{t+1} \\ X_{t+1} \end{pmatrix} \Big| \mathcal{F}_t \sim N\left(\begin{pmatrix} m_t \\ m_t \end{pmatrix}, \begin{pmatrix} v_t + \sigma_\mu^2 & v_t + \sigma_\mu^2 \\ v_t + \sigma_\mu^2 & v_t + \sigma_\mu^2 + \sigma^2 \end{pmatrix} \right)$$

so that

$$(\mu_{t+1} | \mathcal{F}_{t+1}) \sim N\left(\underbrace{\beta_t m_t + (1-\beta_t)(X_{t+1} - m_t)}_{= m_{t+1}}, \frac{(v_t + \sigma_\mu^2)\sigma^2}{v_t + \sigma_\mu^2 + \sigma^2} \right)$$

where $1-\beta = (v_t + \sigma_\mu^2)/(v_t + \sigma_\mu^2 + \sigma^2)$. Let's suppose we've got to steady state so that v_t, β_t don't depend on t.

Thus the likelihood of $(x_0, \ldots, x_t)$ for this agent can be expressed as

$$\prod_{j=0}^{t} \exp\left\{ -\frac{1}{2(\sigma^2+v)} (x_j - m_{j-1})^2 \right\} (\sigma^2+v)^{-1/2}$$

Let's write $\bar{v} \equiv \sigma^2 + v$.

2) Now let's suppose we have CARA agents with common impatience ρ and coefficients γ of ARA. These agents have different ideas what σ_μ^2 should be; agent j ends up with (steady state) $\bar{v}_j$ and β_j. Now

$$\Lambda_t^j = \bar{v}_j^{-(t+1)/2} \exp\left[-\frac{1}{2\bar{v}_j} \sum_{i=0}^{t} (x_i - m_{i-1}^j)^2 \right] / Q_t$$

Let's notice

$$x_t - m_{t-1} = x_t - \sum_{n=0}^{t-1} \beta^n (1-\beta) x_{t-n-1} - \beta^t m_{-1}$$

$$= \sum_{n\geq 0} \beta^n (1-\beta)(x_t - x_{t-n-1}) \qquad (x_k = m_{-1} \ \forall k \leq -1.)$$

$$= \sum_{n\geq 0} \beta^n (1-\beta) \sum_{s=t-n}^{t} \Delta x_s$$

$$= \sum_{s \leq t} \Delta x_s \, \beta^{t-s}$$

$$\exp\left[-\Gamma x - \sum \frac{p_j}{2v_j}(x-m_j)^2\right]$$

$$= \exp\Big[-\Gamma x - \tfrac{1}{2}x^2 \underbrace{\sum p_j/v_j}_{=1/V} + x\underbrace{\Big(\sum \frac{p_j}{v_j} m_j\Big)}_{=\bar{m}/V} - \tfrac{1}{2}\sum p_j m_j^2/v_j\Big]$$

$$= \exp\left[-\frac{x^2}{2V} + \frac{x}{V}(\bar{m} - \Gamma V) - \tfrac{1}{2}\sum p_j m_j^2/v_j\right]$$

$$= \exp\left\{-\frac{1}{2}\frac{(x-\bar{m}+\Gamma V)^2}{V} + \frac{(\bar{m}-\Gamma V)^2}{2V} - \tfrac{1}{2}\sum p_j m_j^2/v_j\right\}$$

$$= \exp\left\{-\frac{(x-\bar{m}+\Gamma V)^2}{2V} - \tfrac{1}{2}\Big(\sum \pi_j m_j^2 - \big(\sum \pi_j m_j\big)^2\Big) - \bar{m}\Gamma + \tfrac{1}{2}\Gamma^2 V\right\}$$

$$= \exp\left\{-\frac{(x-\bar{m}+\Gamma V)^2}{2V} - \tfrac{1}{2}\sum \pi_j (m_j-\bar{m})^2 - \bar{m}\Gamma + \tfrac{1}{2}\Gamma^2 V\right\}$$

which might be useful

Solving for the equilibrium leads to

$$-\rho t - \gamma_j c_t^j + \log \Lambda_t^j = \alpha_j + \log \zeta_t$$

and market clearing gives us

$$X_t = \sum_j c_t^j = -\rho t \sum \gamma_j^{-1} + \sum \gamma_j^{-1} \log \Lambda_t^j - \sum \gamma_j^{-1} \log \zeta_t + \text{const}$$

$$\Rightarrow \Gamma X_t = -\rho t + \sum \Gamma \gamma_j^{-1} \log \Lambda_t^j - \log \zeta_t$$

where $\Gamma^{-1} \equiv \sum \gamma_j^{-1}$. This gives the SPD

$$\boxed{\zeta_t = \exp\left(-\rho t - \Gamma X_t + \sum p_j \log \Lambda_t^j\right)} \qquad p_j \equiv \Gamma / \gamma_j$$

3) We may be able to get something out of this, such as the price of a 1-period bond. We need to calculate

$$\mathbb{E}_t^0\left[\zeta_{t+1}/\zeta_t\right] = \kappa e^{-\rho}\, \mathbb{E}_t^0\left[e^{-\Gamma X_t} \exp\left(-\sum_j \frac{p_j}{2v_j}(x_{t+1} - m_t^j)^2 + \frac{x_{t+1}^2}{2\sigma^2}\right)\right]$$

$$\left[\kappa \equiv \sigma \prod v_j^{-p_j/2}\right]$$

$$= \int \exp\left\{-\Gamma x - \sum_j \frac{p_j}{2v_j}(x - m_t^j)^2\right\} \frac{\kappa\, dx}{\sigma\sqrt{2\pi}}\, e^{-\rho}$$

$$= \frac{\kappa}{\sigma}\sqrt{V} \exp\left[-\frac{1}{2V}\sum \pi_j (m_t^j - \bar{m}_t)^2 - \bar{m}_t \Gamma + \frac{\Gamma^2 V}{2}\right] e^{-\rho}$$

where $V \equiv \left(\sum p_j / v_j\right)^{-1}$, $\frac{\bar{m}}{V} = \sum_j \frac{p_j}{v_j} m_j \equiv \sum_j \pi_j m_j / V$.

This is reasonably explicit, but may still not be easy to work with.

4) Another point which comes from this analysis; in the risk-neutral measure

$$\boxed{\mathcal{L}(X_{t+1} | \mathcal{F}_t) = N(\bar{m}_t - \Gamma V, V)}$$

Model for firms (23/1/11)

This may be very similar to existing stories, but it seems to me that the mathematical modelling of a firm has to look something like this. Work in discrete time. The state of the firm at the end of period $t-1$ is a triple $(K_{t-1}, D_{t-1}, Q_{t-1})$ where K_{t-1} is the capital held, D_{t-1} is the amount borrowed (denominated in cash) and Q_{t-1} is the number of issued shares.

At the start of period t, the firm makes decisions about changing Q, D, K and investing I_t of the proceeds of the output of the coming period. The changes have to be linked by a budget equation

$$\Delta D_t + I_t + \underbrace{S_t \Delta Q_t}_{\text{cash raised by selling } \Delta Q_t \text{ new shares; we might well suppose a non-linear relation here}} = \underbrace{q_t}_{\text{market price of capital}} (\Delta K_t + \underbrace{\delta}_{\text{depreciation rate}} K_{t-1})$$

We also expect a borrowing constraint $D_t \leq \beta q_t K_t$; the amount borrowed may not exceed a fraction β of the current market value of the firm's capital.

Having done the financial adjustments, the firm moves on to production, where we assume a production function $f(K, L)$ (probably homogeneous of degree 1) and that the output $f(K_t, L_t)$ is denominated in units of consumption good. Let p_t denote the price level process. Then the monetary value of the firm's profit is

$$p_t \underbrace{(1-\epsilon)}_{\text{quantity of good required to make 1 unit of good.}} f(K_t, L_t) - w_t L_t - I_t - \underbrace{R_t}_{\text{interest rate on borrowing}} D_t$$

and the objective of management is to maximise profit per issued share

$$\max \frac{1}{Q_t} \left\{ p_t (1-\epsilon) f(K_t, L_t) - w_t L_t - I_t - R_t D_t \right\}$$

We can take this a bit further on if we assume cash raised by issuing ΔQ_t shares is $h_t(\Delta Q_t)$, increasing convex. If we suppose for the moment that ΔQ_t has been chosen and is held fixed, our goal is to

$$\max \quad p_t(1-\epsilon) f(K_t, L_t) - w_t L_t + \left(\Delta D_t + h_t(\Delta Q_t) - q_t(\Delta K_t + \delta K_{t-1})\right) - R_t D_t$$

Now if (as seems likely) $R_t < 1$, the best thing is to push D_t up as far as the bank will allow, so we get $D_t = \beta q_t K_t$, and we need to

$$\max \left[p_t (1-\epsilon) f(K_t, L_t) - w_t L_t + (1-R_t) \beta q_t K_t - q_t \Delta K_t \right]$$

so we get the FOCs

$$\begin{cases} p_t (1-\epsilon) f_K (K_t, L_t) = q_t (1 - \beta(1-R_t)) \\ p_t (1-\epsilon) f_L (K_t, L_t) = w_t \end{cases}$$

For (say) Cobb-Douglas f we could now get K, L quite explicitly

I guess if the firm finds it impossible to make a positive profit, it has to go bankrupt. We'd also need to allow the government to tax the profits at some stage if this was in the story.

For $0<\alpha<1$

$$\sup_{s}\left[A s^{\alpha} - q s\right] = A(1-\alpha)\left(\frac{q}{\alpha A}\right)^{\alpha/(\alpha-1)}$$

<u>History-dependent preferences again (27/1/11)</u>

(i) This optimal-investment problem was studied already in WN XXVIII p 51. We have

$$\begin{cases} dw_t = rw_t\,dt + \theta_t(\sigma dW_t + (\mu - r)dt) - c_t\,dt \\ d\xi_t = \lambda(c_t^\alpha - \xi_t)\,dt \end{cases}$$

for some $0 < \alpha \le 1$, and the objective is

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

As we saw, the value $V(w,\xi)$ scales: $V(\beta w, \beta^\alpha \xi) = \beta^{\alpha(1-R)} V(w,\xi)$ so we get $V(w,\xi) = \xi^{1-R} v(x)$, $x = w/\xi^{1/\alpha}$.

The HJB is

$$0 = U(1) - \rho v + rxv' - \lambda\left((1-R)v - \tfrac{1}{\alpha} xv'\right) - \tfrac12 \kappa^2 \frac{v'^2}{v''}$$
$$+ \sup_s \left[\lambda s^\alpha\left((1-R)v - \tfrac{1}{\alpha} xv'\right) - s v'\right]$$

(ii) We can do the dual variables thing: $z = v'(x)$, $J(z) = v(x) - xz$, $J' = -x$, $J'' = -1/v''$ and we find

$$0 = U(1) - \rho(J - zJ') - rzJ' + \tfrac12 \kappa^2 z^2 J''$$
$$+ \sup_s \left[\lambda s^\alpha \left\{(1-R)v - \tfrac{1}{\alpha} xv'\right\} - s v'\right]$$

For $\alpha \in (0,1)$, this is as far as we can go, but if we take $\alpha = 1$, and suppose that $c \ge 0$, we shall have to have

$$\lambda\left((1-R)v - xv'\right) - v' \le 0$$

that is $(1+\lambda x)^{R-1} v$ is increasing, or $z^{(1-R)/R} J - \frac{1}{\lambda} z^{1/R}$ decreasing:

$$J' + \frac{1-R}{Rz} J \le \frac{1}{\lambda R}.$$

What we expect will happen is for $x \ge x_*$ we immediately push ξ up, w down, to restore $x = w/\xi$ to x_*. Using scaling form of V, this gives

$$v(x) = \left(\frac{1+\lambda x}{1+\lambda x_*}\right)^{1-R} v(x_*) \qquad \text{for } x \ge x_*$$

Thus for $x \geq x_*$ we shall have

$$\cancel{v(x) = a \frac{(1+\lambda x)^{1-R}}{1-R}}$$

for some $a > 0$ and so for $z \leq \lambda a(1+\lambda x_*)^{-R} \equiv z_*$ we shall have

$$J(z) = \sup\left[a\frac{(1+\lambda x)^{1-R}}{1-R} - zx \right]$$

$$= \sup\left[a\frac{(1+\lambda x)^{1-R}}{1-R} - \frac{z}{\lambda}(1+\lambda x)\right] + z/\lambda$$

$$= a\tilde{U}(z/\lambda a) + z/\lambda$$

To the right of $z = z_*$, the solution to the ODE has the form

$$J(z) = \frac{U(1)}{\rho} + Az^{-\alpha} + Bz^{\beta}$$

where $-\alpha < 0 < 1 < \beta$ solve $\frac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho = 0$. In fact, since J must be convex decreasing, we have $B = 0$, and

$$J(z) = \begin{cases} a\,\tilde{U}(z/\lambda a) + z/\lambda & (z \leq z_*) \\ \dfrac{U(1)}{\rho} + A z^{-\alpha} & (z \geq z_*) \end{cases}$$

with C^2 condition at z_*, which provides three conditions to fix the three unknowns a, A, z_*.

(iii) This story can be instead interpreted as an agent who gains utility from holding stocks of a perishable good. If he has no price impact in the sense that $\lambda = 1$, he just goes shopping when $x = W/\xi$ falls to some threshold value, at which he does 'local time shopping'.

How would it look if he gained utility from holding stocks of several goods? The asset dynamics now will be

$$\begin{cases} dW_t = rW_t\,dt + \theta_t(\sigma\,dW_t + (\mu - r)dt) - p\cdot c_t\,dt \\ d\xi^i_t = \lambda^i(c^i_t - \xi^i_t)\,dt \end{cases}$$

$$V_w = w^{b-1}\left(bv - x \cdot Dv \right)$$

$$V_{ww} = w^{b-2}(b-1)\left(bv - x\cdot Dv\right) - w^{b-2}x \cdot \left(b\, Dv - Dv - x \cdot D^2 v \right)$$

$$V_{\xi} = w^{b-1}\, Dv$$

$$v(x) = K_0 U(x)$$

$$x \cdot Dv(x) = \left(\sum \alpha_i\right) v(x) \equiv K_1 U(x)$$

$$x \cdot Dv(x)) \cdot x = \left\{ \left(\sum \alpha_i\right)^2 - \left(\sum \alpha_i^2\right) \right\} v(x) \equiv K_2 U(x)$$

$$0 = -\rho K_0 + r\left(bK_0 - K_1\right) - \sum \lambda_i \alpha_i$$

$$- \frac{1}{2}\kappa^2 \frac{\left(bK_0 - K_1\right)^2}{b(b-1)K_0 - 2(b-1)K_1 + K_2} + 1$$

(ii) Suppose for simplicity that we take some scalable U, either

$$U(x) = \prod_{i=1}^{J} x_i^{\alpha_i} \qquad \left(\alpha_i > 0, \ \sum \alpha_j \le 1\right)$$

or

$$U(x) = -\prod_{i=1}^{J} x_i^{-a_i} \qquad (a_i > 0)$$

(These U are increasing in all arguments, and concave). Then there is a scaling relationship

$$V(tw, t\xi) = t^{b}\, V(w,\xi)$$

where $b = \sum \alpha_i$ for the first case, $b = -\sum a_i$ for the second case. So we get

$$V(w,\xi) = w^b v(x) \equiv w^b v(\xi/w)$$

for some function v, and the HJB

$$0 = \sup\Big[-\rho V + \big(rw + \theta(\mu - r) - p\cdot c\big)V_w + (D_\xi V)\,\lambda(c-\xi) + \tfrac{1}{2}\sigma^2\theta^2 V_{ww} + U(\xi)\Big]$$

Simplifies to

$$0 = \sup\Big[-\rho v + \big(r + \tilde\theta(\mu - r) - p\cdot\tilde c\big)(bv - x\cdot Dv) + (Dv)\cdot\lambda(\tilde c - x) + \tfrac{1}{2}\sigma^2\tilde\theta^2\big\{(b-1)(bv - x\cdot Dv) - (b-1)\,x\cdot Dv + x\cdot D^2 v\cdot x\big\} + U(x)\Big]$$

$$= \sup\Big[-\rho v + \big(r + \tilde\theta(\mu - r) - p\cdot\tilde c\big)(bv - x\cdot Dv) + (Dv)\cdot\lambda(\tilde c - x) + \tfrac{1}{2}\sigma^2\tilde\theta^2\big\{b(b-1)v - 2(b-1)\,x\cdot Dv + x\cdot D^2 v\cdot x\big\} + U(x)\Big]$$

If we have the constraint $\tilde c \ge 0$, this will require

$$\boxed{p_i(bv - x\cdot Dv) \ge \lambda_i D_i v}$$

with equality when you are buying. The optimization over $\tilde\theta$ gives us

$$\boxed{0 = -\rho v + r(bv - x\cdot Dv) - (Dv)\cdot\lambda x - \tfrac{1}{2}\kappa^2 \frac{(bv - x\cdot Dv)^2}{b(b-1)v - 2(b-1)\,x\cdot Dv + x\cdot D^2 v\cdot x} + U(x)}$$

This would suggest a guess of the form $\boxed{v(x) = K_0 U(x)}$. Then we have $x\cdot Dv(x) = (\sum\alpha_i)\, v(x) \equiv K_1 U(x)$, and $x\cdot(D^2 v)\cdot x = \big((\sum\alpha_i)^2 - \sum\alpha_i^2\big) v(x) \equiv K_2 U(x)$. To fix the constant K_0, we have the equation

$$\boxed{0 = -\rho K_0 + r(bK_0 - K_1) - \sum \lambda_i \alpha_i K_0 - \frac{1}{2}\kappa^2 \frac{(bK_0 - K_1)^2}{b(b-1)K_0 - 2(b-1)K_1 + K_2} + 1}$$

This should allow us to identify what K_0 will be, but the remaining issue concerns what v looks like in the region where we do some sudden buying. Can we extend this guess for v off the no-buy region so that the HJB holds universally?

LAF: the second-order effect (4/2/11)

In the LAF story, we identified the least action = ML path, and then considered perturbing by ξ around that least action path. The second-order effect is

$$Q(\xi) = \tfrac{1}{2} D_i D_j \varphi(X_0^*)\, \xi_0^i \xi_0^j + \int_0^T \left\{ \tfrac{1}{2} \xi_s^i \xi_s^j A_s^{ij} + \xi_s^i B_s^{ij} \dot{\xi}_s^j + \tfrac{1}{2}\dot{\xi}_s^i q_s^{ij} \dot{\xi}_s^j \right\} ds$$

Let's for the moment just assume the scalar case $d=1$. We know that $Q(\xi) \geq 0$ whatever ξ, so let's try to find a ξ which will minimise $Q(\cdot)$ subject to $\xi_0 = 1$, say. The variational problem would perturb the optimal ξ to $\xi+\eta$, where the first-order story would say

$$0 = \int_0^T \left(\xi_s A_s \eta_s + \xi_s B_s \dot{\eta}_s + \eta_s B_s \dot{\xi}_s + \dot{\xi}_s q_s \dot{\eta}_s \right) ds$$

$$= \int_0^T \eta_s \left(\xi_s A_s + B_s \dot{\xi}_s - (B_s \xi_s + \dot{\xi}_s q_s)^{\cdot} \right) ds + \Big[\eta_s (B_s \xi_s + \dot{\xi}_s q_s) \Big]_0^T$$

$$= \int_0^T \eta_s \left\{ \xi_s A_s + \cancel{B_s \dot{\xi}_s} - \dot{B}_s \xi_s - \cancel{\dot{\xi}_s B_s} - (\dot{\xi} q)^{\cdot} \right\} ds + \eta_T (B_T \xi_T + \dot{\xi}_T q_T)$$

So we would have to have

$$\boxed{\frac{d}{dt}(\dot{\xi} q) - \xi A + \xi \dot{B} = 0} \qquad B_T \xi_T + \dot{\xi}_T q_T = 0$$

When we tried to make $Q(\cdot)$ a perfect square, we added and subtracted $\left[\tfrac{1}{2}\theta_t \xi_t^2\right]_0^T$ where we wanted θ to solve

$$A + \dot{\theta} = (B+\theta)^2/q$$

If we now put $\boxed{(B+\theta) = -q\dot{\psi}/\psi}$ in the usual Riccati style then

$$\dot{B} + \dot{\theta} = +\frac{\dot{q}}{q}(B+\theta) - q\left(\frac{\ddot{\psi}}{\psi} - \frac{(\dot{\psi})^2}{\psi^2} \right) = \frac{\dot{q}}{q}(B+\theta) - q\frac{\ddot{\psi}}{\psi} + \frac{1}{q}(B+\theta)^2$$

$$= -\dot{q}\frac{\dot{\psi}}{\psi} - q\frac{\ddot{\psi}}{\psi} + A + \dot{\theta}$$

$$\Rightarrow \boxed{\frac{d}{dt}(q\dot{\psi}) - A\psi + \dot{B}\psi = 0}$$

which is the same ODE!! even with the same BC at T !!!

Continuous-time contracting problem again (4/2/11)

(i) My analysis of the problem on pp 72-73 of WN XXXI was incorrect, as Moritz has explained; the analysis assumed that the agent's value is of the form $\kappa_A U_A(w)$, for some constant κ_A, and this is erroneous – the value of κ_A may depend on the action of the principal.

So let's try again. We have

$$dw_t = w_t(\sigma dW_t + (\mu - \theta_t)dt) - (c_t + q_t)dt$$

with agent's choice to be $\theta_t \geq 0$, principal's choice to be $c_t, q_t \geq 0$, and objectives

$$E\int_0^\infty e^{-\delta t} U_A(q_t + \beta\theta_t w_t)\,dt \quad \text{for } A, \quad E\int_0^\infty e^{-\rho t} U_p(c_t)\,dt \quad \text{for } P.$$

(ii) Let's suppose that P decides to use $q_t = \nu w_t$, $c_t = \lambda w_t$ for some constants $\lambda, \nu > 0$ and now we consider the HJB for the agent

$$0 = \sup_\theta \Big[-\gamma V + (w(\mu-\theta) - c - q)V_w + \tfrac{1}{2}\sigma^2 w^2 V_{ww} + U_A(q + \beta\theta w) \Big]$$

$$= \sup_\theta U_A(w)\Big[-\gamma\kappa_a + (1-R_a)(\mu - \theta - \lambda - \nu)\kappa_a - R_a(1-R_a)\kappa_a\sigma^2/2 + \left(\tfrac{q}{w} + \beta\theta\right)^{1-R_a} \Big]$$

When this is maximized, get FOC for θ

$$\kappa_a = \beta(\nu + \theta\beta)^{-R_a}$$

so the equation to be solved for κ_a (or equivalently θ) would be

$$0 = -\gamma - (R_a - 1)(\mu - \lambda - \nu) + (R_a - 1)R_a\sigma^2/2 + R_a\theta + \nu/\beta$$

$$\Rightarrow \theta = R_a^{-1}\{\gamma + (R_a - 1)(\mu - \lambda - \nu) - (R_a - 1)R_a\sigma^2/2 - \nu/\beta\}^+$$

As ν gets bigger, θ gets smaller. The slope is $-(R_a - 1 + \beta^{-1})/R_a$ which is bigger than 1 in modulus

The principal's problem is now

$$0 = \sup_{\lambda,\nu}\Big[-\rho\kappa_p + (1-R_p)(\mu - \theta - \lambda - \nu)\kappa_p + (R_p - 1)R_p\kappa_p\sigma^2/2 + \lambda^{1-R_p}\Big] U_p(w)$$

The FOC for λ will be

$$\lambda^{-R_p} = \kappa_p / R_a$$

$$\lambda\left(\frac{1}{R_a} + R_p - 1\right) + \nu\,(R_p - 1) = \rho + (R_p - 1)\,\mu - R_p(R_p - 1)\,\sigma^2/2$$

$$\lambda\,(R_a - 1) + \nu\left(R_a - 1 + \frac{1}{R_p}\right) = \gamma + (R_a - 1)\,\mu - R_a(R_a - 1)\,\sigma^2/2$$

and optimizing over ν we would set ν to make Θ vanish. This means we must have

$$\nu = \frac{\gamma + (R_a - 1)(\mu - \lambda - \sigma^2 R_a/2)}{R_a - 1 + \beta^{-1}}$$

We get then the two equations for λ, ν:

$$\left.\begin{aligned} 0 &= -\rho + (1-R_p)(\mu-\lambda-\nu) + R_p(R_p-1)\frac{\sigma^2}{2} + \frac{\lambda}{R_a} \\ 0 &= -\gamma + (1-R_a)(\mu-\lambda-\nu) + R_a(R_a-1)\frac{\sigma^2}{2} + \frac{\nu}{\beta} \end{aligned}\right\}$$

We can solve these linear equations explicitly:

$$\big((1-R_a)(1-\beta) + R_a R_p\big)\lambda = R_a\Big\{(R_p-1)\frac{\sigma^2}{2}\big(R_a\beta(R_a-1-R_p) - (1-\beta)R_p\big) + (R_p-1)\mu + \gamma\beta(1-R_p) + \rho(1-\beta) + \beta\rho R_a\Big\}$$

$$\big((1-R_a)(1-\beta) + R_a R_p\big)\nu = \beta\Big\{-R_a(R_a-1)\frac{\sigma^2}{2}\big(R_a(R_p-1) + 1 + R_p - R_p^2\big) + \mu(R_a-1) + \gamma(1-R_a+R_a R_p) - R_a\rho(R_a-1)\Big\}$$

A question from Katsumasa Nishide (7/2/11)

(i) Katsumasa has been looking at a paper by Carmona et al on emissions trading where he tries to extend to a continuous-time setting the questions studied there. To summarize, the idea is to take the view of a central planner and suppose that various technologies are available to generate electricity, technology i costing $c_i(q)$ to generate at rate q, and producing emissions at rate $e_i(q)$ which have to be controlled in some way. If p_t is the price at time t of electricity, then the aim is to

$$\max E\left[\int_0^T e^{-rs} \sum_i (p_t q_{is} - c_i(q_{is}))\, ds - e^{-rT} L(E_T)\right]$$

where L is some penalty function, and $E_t = \int_0^t \sum_j e_j(q_{js})\, ds$.

(ii) Let's suppose that the demand X_t at time t is some exogenous diffusion process. Then the HJB for the value $G(t, X, e)$ will be

$$0 = \sup\left[-rG + G_t + \tfrac{1}{2}\sigma(x)^2 G_{xx} + \mu(x) G_x + \Sigma e(q) G_E + \Sigma p q_i - \Sigma c_i(q_i)\right]$$

where we want to max over q. Notice that if we know the current demand X, then the price p must be set so as to clear the market.

This therefore means that

$$\sup_q \left\{ \Sigma e(q) G_E + p \Sigma q_j - \Sigma c_j(q_j) \right\} = h(t, X, G_E)$$

for some function h which may be made more explicit in nice examples.

The HJB then says

$$0 = -rG + G_t + \tfrac{1}{2}\sigma(x)^2 G_{xx} + \mu(x) G_x + h(t, X, G_E)$$

where the BC at time T is $G(T, X, E) = -L(E)$.

(iii) A nice special case arises when the loss function is linear, $L(E) = -aE$. Then clearly

$$G(t, X, E) = G(t, X, 0) - a e^{-r(T-t)} E$$

so we get ($g(t,X) \equiv G(t,X,0)$)

$$\boxed{0 = -rg + g_t + \tfrac{1}{2}\sigma^2 g_{xx} + \mu g_x + h(t, x, -a e^{-r(T-t)})}$$

The solution g now has a F-K style representation

$$\boxed{g(t,x) = E\left[\int_t^T e^{-r(s-t)} h(s, X_s, -\alpha e^{-r(T-s)})\, ds \,\Big|\, X_t = x\right]}$$

So this may be tractable for suitably pleasant h. Certainly OK numerically.

LAF again: the multivariate case (9/2/11)

(i) In the LAF story, the second order effect in the action if we perturb away from optimal x^* by ξ comes out to be

$$Q(\xi) = \tfrac{1}{2}\,\xi_0\cdot D_{xx}\varphi(x_0^*)\,\xi_0 + \int_0^T \left\{ \tfrac{1}{2}\xi\cdot A\xi + \xi\cdot B\dot\xi + \tfrac{1}{2}\dot\xi\, q\,\dot\xi \right\} ds$$

and we seek to extend the story on page 11 to this multivariate context. So suppose that we fix $\xi_0 \neq 0$ for now, and ask for the path ξ which will minimize Q. We find this by calculus of variations, perturbing ξ to $\xi+\eta$ for small η and then studying the first-order condition

$$0 = \int_0^T \left\{ \eta A\xi + \eta B\dot\xi + \xi B\dot\eta + \dot\eta q\dot\xi \right\} ds$$

$$= \int_0^T \eta\left\{ (A\xi + B\dot\xi) - (q\dot\xi + B^T\xi)^{\cdot} \right\} ds + \left[\eta\,(B^T\xi + q\dot\xi) \right]_0^T$$

so we deduce the boundary condition

$$B_t^T\xi_t + q_t\dot\xi_t = 0 \qquad \text{at } t = T$$

and the ODE

$$\boxed{0 = (A - \dot B^T)\xi + (B - B^T - \dot q)\dot\xi - q\ddot\xi}$$

which is second order, linear. The solution therefore looks like

$$\xi_t = F_t\,\xi_0$$

for some matrix-valued function of time.

(ii) Assuming that F_t^{-1} exists for all time, we can define a matrix valued f^n of time θ_t by

$$\boxed{q_t^{-1}(B_t^T + \theta_t) = -\dot F_t F_t^{-1}}$$

Now multiply by q_t, differentiate and develop:

$$\dot B^T + \dot\theta = -\dot q\,\dot F F^{-1} - q(\ddot F F^{-1} - \dot F F^{-1}\dot F F^{-1})$$

so that

$$(\dot B^T + \dot\theta)F = -\dot q\dot F - q\ddot F + q\dot F F^{-1}\dot F$$

$$= -(A-\dot{B}^T)F - (B-B^T)\dot{F} + q\dot{F}F^{-1}\dot{F}$$ using ODE for ξ

$\Rightarrow$

$$(A+\dot{\theta})F = (B^T-B)\dot{F} - (B^T+\theta)\dot{F}$$

$$= -(B+\theta)\dot{F}$$

$$= (B+\theta)q^{-1}(B^T+\theta)F$$

and hence we deduce that

$$\boxed{A+\dot{\theta} = (B+\theta)q^{-1}(B^T+\theta)}$$

with BC $\theta_T = 0$, just as in the LAF notes!

(iii) What could go wrong? For the solution of the ODE for ξ, provided that A, B, q, $\dot{q}$ are all continuous on $[0,T]$, they are bounded and there is no problem solving for ξ out from initial conditions $(\xi_0, \dot{\xi}_0)$, or perhaps more simply, solving back from ξ_T, using the BC at $t=T$ to fix $\dot{\xi}_T$.

So the only thing which could go wrong is that F_t fails to be invertible somewhere. The right way to think about this is solving back from $t=T$, because for some time just before T, F_t will be invertible. But suppose as t decreases we hit some t^* where F_t becomes singular. What this means is that there is some solution ξ^* starting from some value $\xi^*_T \neq 0$ which hits zero at $t=t^*$. Now minimality implies that

$$\int_{t^*}^T \left\{ \tfrac{1}{2}\xi^*_s A_s \xi^*_s + \xi^*_s B_s \dot{\xi}^*_s + \tfrac{1}{2}\dot{\xi}^*_s q_s \dot{\xi}^*_s \right\} ds < 0$$

because we could always replace ξ^* on $[t,T]$ by the constant $0 \in \mathbb{R}^d$, and this would give a larger value of Q. But if we now make

$$\tilde{\xi}_t = 0 \qquad \forall\, t \le t^*$$
$$= \xi^*_t \qquad \forall\, t \ge t^*$$

we see that $Q(\tilde{\xi}) < 0$!! This contradiction guarantees that F_t is always invertible, and so we do have a solution (θ_t) to the Riccati equation, $\theta_T = 0$

A question of Phil D & Yajun Wang (14/2/11)

(i) This is a story with conventional wealth dynamics

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

and objective

$$\max E \int_0^\infty e^{-\rho s} U(c_s)\, ds$$

but subject to some constraint of the form

$$\boxed{\theta_t \leq K w_t}$$

for some K. We do not of course assume that U is CRRA, and we could even allow U to be nonconcave, at the price of replacing U with its GCM.

The HJB is just

$$0 = \sup_{\substack{c \geq 0 \\ \theta \leq Kw}} \left[-\rho V + (rw + \theta(\mu - r) - c)V' + \tfrac{1}{2}\sigma^2\theta^2 V'' + U(c) \right]$$

(ii) We can optimise by the usual things:

$$c^* = I(V')$$

$$\theta^* = \left(\frac{-(\mu - r)V'}{\sigma^2 V''} \right) \wedge (Kw)$$

This might suggest it could be interesting to have an increasing coeff of RRA so that the constraint could bind somewhere. So we might try having this for U as well, as in

$$\frac{-xU''}{U'} = \frac{1}{a + bx}$$

for some a, b > 0, which would lead to

$$U'(x) = \left(\frac{a + bx}{x} \right)^{1/a}.$$

Probably we can't integrate this to obtain U. Also, $U'(\infty) = b^{1/a} > 0$. Not so good... Another try would be

$$\boxed{I(y) = y^{-1/R_1} + q\, y^{-1/R_2}}$$

for $R_2 > R_1 > 1$ (say) for well-posed problem, some $q > 0$. This would imply that $U'(x) \sim x^{-R_2}$ for small x, $U'(x) \sim x^{-R_1}$ for large x.

(iii) This appears to be numerically delicate if we don't get the BCs correctly, for example. Let's suppose that

$$\frac{\mu - r}{\sigma^2 R_2} < K < \frac{\mu - r}{\sigma^2 R_1}$$

so that the constraint on the portfolio will bite at high wealth levels, but not at low wealth levels. If x is very small, the x^{1-R_1} bit doesn't ever get seen, and we may as well suppose that we are just a CRRA investor with R_2, $I(y) = q\, y^{-1/R_2}$, which would give

$$U(x) = q^{R_2} \frac{x^{1-R_2}}{1-R_2}$$

so the value function for small initial wealth levels would be

$$V(x) \approx q^{R_2} \gamma_{M_2}^{-R_2} \frac{x^{1-R_2}}{1-R_2}$$

where here

$$\gamma_{M_2} = R_2^{-1} \left\{ \rho + (R_2 - 1)(r + K^2/2R_2) \right\}$$

For large wealth levels, we expect to have $I(y) \approx y^{-1/R_1}$, $U(x) \approx x^{1-R_1}/(1-R_1)$ and HJB which looks like

$$0 = \sup_c \left[-\rho V + (rw + Kw(\mu - r) - c) V' + \tfrac{1}{2}\sigma^2 K^2 w^2 V'' + U(c) \right]$$

so if $\boxed{V(w) = a\, U(w)}$, we get $V' = c^{-R} \Rightarrow c = a^{-1/R} w$, and

$$0 = -a\rho + (r + K(\mu - r))a(1-R_1) - R_1(1-R_1)\frac{\sigma^2 K^2}{2} a + R_1 a^{1-1/R_1}$$

$$\Rightarrow \boxed{a^{-1/R_1} = R_1^{-1} \left\{ \rho + (R_1 - 1)\left(r + K(\mu - r) - \tfrac{1}{2}\sigma^2 K^2 R_1 \right) \right\}}$$

This tells us what approximately the value function will be for large wealth levels.

Merton problem with CARA utility (17/2/11)

(i) Suppose we have standard wealth dynamics

$$dw_t = rw_t\,dt + \theta_t(\sigma dW_t + (\mu - r)dt) - c_t\,dt$$

with

$$U(x) = -\exp(-\gamma x) \qquad \left(\text{so } \tilde{U}(\lambda) = \tfrac{\lambda}{\gamma}\log\left(\tfrac{\lambda}{\gamma e}\right)\right)$$

and the usual objective. Solving the HJB equation

$$0 = \sup\left[-\rho V + (rw + \theta(\mu - r) - c)V' + \tfrac{1}{2}\sigma^2\theta^2 V'' + U(c)\right]$$

we find a possible solution is to take

$$V(x) = -A\exp(-r\gamma x)$$

using

$$\theta^* = \frac{\mu - r}{\sigma^2 r\gamma}, \qquad c^* = rw - \log(rA)/\gamma$$

$$A = r^{-1}\exp\left[-\frac{(\mu - r)^2 + 2(\rho - r)\sigma^2}{2r\sigma^2}\right]$$

Thus the optimally-controlled wealth process is a drifting Brownian motion, which can of course go unboundedly negative... so what problem if any does this solve? A question I put to Alex Selby!

(ii) The dual form of this is

$$0 = -\rho J + (\rho - r)z J' + \tfrac{1}{2}z^2\kappa^2 J'' + \tfrac{z}{\gamma}\left\{\log\left(\tfrac{z}{\gamma}\right) - 1\right\}$$

Solutions to the homogeneous ODE are of the form $Az^{-\alpha} + Bz^{\beta}$, where $-\alpha < 0 < 1 < \beta$ solve $\tfrac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho = 0$. But there's also a particular solution of the form

$$\frac{1}{r\gamma}z\log z + qz$$

where

$$q = (\gamma r)^{-1}\left\{\frac{\tfrac{1}{2}\kappa^2 + (\rho - r)}{r} - \log\gamma - 1\right\}$$

Suppose we now tell the story that when wealth hits $-K$, you get banned from the stock market and have to pay back. This would imply that $V(-K) = -\frac{1}{\rho}\exp(\gamma r K)$ so in dual variables, when $J'(z) = K$ we have $V = J(z) - Kz = -\frac{1}{\rho}e^{\gamma r K}$

Selby tells a story where $w_t \geq -K$ always, but if you reach $-K$ you may come out again. This requires that at critical z^*, $J'(z^*) = K$, $J''(z^*) = 0$

With utility $-\gamma^{-1} e^{-\gamma x}$, he gets

$$J(z) = \frac{1}{\gamma r}\left[z\left(\log z - 1 + \frac{\delta}{r}\right) - \frac{(z/z^*)^\beta}{\beta(\beta-1)} z^* \right]$$

Now check $J'(z) = \frac{1}{\gamma r}\left[\log z + \frac{\delta}{r} - \frac{1}{\beta-1}\left(z/z^*\right)^{\beta-1}\right]$, $J''(z) = \frac{1}{\gamma r}\left[\frac{1}{z} - \frac{1}{z_*}\left(\frac{z}{z_*}\right)^{\beta-2}\right]$

so at $z = z_*$ we get $\left(\log z_* = \frac{1}{\beta-1} - \frac{\delta}{r} + r\gamma K\right)$ that $J'(z_*) = K$, and $J''(z_*) = \frac{1}{\gamma r}\left[\frac{1}{z_*} - \frac{1}{z_*}\right] = 0$. So for $z \leq z_*$ we have $J(\cdot)$ given by the above formula, but for $z > z_*$ we get $J(z) = V(-K) + zK$

If we write $z = z_* - t$, then to highest order near z_*

$$J(z) = J(z_*) - Kt + \frac{t^3}{6}\,\frac{\beta-1}{z_*^2}\,\frac{1}{\gamma r}$$

so that $$w + K \doteq \frac{t^2}{2}\,\frac{\beta-1}{z_*^2}\,\frac{1}{\gamma r}$$

$$\theta = -\frac{\mu - r}{\sigma^2}\frac{V'}{V''} = \frac{\mu - r}{\sigma^2} z J'' \approx \frac{\beta-1}{z_*}\,\frac{1}{\gamma r}\,\frac{\kappa}{\sigma}\, t$$

$$c = -\frac{1}{\gamma}\log(z) \approx \frac{1}{\gamma}\left\{ -\frac{1}{\beta-1} + \frac{\delta}{r} - r\gamma K\right\}$$

So for w near $-K$ we get

$$\theta \approx \sqrt{\frac{2(\beta-1)}{\gamma r}(w+K)}\;\frac{\kappa}{\sigma}$$

so the SDE for w looks like $\left[A \equiv \kappa\sigma^{-1}\left((\beta-1)/2\gamma r\right)^{1/2}\right]$

$$dw_t = 2A\sigma\sqrt{w_t + K}\,(dW_t + \kappa dt) + (rw - c)dt$$

$$= 2A\sigma\sqrt{w_t + K}\,(dW_t + \kappa dt) + r(w+K)dt + \underbrace{\left(\frac{1}{\beta-1} - \frac{\delta}{r}\right)}_{= q, \text{ say}}\frac{dt}{\gamma}$$

Dimension is

$$\frac{q}{A^2\sigma^2} = 1 \text{ after some calculations.}$$

So once your wealth hits $-K$, you can't continue thereafter, since $V'(-K) \neq 0$, so Selby's story becomes the one I was telling

We know that as $z \to 0$, $J(z) \to 0 = J(0) = \sup V(w)$, so the term in z^{α} is not present in the solution! So we get

$$J(z) = Bz^{\beta} + \frac{z \log z}{r\gamma} + qz$$

To satisfy the boundary condition that at some $z(K)$ we have

$$J'(z(K)) = K, \quad J(z(K)) - Kz(K) = -\frac{1}{\rho}\exp(\gamma r K)$$

there has to be some $B(K)$ to do the job. Suppose that as $K \to \infty$ we have $\limsup B(K) = \varepsilon > 0$. Then for arbitrarily large K_n we will have

$$z(K_n) \sim \left(\frac{K_n}{\beta\varepsilon}\right)^{1/(\beta-1)}$$

$$J(z(K_n)) \sim \varepsilon\, z(K_n)^{\beta} \sim \varepsilon \left(\frac{K_n}{\beta\varepsilon}\right)^{\beta/(\beta-1)} = \varepsilon \left(\frac{K_n}{\beta\varepsilon}\right)^{1+1/(\beta-1)}$$

Likewise

$$K_n\, z(K_n) \sim K_n^{1+1/(\beta-1)} \left(\frac{1}{\beta\varepsilon}\right)^{1/(\beta-1)}$$

and it's not possible then that $J(z(K_n)) - K_n z(K_n) = -\frac{1}{\rho}\exp(\gamma r K_n)$

So we have to have

$$\lim_{K\to\infty} B(K) = 0$$

and the desired solution really can be interpreted as a limiting form of the problem stopped at $-K$.

Merton problems when utility is bounded below (20/2/11)

(1) Let's look at the problem of a Merton-type terminal wealth problem $\max E\, U(w_T)$, where we realise that there is no difference as far as the agent is concerned between finishing with -10^{10} USD and finishing with -10^{100} USD. So unless there is a hard lower bound for wealth, the optimal behaviour would involve infinite risk taking...

So let's consider a story where you might go into $w<0$, but that as you do so there is a risk of discovery and a big penalty $-A$; we'll suppose the rate at which discovery happens is $\alpha(w_t)(a+\frac{1}{2}b\,\theta_t^2)$, where $\alpha(w)=0$ for $w \geq 0$, and $\alpha(\cdot)$ is decreasing. It seems fair to allow that if you have a large position in risky assets, you are more likely to be found out.

The MPOC would say

$$V(t\wedge\tau, w_{t\wedge\tau}) - (A+V(\tau, w_\tau))\, I_{\{\tau\le t\}}$$

is a supermartingale etc, with $V(T,w) = U(w)$. Hence

$$V(t\wedge\tau, w_{t\wedge\tau}) - \int_0^{t\wedge\tau}(A+V)\,\alpha(w_s)\,(a+\tfrac{1}{2}b\,\theta_s^2)\,ds \quad \text{is a supermartingale etc}$$

So we get HJB

$$0 = \sup_\theta\Big[\, V_t + (rw+\theta(\mu-r))V_w + \tfrac{1}{2}\theta^2\sigma^2 V_{ww} - \alpha(w)(a+\tfrac{1}{2}b\theta^2)(A+V)\,\Big]$$

which modifies the agent's investment,

$$\theta(\, b\alpha\,(A+V) - \sigma^2 V_{ww}\,) = (\mu-r)V_w$$

So we have

$$0 = V_t + rwV_w - \alpha(w)(A+V)a + \frac{(\mu-r)^2 V_w^2}{2\,(b\alpha(w)(A+V) - \sigma^2 V_{ww})}$$

as HJB, probably completely hopeless except numerically.

(2) Could also tell an infinite-horizon version with running consumption: perhaps this time we just have detection rate $\alpha(w,c)$ as the portfolio might be less variable. Then

$$0 = \sup\Big[-\rho V + (rw+\theta(\mu-r)-c)V_w + \tfrac{1}{2}\sigma^2\theta^2 V_{ww} - \alpha(w,c)(A+V) + U(c)\Big]$$

which would give as first-order conditions the familiar portfolio condition:

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

and the condition

$$V_w = U'(c) - \alpha_c (A+V)$$

for the consumption rate. We may as well suppose that α is convex increasing inc, decreasing in w.

$$D(R_\lambda f) = D\psi_\lambda^- \cdot A + D\psi_\lambda^+ \cdot B$$

$$D^2(R_\lambda f) = A\, D^2\psi_\lambda^- + B\, D^2\psi_\lambda^+ + hf\left(\psi_\lambda^+ D\psi_\lambda^- - \psi_\lambda^- D\psi_\lambda^+\right)$$

$$\therefore \left(\lambda R_\lambda f - \tfrac{1}{2}\sigma^2 D^2(R_\lambda f) - \mu D(R_\lambda f)\right) = \frac{h\sigma^2}{2}\left(\psi_\lambda^- D\psi_\lambda^+ - \psi_\lambda^+ D\psi_\lambda^-\right) f$$

Boundary conditions for solving stochastic optimal control problems (28/2/4.)

1) Numerical solution of one-dimensional stochastic optimal control problems is in principle quite straightforward using policy improvement, but in practice it appears to require a careful treatment of the boundary conditions.

What I typically do is set down some grid $\underline{x} = x_1 < \dots < x_N = \bar{x}$ and build a MC approximation on that grid, running policy improvement over the problem enough times. To do this, we need good boundary conditions at $\underline{x}$, $\bar{x}$. In order to do this, what I'd suggest is that we assume a fixed diffusion outside $[\underline{x}, \bar{x}]$ with generator $\mathcal{G}$, and then if we take a starting point $x > \bar{x}$ (say) we shall have

$$V(x) = E^x\left[\int_0^\tau e^{-\rho s} f(X_s)\,ds\right] + E^x\left[e^{-\rho\tau}\right] V(\bar{x})$$

where $\tau = \inf\{t : X_t = \bar{x}\}$. This will give us

$$\frac{V(x) - V(\bar{x})}{x - \bar{x}} = \frac{E^x \int_0^\tau e^{-\rho s} f(X_s)\,ds - E^x(1 - e^{-\rho\tau})\, V(\bar{x})}{x - \bar{x}} \qquad (1)$$

2) Suppose we have generator $\mathcal{G} = \frac{1}{2}\sigma(x)^2 D^2 + \mu(x) D$, and resolvent is

$$(R_\lambda f)(x) = \int f(y)\, \psi_\lambda^+(x \wedge y)\, \psi_\lambda^-(x \vee y)\, h(y)\,dy$$

$$= \psi_\lambda^-(x) \int_{-\infty}^x f(y)\, \psi_\lambda^+(y)\, h(y)\,dy + \psi_\lambda^+(x) \int_x^\infty f(y)\, \psi_\lambda^-(y)\, h(y)\,dy$$

$$= \psi_\lambda^-(x)\, A(x) + \psi_\lambda^+(x)\, B(x)$$

So we get

$$(\lambda - \mathcal{G})(R_\lambda f)(x) = \frac{h(x)\sigma(x)^2}{2} \left(\psi_\lambda^- D\psi_\lambda^+ - \psi_\lambda^+ D\psi_\lambda^-\right)(x)\, f(x)$$

which tells us what h must be:

$$\boxed{h(x) = \left(\tfrac{1}{2}\sigma(x)^2\, W(x)\right)^{-1}} \qquad (2)$$

where $W(x) = \psi_\lambda^-(x)\, D\psi_\lambda^+(x) - \psi_\lambda^+(x)\, D\psi_\lambda^-(x)$ is the Wronskian.

3) Now suppose we want the derivative as $x \downarrow \bar{x}$ of the resolvent density killed at $\bar{x}$. We find we need $(\bar{x} < x < y)$

$$\lim_{x \downarrow \bar{x}} \frac{1}{x - \bar{x}} \left\{ h(y)\psi_\lambda^-(y)\, \psi_\lambda^+(x) - h(y)\, \psi_\lambda^+(\bar{x})\, \psi_\lambda^-(y) \cdot \frac{\psi_\lambda^-(x)}{\psi_\lambda^-(\bar{x})} \right\}$$

For $y < x < \underline{x}$, killed resolvent density derivatives

$$\lim_{x \uparrow \underline{x}} \frac{1}{\underline{x} - x} \left\{ h(y)\psi_\lambda^+(y)\psi_\lambda^-(x) - h(y)\psi_\lambda^+(y)\psi_\lambda^-(\underline{x}) \cdot \frac{\psi_\lambda^+(x)}{\psi_\lambda^+(\underline{x})} \right\}$$

$$= \lim_{x \uparrow \underline{x}} \frac{1}{\underline{x} - x} h(y)\psi_\lambda^+(y)\psi_\lambda^-(\underline{x}) \left\{ \frac{\psi_\lambda^-(x)}{\psi_\lambda^-(\underline{x})} - \frac{\psi_\lambda^+(x)}{\psi_\lambda^+(\underline{x})} \right\}$$

$$= h(y)\psi_\lambda^+(y)\psi_\lambda^-(\underline{x}) \lim \frac{1}{\underline{x} - x} \left\{ \frac{\psi_\lambda^-(x) - \psi_\lambda^-(\underline{x})}{\psi_\lambda^-(\underline{x})} - \frac{\psi_\lambda^+(x) - \psi_\lambda^+(\underline{x})}{\psi_\lambda^+(\underline{x})} \right\}$$

$$= h(y)\psi_\lambda^+(y)\psi_\lambda^-(\underline{x}) \left(- \frac{D\psi_\lambda^+(\underline{x})}{\psi_\lambda^+(\underline{x})} + \frac{D\psi_\lambda^-(\underline{x})}{\psi_\lambda^-(\underline{x})} \right)$$

$$= - h(y) \frac{\psi_\lambda^+(y)}{\psi_\lambda^+(\underline{x})} W(\underline{x})$$

$$= h(y)\psi_\lambda^-(y) \lim \frac{1}{x-\bar{x}} \left\{ \psi_\lambda^+(x) - \psi_\lambda^+(\bar{x}) \cdot \frac{\psi_\lambda^-(x)}{\psi_\lambda^-(\bar{x})} \right\}$$

$$= h(y)\psi_\lambda^-(y) \frac{W(\bar{x})}{\psi_\lambda^-(\bar{x})} = \frac{2}{\sigma(y)^2} \frac{\psi_\lambda^-(y)}{W(y)} \bigg/ \frac{\psi_\lambda^-(\bar{x})}{W(\bar{x})} \qquad (3_+)$$

For the region below $\underline{x}$, a similar story gives density

$$- \frac{2}{\sigma(y)^2} \frac{\psi_\lambda^+(y)}{W(y)} \bigg/ \frac{\psi_\lambda^+(\underline{x})}{W(\underline{x})} \qquad (3_-)$$

4) Without explicitly finding $\psi_\lambda^\pm$ we get no further. If we have drifting BM,

$$\mathcal{G} = \tfrac{1}{2}\sigma^2 D^2 + \mu D$$

we get $\psi_\lambda^\pm(x) = \exp(\alpha_+ x),\ \exp(-\alpha_- x)$, where $\alpha_+, -\alpha_-$ solve

$$\tfrac{1}{2}\sigma^2 t^2 + \mu t - \lambda = 0$$

So $W(x) = (\alpha_+ + \alpha_-)\, e^{(\alpha_+ - \alpha_-)x} = \dfrac{2\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^{2}}\, e^{-2\mu x/\sigma^2}$

So we get the derivative at $\bar{x}$ of the resolvent density is

$$\boxed{\frac{2}{\sigma^2} \exp(-\alpha_+(y-\bar{x})),\ (y > \bar{x})} \qquad (4)$$

5) Similarly, if the diffusion generator is

$$\mathcal{G} = \tfrac{1}{2}\sigma^2 x^2 D^2 + \mu x D$$

then there are eigenfunctions

$$\psi_\lambda^+(x) = x^{\beta(\lambda)}, \quad \psi_\lambda^-(x) = x^{-\alpha(\lambda)}$$

where $-\alpha$, β are roots of $\tfrac{1}{2}\sigma^2 t(t-1) + \mu t - \lambda = 0$. Then we have $W(x) = (\alpha+\beta)x^{\beta-\alpha-1}$ and the resolvent density derivative in $(\bar{x}, \infty)$ is

$$\boxed{\frac{2}{\sigma^2 y^2} \left(\frac{y}{\bar{x}}\right)^{1-\beta(\lambda)}} \qquad (5_+)$$

and in $(0, \underline{x})$ we get

$$- \boxed{\frac{2}{\sigma^2 y^2} \left(\frac{y}{\underline{x}}\right)^{1+\alpha(\lambda)}} \qquad (5_-)$$

$$dw_t = w_t \left\{ r\,dt + K(\sigma\, dW_t + (\mu - r)\,dt) - b\,dt \right\}$$

$$\Rightarrow w_t = w_0 \exp\left\{ \sigma K W_t + \nu t \right\} \qquad \nu \equiv r + K(\mu - r) - b - \tfrac{1}{2}\sigma^2 K^2$$

So the value of using this is

$$U_1(b) \; E \int_0^\infty e^{-\rho t}\, w_t^{1-R}\, dt$$

$$= U_1(b) \int_0^\infty \exp\left\{ -\rho t + (1-R)\nu t + \tfrac{1}{2}\sigma^2 K^2 (1-R)^2 t \right\} dt$$

$$= U_1(1)\; b^{1-R} \Big/ \left\{ \rho + (R-1)\nu - \tfrac{1}{2}\sigma^2 K^2 (1-R)^2 \right\}$$

So to maximise b we shall have

$$\frac{1-R}{b} = \frac{(1-R)}{\rho + (R-1)\nu - \tfrac{1}{2}\sigma^2 K^2 (1-R)^2}$$

$$\Rightarrow b = \left\{ \rho + (R-1)\left(r + K(\mu - r) - \tfrac{1}{2}\sigma^2 K^2 R \right) \right\} - b(R-1)$$

$$\Rightarrow b = R^{-1}\left\{ \rho + (R-1)\left(r + K(\mu - r) - \tfrac{1}{2}\sigma^2 K^2 R \right) \right\}$$

6) How would this apply to the problem on p 18–19? Let's suppose we take $R_2 > R_1 > 1$ and do

$$I(y) = \begin{cases} y^{-1/R_1} & y \le 1 \\ y^{-1/R_2} & y \ge 1 \end{cases}$$

so that

$$U'(x) = \begin{cases} x^{-R_2} & x \le 1 \\ x^{-R_1} & x \ge 1 \end{cases},$$

so we may take
$$U(x) = \begin{cases} \dfrac{x^{1-R_1}}{1-R_1} & (x \ge 1) \\ \dfrac{x^{1-R_2}}{1-R_2} + \dfrac{1}{1-R_1} - \dfrac{1}{1-R_2} & (x \le 1) \end{cases}$$

If we insist that $\theta \le K w$, where $\pi_2 = \dfrac{\mu-r}{\sigma^2 R_2} < K < \dfrac{\mu-r}{\sigma^2 R_1} = \pi_1$ then for large wealth levels there will we expect be an effective constraint on wealth. Suppose then that when wealth is $\ge \bar{x}$ we shall require investment $\theta = Kw$ and consumption bw, where b is optimized by

$$b = R_1^{-1}\{\rho + (R_1-1)(r + K(\mu-r) - \tfrac{1}{2}\sigma^2K^2R_1)\}$$

Then $$dw_t = w_t\{\underbrace{\sigma K}_{=s}\, dW_t + \underbrace{(K(\mu-r)+r-b)}_{=m}\, dt\}$$

If we now define
$$F_+(s, m, q, \lambda, x_0) = \frac{d}{dx} E^x\Big[\int_0^\tau e^{-\lambda s} X_s^q\, ds\Big]\Big|_{x=x_0+}$$

(where $dX = X(s\,dW + m\,dt)$), we can express $F_\pm$ quite explicitly in terms of the roots $-\alpha < 0$ and $\beta > 1$ of $\frac{1}{2}s^2 t(t-1) + mt - \lambda = 0$, for then

$$\left.\begin{aligned} F_+(s,m,q,\lambda,x_0) &= \frac{2}{\sigma^2}\int_{x_0}^\infty y^q \Big(\frac{y}{x_0}\Big)^{1-\beta} \frac{dy}{y^2} = \frac{2}{\sigma^2}\,\frac{x_0^{q-1}}{\beta - q} \\ F_-(s,m,q,\lambda,x_0) &= -\frac{2}{\sigma^2}\int_0^{x_0} y^q \Big(\frac{y}{x_0}\Big)^{1+\alpha} \frac{dy}{y} = -\frac{2}{\sigma^2}\,\frac{x_0^{q-1}}{\alpha+q} \end{aligned}\right\}$$

Returning to (1), we get at the upper boundary the condition

$$V'(\bar{x}) = b^{1-R_1}\,\frac{F_+(s_+, m_+, 1-R_1, \rho, \bar{x})}{1-R_1} - \rho F_+(s_+, m_+, 0, \rho, \bar{x})\; V(\bar{x})$$

with $s_+ = \sigma K$, $m_+ = r + K(\mu - r) - b$; and we get the lower boundary condition

$$V'(\underline{x}) = \gamma_2^{1-R_2}\,\frac{F_-(s_-, m_-, 1-R_2, \rho, \underline{x})}{1-R_2} + \left(\frac{1}{1-R_1} - \frac{1}{1-R_2}\right)F_-(s_-, m_-, 0, \rho, \underline{x})$$
$$- \rho\, F_-(s_-, m_-, 0, \rho, \underline{x})\; V(\underline{x})$$

with $s_- = \pi_2\sigma$, $m_- = r + \pi_2(\mu - r) - \gamma_2$ consumption rate for Merton problem

7) There's another way we could try to handle things, which is to suppose that we have boundary at $x_0 < x_1$, and we do a controlled diffusion process down to x_0+, but when we get to x_0 then a fixed uncontrolled diffusion process takes over until we get to x_1. Thus

$$V(x_0) = E^{x_0}\left[\int_0^{\tau} e^{-\rho s} f(X_s)\, ds\right] + E^{x_0}\left[e^{-\rho\tau}\right]\; V(x_1)$$

where $\tau = \inf\{t : X_t = x_1\}$. So we could do this by modifying the first row of the Q-matrix:

$$\begin{pmatrix} -1 & E^{x_0} e^{-\rho\tau} & 0 & 0 \cdots \\ \cdot & - & - & - \\ & & & \end{pmatrix} V = \begin{pmatrix} -E^{x_0}\left[\int_0^{\tau} e^{-\rho s} f(X_s)\, ds\right] \\ \vdots \\ \vdots \end{pmatrix}$$

This is nice because it is wholly probabilistic!

Realistically, there won't be many cases we can do in closed form, probably just BM and log BM

(a) Suppose the generator is $\frac{1}{2} s^2 D^2 + m D$. Then the e-f's are $\exp(\beta x)$, $\exp(-\alpha x)$, where $-\alpha < 0 < \beta$ solve

$$Q(t) \equiv \tfrac{1}{2} s^2 t^2 + m t - \rho = 0$$

If we have the diffusion in $(-\infty, x_*]$ the killed resolvent applied to $f(x) \equiv \exp(\theta x)$ will be

$$R_\rho^{\partial} f(x) = -e^{\theta x}/Q(\theta) + e^{\beta(x - x_*) + \theta x_*}/Q(\theta)$$

$$= \frac{e^{\theta x_*}}{Q(\theta)} \left\{ e^{\beta(x-x_*)} - e^{\theta(x-x_0)} \right\}$$

If we work with the diffusion in $[x^*, \infty)$ we shall likewise have

$$R_\rho^\partial f(x) = \frac{e^{\theta x^*}}{Q(\theta)} \left\{ e^{-\alpha(x-x^*)} - e^{\theta(x-x^*)} \right\}$$

(b) If now we work with GBM, generator $\frac{1}{2}s^2x^2D^2 + mxD$, then the quadratic $\tilde{Q}(t) = \frac{1}{2}s^2t(t-1) + mt - \rho$ has roots $-\alpha < 0 < \beta$ and we have for $f(x) = x^\nu$ that in $(0, x_*]$

$$R_\rho^\partial f(x) = \frac{-x^\nu}{Q(\nu)} + \left(\frac{x}{x_*}\right)^\beta \frac{x_*^\nu}{Q(\nu)}$$

and in $[x^*, \infty)$ we will have

$$R_\rho^\partial f(x) = \frac{-x^\nu}{Q(\nu)} + \left(\frac{x}{x^*}\right)^{-\alpha} x_*^\nu / Q(\nu).$$

A model for fund management (3/3/11)

Here's a model for the evolution of the assets in a (long-only) fund, perhaps a unit trust. The level of the fund's investments will evolve like

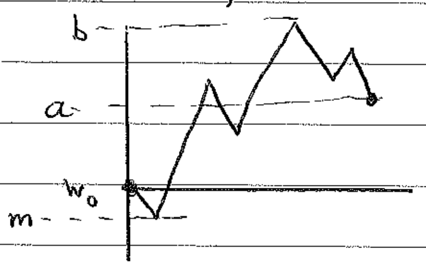
$$dw_t = r w_t\,dt + \theta_t\left(\sigma\,dW_t + (\mu - r)\,dt\right)$$

but the size of the assets under management is given by

$$A_t = \varphi(\bar{w}_t)\, h(w_t/\bar{w}_t)$$

where φ, h are increasing, $h(1) = 1$, and $\bar{w}_t = \sup\{w_s : s \le t\}$.

The fund's managers get paid a fee at time T based on both the total AUM, and the performance. How does it work?



Performance fee earned in this case

$$= \lambda \int_m^a \varphi'(x)\,(a-x)\,dx$$

$$= \lambda \int_m^a \varphi'(x) \int_x^a dy\,dx$$

$$= \lambda \int_m^a dy\,\{\varphi(y) - \varphi(m)\}$$

$$\equiv \lambda\, F(m, a)$$

There is a small fee for total AUM, so the reward paid to the fund manager at time T will be

$$\boxed{\varepsilon\, \varphi(\bar{w}_T)\, h(w_T/\bar{w}_T) + \lambda\, F(\underline{w}_T, w_T)}$$

where $\underline{w}_t \equiv \inf\{w_s;\, s \le t\}$.

Suppose we try $\varphi(x) = x^\nu$. Then

$$F(\underline{w}, w) = \lambda\, \underline{w}^{\nu+1} \left\{ \frac{(w/\underline{w})^{\nu+1} - 1}{\nu+1} - \left(\frac{w}{\underline{w}} - 1\right) \right\}$$

so if $\underline{w} = q\,w$ we get

$$\boxed{F(\underline{w}, w) = \lambda w^{\nu+1} q^{\nu+1} \left\{ \frac{q^{-\nu-1} - 1}{\nu+1} - (q^{-1} - 1) \right\}}$$

Not obvious how to do the optimization here. Any easier if $r = 0$?

Merton problem with option to stop early (4/3/11)

1) This is like the standard Merton optimal investment/consumption problem, except that a stopping time τ of our choosing we may stop and receive reward $U_0(W_\tau)$. What we therefore will have for the value function $V(w)$ is that $V(w) \geq U_0(w)$ for all w, and

$$0 \geq \sup \left[-\rho V + (rw + \theta(\mu - r) - c) V_w + \tfrac{1}{2}\sigma^2 \theta^2 V_{ww} + U(c) \right].$$

Now when we go across to the dual variables, we shall have as usual

$$\tilde{U}(z) + \tfrac{1}{2}\kappa^2 z^2 J'' + (\rho - r) z J' - \rho J \leq 0$$

with equality where z is the slope of tangent to V in some place where we optimally chose not to stop. So altogether this is like an optimal stopping problem for the diffusion Z_t with generator

$$\tfrac{1}{2}\kappa^2 z^2 D_z^2 + (\rho - r) z D_z$$

with running objective $\int_0^\tau e^{-\rho s} \tilde{U}(Z_s)\, ds$ and stopping objective $\tilde{U}_0(z)$.

2) Is it evident that V is concave? In general it will not be, which means that identifying the dual function isn't solving the problem? But if we ever got $V_{ww} > 0$, we would do $\theta = +\infty$ for a while, so in fact passing to the dual will be OK.

3) Suppose that we have $U(w) = w^{1-R}/(1-R)$ for some $R > 1$, and we select $U_0(w) = A\, w^{1-R_0}/(1-R_0)$ where $R_0 > R$. What we expect therefore is that there will be a critical value of wealth w_* so that we stop as soon as w rises to w_* or higher. In the dual formulation, this means there is some z_* such that

$$J(z) = \begin{cases} \tilde{U}_0(z) & (z \leq z^*) \\[1em] -\dfrac{\tilde{U}(z)}{Q(1-1/R)} + B\left(\dfrac{z}{z_*}\right)^{-\alpha} & (z > z^*) \end{cases}$$

where the quadratic $Q(t) = \tfrac{1}{2}\kappa^2 t(t-1) + (\rho - r)t - \rho$ has roots $-\alpha < 0 < \beta$.

Expect to see a C^1 join at z^*. $\tilde{U}_0(y) = A\, \tilde{u}(y/A)$ where $\tilde{u}(y) = -y^{1-1/R_0}/(1-1/R_0)$.

This has a fairly explicit solution: we find that z_* must satisfy

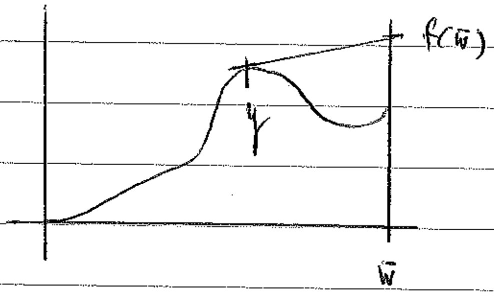
$$\boxed{\left(1 - \frac{1}{R_0} + \alpha\right) \tilde{U}_0(z_*) = -\frac{\tilde{U}(z_*)}{Q(1-1/R)} \left\{1 - \frac{1}{R} + \alpha\right\}}$$

which uniquely fixes z_*.

Fund management model again (8/3/11)

(1) We can firstly study a simpler problem where we seek to optimally stop a BM which gets stopped at 0, when the stopping reward is $\varphi(w, \bar{w})$. Suppose we have $V(\bar{w}, \bar{w}) \equiv f(\bar{w})$. If we are on an excursion down from the maximum $\bar{w}$, the optimal stopping problem would stop when either we return to $\bar{w}$, or we fall to $\eta(\bar{w})$, where $\eta(\bar{w}) = y$ is the value of y which minimizes the slope

$$\frac{f(\bar{w}) - \varphi(y, \bar{w})}{\bar{w} - y}$$



If we just consider the martingale $M_t \equiv V(w_{t\wedge T}, \bar{w}_{t\wedge T})$ time changed by $\tau_a = \inf\{t : \bar{w}_t > a\}$ then we have

$$M_{\tau_t} = f(t)\, I_{\{\tau_t < T\}} + \varphi(\eta(\bar{w}_T), \bar{w}_T)\, I_{\{\tau_t \geq T\}}.$$

Since the martingale is experiencing a drop of size $f(t) - \varphi(\eta(t), t)$ with intensity $1/(t - \eta(t))$, for M to be a martingale it must be that

$$\boxed{f'(t) = \frac{f(t) - \varphi(\eta(t), t)}{t - \eta(t)}}$$

So f solves a nonlinear first-order ODE (because $\eta(t)$ depends on $f(t)$ in a fairly complicated way).

(2) If we have the situation considered earlier with stopping reward $\varphi(w, \underline{w}, \bar{w})$, we need to consider $f_+(\bar{w}, \underline{w}) = V(\bar{w}, \underline{w}, \bar{w})$ and $f_-(\bar{w}, \underline{w}) = V(\underline{w}, \underline{w}, \bar{w})$. If we are at $w = \bar{w}$ with $\underline{w}$ the current value of the min, we can define $\eta(\bar{w}, \underline{w})$ to be the value y which minimizes the slope

$$\frac{f_+(\bar{w}, \underline{w}) - \varphi(y, \underline{w}, \bar{w})}{\bar{w} - y}$$

and similarly $\xi(\bar{w}, \underline{w})$ is the value of y which maximizes the slope

$$\frac{-f_-(\bar{w}, \underline{w}) + \varphi(y, \underline{w}, \bar{w})}{y - \underline{w}}$$

By analogous reasoning to that applied in (1), we can deduce that

$$\boxed{\begin{aligned}\frac{\partial f_+}{\partial \bar{w}}(\bar{w},\underline{w}) &= \frac{f_+(\bar{w},\underline{w}) - \varphi(\eta(\bar{w},\underline{w}),\underline{w},\bar{w})}{\bar{w} - \eta(\bar{w},\underline{w})} \\ \frac{\partial f_-}{\partial \underline{w}}(\bar{w},\underline{w}) &= + \frac{\varphi(\xi(\bar{w},\underline{w})) - f_-(\bar{w},\underline{w})}{\xi(\bar{w},\underline{w}) - \underline{w}}\end{aligned}}$$

How do we determine boundary conditions?

(3) Notice that this analysis does not account for the possibility that the individual optimal stopping problems might involve continuing until $\bar{w}$ or $\underline{w}$ is reached. In that case we would have

$$\begin{cases} \dfrac{\partial f_+}{\partial \bar{w}} = \dfrac{f_+(\bar{w},\underline{w}) - f_-(\bar{w},\underline{w})}{\bar{w} - \underline{w}} \\[2ex] \dfrac{\partial f_-}{\partial \underline{w}} = + \dfrac{f_+(\bar{w},\underline{w}) - f_-(\bar{w},\underline{w})}{\bar{w} - \underline{w}} \end{cases}$$

(4) Back to the case of stopping at the max only. If $\eta(x)$ is the tangent point to φ when the max is x, we shall have

$$\boxed{f(x) = \varphi(\eta(x), x) + \varphi_y(\eta(x), x)(x - \eta(x))}$$

as well as $f'(x) = \varphi_y(\eta(x), x)$. Differentiating this wrt x gives us

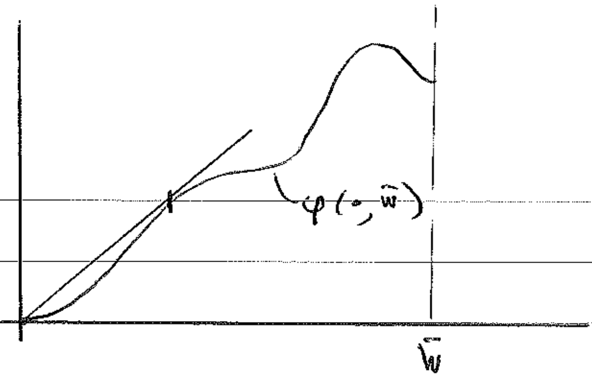
$$f'(x) = \frac{d\eta}{dx} \cdot \varphi_{yy} \cdot (x-\eta) + \varphi_x + \varphi_{xy}(x-\eta) + \varphi_y$$

hence

(1) $$\boxed{\frac{d\eta}{dx} = - \frac{\varphi_x + \varphi_{xy}(x-\eta)}{(x-\eta)\varphi_{yy}}}$$

at least while the optimal stopping problem on the slice $\{x : x \leq \bar{x}\}$ is solved internally

(5) How would this get started? Suppose we begin at $W_0 = \bar{W}_0 = 1$, and suppose that $\varphi(0, \bar{w}) = 0$ for all $\bar{w}$; once the process reaches 0, there is no



value left. What I expect is that for $\bar w \le \xi_0$ we do not stop, but that once $\bar w$ has got to ξ we may then stop.

How do we decide ξ? For $\bar w \le \xi_0$ we shall have that $V(w,\bar w) = w$ (since up till the first hitting time of ξ_0, w_t is a martingale) so we shall want that for any $0 \le w \le \bar w \le \xi_0$ we get

$$\varphi(w,\bar w) \le w$$

So if we set

$$\Lambda = \sup_{w,\bar w} \frac{\varphi(w,\bar w)}{w}$$

which we assume is finite, and attained at $\bar w = \xi_0$, $w = \eta_0$, then what we get is that $V(w,\bar w) = w$ for $0 \le w \le \bar w \le \xi_0$, and then the upper exercise boundary $\eta(\cdot)$ starts at η_0 when $\bar w = \xi_0$, and solves the ODE (1) from that initial condition. There is also going to be a lower exercise boundary $\zeta(\bar w)$ which is characterised by

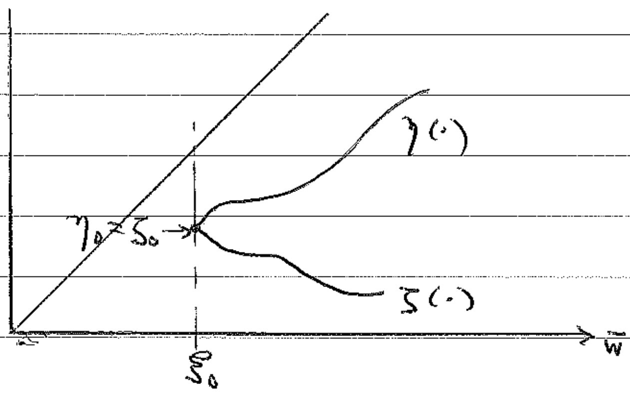
$$\frac{\varphi(\zeta,\bar w)}{\zeta} = \varphi_y(\zeta,\bar w)$$

that is, $\varphi(\zeta,\bar w) - \zeta\varphi_y(\zeta,\bar w) = 0$. Differentiating wrt $\bar w$ gives

$$-\zeta\frac{d\zeta}{d\bar w}\varphi_{yy} + \varphi_x - \zeta\varphi_{xy} = 0$$

so

$$\boxed{\frac{d\zeta}{d\bar w} = \frac{\varphi_x - \zeta\varphi_{xy}}{\zeta\varphi_{yy}}}$$



Assume wlog $r = 0$.

If $A = \sum a_j$, we deduce that for $t \geq \tau_J$

$$(\Pi)^{1+A} = \left(\Pi \xi_j(t)^{-a_j}\right)^{(1+A)} = e^{A(\rho - r)t} y_0^A \, \Pi(p_j/a_j)^{a_j}$$

$$\Rightarrow (\Pi) = y_0^{A/(1+A)} e^{\rho A t/(1+A)} K, \qquad K = \Pi(p_j/a_j)^{a_j/(1+A)}$$

$$\Rightarrow \boxed{\xi_j = K \frac{a_j}{p_j} y_0^{-1/(1+A)} e^{-\rho t/(1+A)}}$$

History-dependent preferences: a special case (23/3/11)

Suppose we work with a deterministic form of the problem, where

$$\dot{w}_t = rw_t - p\cdot c_t\,dt, \qquad \dot{\xi}_j = \lambda_j(c_j - \xi_j)$$

and we seek

$$\max \int_0^\infty e^{-\rho t} U(\xi)\,dt$$

where $U(\xi) = -\prod_j \xi_j^{-a_j}$, and $\sum a_j = b$. If we write this in P-L form, we get

$$\sup \int_0^\infty \{ e^{-\rho t} U(\xi_t) + y_t(rw_t - p\cdot c_t) + w_t\dot{y}_t + \eta_t\Lambda(c-\xi) + \xi\dot{\eta}\}dt + y_0w_0 + \eta_0\xi_0$$

$$= \sup \int_0^\infty \{e^{-\rho t}U(\xi_t) + \xi_t(\dot{\eta}_t - \Lambda\eta_t) + w_t(\dot{y}_t + ry_t) + \eta_t\Lambda c - y_t p c\}dt + -$$

so we shall have

$$\boxed{\begin{aligned} &\dot{y}_t + ry_t = 0 \\ &\lambda_j\eta_j(t) \le p_j y(t) \\ &e^{-\rho t}\frac{a_j}{\xi_j}(\Pi) + \dot{\eta}_j - \lambda_j\eta_j = 0 \end{aligned}}$$

(assuming $w_t > 0$ always)
(equality when $c_j(t) > 0$)

where (Π) is short for $\Pi\,\xi_i^{-a_i}$. So we have $y_t = y_0 e^{-rt}$, and then $\eta_j(t) \le p_j e^{-rt} y_0/\lambda_j$. It seems obvious that once you get $\xi_j(t)$ down to such a low level that you want to buy then you will buy only just enough to stay at the margin, and therefore you will be buying thereafter always. So we expect that for some τ_J we shall have $\eta_j(t) = p_j y(t)/\lambda_j$ $\forall t \ge \tau_J$, $\forall j$ and hence for $t \ge \tau_J$

$$e^{-\rho t}\frac{a_j}{\xi_j(t)}(\Pi) = (\lambda_j + r)\frac{p_j y_0}{\lambda_j}e^{-rt}$$

So for $t \ge \tau_J$ must have

$$\boxed{\xi_j(t) = e^{(r-\rho)t}(\Pi)\cdot\frac{\lambda_j a_j}{y_0 p_j(\lambda_j + r)}}$$

In particular, after τ_J, we get $\xi_i(t)/\xi_k(t)$ never changes. So there are some ideal ratios for ξ_i/ξ_k, indeed, we want ideally that

Value function approach: $V(w,\xi) = w^{-A} v(\xi/w) \equiv w^{-A} v(x)$
for $A = \sum a_j$: we find that

$$\max_j \Big[\lambda_j D_j v + A v + x \cdot Dv \Big] \vee \{ U(x) - \rho v(x) - x \cdot \Lambda Dv(x) \} = 0$$

If one of the first terms is zero, we can be doing a jump j; for any j which gives a strictly negative term, there can be no purchasing of asset j

To solve the DE $\quad U(x) - \rho v(x) - x \cdot \Lambda Dv(x) = 0 \quad$ in two dimensions
we could introduce new variables

$$z_1 = x_1^{\lambda_2} x_2^{-\lambda_1}, \quad z_2 = x_1^{-a_1} x_2^{-a_2}$$

and write

$$v(x) = g(z_1, z_2)$$

so that

$$x \cdot \Lambda Dv(x) = (-\lambda_1 a_1 \;\; -\lambda_2 a_2)\; z_2 D g(z)$$

and the ODE is

$$-z_2 - \rho g + (\lambda_1 a_1 + \lambda_2 a_2) z_2 D_2 g(z) = 0$$

so if we set $k \equiv \rho / (\lambda_1 a_1 + \lambda_2 a_2)$ we get

$$\frac{\partial g}{\partial z_2} - \frac{k}{z_2} g = \frac{1}{(\lambda_1 a_1 + \lambda_2 a_2)}$$

$$\Rightarrow \quad \frac{\partial}{\partial z_2}\left(z_2^{-k} g \right) = z_2^{-k} / (\lambda_1 a_1 + \lambda_2 a_2)$$

$$\Rightarrow \quad z_2^{-k} g = h(z_1) + \frac{z_2^{1-k}}{(1-k)(\lambda_1 a_1 + \lambda_2 a_2)}$$

$$\Rightarrow \quad \boxed{g(z_1, z_2) = z_2^{k} h(z_1) - \frac{z_2}{\rho - \lambda_1 a_1 - \lambda_2 a_2}}$$

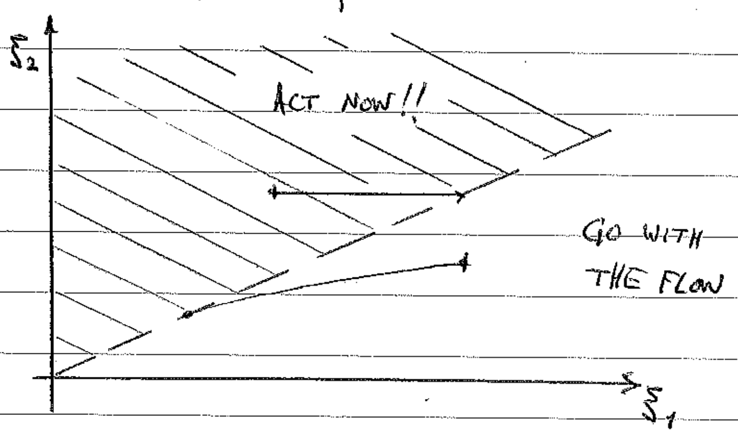
If we were so fortunate as to have $h \equiv 0$, we'd be ok, else not much hope.

$$\boxed{\frac{\xi_i}{\xi_k} = \frac{\lambda_i a_i}{p_i(\lambda_i + r)} \frac{p_k(\lambda_k + r)}{\lambda_k a_k}} \equiv q_{ik}, \text{ say.}$$

Notice that without any buying, $\xi_i(t)/\xi_k(t) = \text{const} \exp((\lambda_k - \lambda_i)t)$
So if we start with

$$\xi_i(0)/\xi_k(0) > q_{ik}$$

and $\lambda_k > \lambda_i$, then just leaving these alone will take you away from the ideal proportions, so what I think happens is that you would immediately buy some of commodity k to bring yourself to $\xi_i(0)/\xi_k(0+) = q_{ik}$
On the other hand, if you start with $\xi_i(0)/\xi_k(0) > q_{ik}$ and now $\lambda_k < \lambda_i$, you can just sit tight and let depreciation bring you in to the ideal ratio. So the picture for two dimensions, where $\lambda_1 > \lambda_2$, is



A dynamic contracting problem with risk aversion (24/3/11)

(1) Suppose wealth process w evolves as

$$dw_t = w_t\{\sigma dW_t + a_t dt\} - (c_t + q_t)dt$$

where a_t is the agent's effort ($a_t \geq 0$), c_t is principal's rate of consumption, and q_t is the rate at which wages are paid to the agent. We'll suppose that the principal may choose to terminate at some time τ, so that his objective is

$$\sup E^w\left[\int_0^\tau e^{-\rho s}U(c_s)ds + e^{-\rho\tau}U_F(w_\tau)\right] \equiv F(w)$$

and the agent's objective is

$$\sup E^w\left[\int_0^\tau e^{-\gamma s}\varphi(q_s, a_s)ds - Ae^{-\gamma\tau}\right] \equiv G(w)$$

where φ is concave inc in q, concave dec in a; we might as well take

$$\varphi(q,a) = -q^{-\nu}a^{\beta}$$

for some $\nu > 0$, $\beta > 1$.

(2) If we think about the agent's HJB, he will have

$$0 = \sup_a\left[-\gamma G + \tfrac{1}{2}\sigma^2 w^2 G'' + (aw - c - q)G' + \varphi(q,a)\right]$$

which would lead to FOCs

$$wG' = \beta q^{-\nu}a^{\beta-1} \Rightarrow \boxed{a = \left(\frac{wG'q^{\nu}}{\beta}\right)^{1/\beta-1}.}$$

Then we would see the equations

$$\boxed{0 = -\gamma G + \tfrac{1}{2}\sigma^2 w^2 G'' - (c+q)G' + (wG')^{\beta/\beta-1} q^{\nu/\beta-1}\left\{\beta^{-1/\beta-1} - \beta^{-\beta/(\beta-1)}\right\}}$$

Looking at the HJB for the principal would give

$$0 = \sup_{q,c}\left[-\rho F + U(c) + (aw - c - q)F' + \tfrac{1}{2}\sigma^2 w^2 F''\right] \qquad (*)$$

and it's TEMPTING BUT WRONG just to optimize over c, q using the form of a as a function of q given above; the problem is that G' in a will depend on c, q, so the dependence of $(*)$ on c, q is extremely complicated...

Bayesian analysis turned around (26/3/11)

(1) In a Bayesian analysis, it's very common to suppose that $Y \sim N(\mu, V)$ where we put a $N(0, v)$ prior on μ. Thus we suppose that when we see Y, we represent $Y = \mu + \varepsilon$ for $\varepsilon \sim N(0,V)$, and then we deduce that the mean $E(\mu|Y) = \theta Y$, and the conditional dist^n is Gaussian $[\theta = E(\mu Y)/EY^2]$.

(2) It becomes harder to do much if the prior for μ is not Gaussian, because the conditional law for $(\mu|Y)$ is a mess. But we can do something else. What we want is to be able to choose from a wide class of distributions for μ, and to say that the conditional law of $(\mu|Y)$ is simple. Here's a way to do it (assume for now all RVs have mean 0, the extension is not difficult). We shall write

$$\mu = \theta Y + Z$$

for some (zero-mean) RV Z independent of Y! With this, it is clear that $E[\mu|Y] = \theta Y$, exactly as for the Gaussian case. But can we fix it so that $E[Y|\mu] = \mu$?? Yes we can! Suppose that the CGF of Y is ψ_Y, and of Z is ψ_Z. We want to have $\forall k \in \mathbb{R}^d$ that

$$0 = E\left[(\mu - Y)e^{k\cdot\mu}\right] = E\left[(Z-(1-\theta)Y)e^{k\cdot(Z+\theta Y)}\right]$$

$$= e^{\psi_Z(k)+\psi_Y(\theta k)}\left\{(D\psi_Z)(k) - (1-\theta)(D\psi_Y)(\theta k)\right\}$$

which we achieve if $\boxed{\psi_Z(k) = \frac{1-\theta}{\theta}\psi_Y(\theta k)}$

(3) Suppose now that $\theta Y \overset{D}{=} X_1$, where $(X_t)_{t\geq 0}$ is a Lévy process with characteristic exponent $\psi \equiv \psi_X$. What the condition on the law of Z says is that $Z \sim X((1-\theta)/\theta)$, and therefore

$$E\, e^{k\cdot\mu} = E\, e^{k\cdot\theta Y + k\cdot Z} = \exp\left\{\psi(k) + \frac{1-\theta}{\theta}\psi(k)\right\}$$

$$= \exp\left\{\frac{1}{\theta}\psi(k)\right\}$$

so that the law of μ is the law of $X_{1/\theta}$, and the structure is revealed. So in this way, μ may have any infinitely-divisible law, and the conditional dist^n of

$\mu | Y$ is $\theta Y + \mathcal{L}(Z)$ where $Z \sim X(\frac{1}{\theta} - 1)$.

(4) If we write $Y = \mu + \varepsilon$ in the oldway, we no longer have ε independent of μ. What do we know about the covariance?

$$\cancel{E[\mu \varepsilon^T] = E[(\theta Y + Z)((1-\theta)Y - Z)^T]}$$

$$= \theta(1-\theta) \operatorname{cov}(Y) - \operatorname{cov} Z.$$

In the Lévy case, what we see here is $\theta^{-1}(1-\theta) \operatorname{cov}(X_1) - (\frac{1}{\theta} - 1) \operatorname{cov}(X_1) = 0$ so that μ and ε are uncorrelated even though not independent.

(5) How would some examples look? If the Lévy process is BM, it's just the usual story. Another example we might like to try would be variance-gamma:

$$\int_0^\infty \frac{(\beta v)^{t-1}}{\Gamma(t)} \beta \, dv \, \exp\{-|x|^2/2v\} (2\pi v)^{-d/2} e^{-\beta v}$$

$$= K_{t-d/2}(\sqrt{2\beta}\,|x|) \, |x|^{t-d/2} \left(\frac{\beta}{2}\right)^{(2t+d)/4} \frac{2}{\pi^{d/2} \Gamma(t)}$$

The characteristic function of this density will be

$$\left(\frac{\beta}{\beta + \frac{1}{2}|k|^2}\right)^t$$

If we write $\varphi_t(x;\beta)$ for the VG density at time t, we see that

$$\varphi_t(x;\beta) = K_{t-\frac{d}{2}}(\sqrt{2\beta}\,|x|)(\sqrt{2\beta}\,|x|)^{t-d/2} \frac{2\,\beta^{d/2}}{2^t \, \pi^{d/2} \Gamma(t)}$$

$$= K_{t-\frac{d}{2}}(\sqrt{2\beta}\,|x|)(\sqrt{2\beta}\,|x|)^{t-d/2} \frac{2\,\beta^{d/2}}{2^{t-d/2}\,\Gamma(t)} \cdot \frac{1}{(2\pi)^{d/2}}$$

we know that $\int e^{ik\cdot x} \varphi_t(x;\beta)\,dx = (\beta/(\beta + \frac{1}{2}|k|^2))^t$. Now provided this last is integrable, that is, $2t > d$, we can recover the original density by inverse FT. However, it doesn't seem that convoluting two t-dist^ns gives you anything very pleasant...

(6) The basic story seems to be to take some Lévy process (ξ_t), suppose that $\mu \sim \xi_T$, and write μ as

$$\mu = \theta Y + Z$$

where $Z \sim \xi_{T-a}$, $\theta Y \sim \xi_a$ and we insist that $\theta = a/T$ so as to keep $E(Y|\mu) = \mu$.

How does this look in a dynamic context? We shall suppose that

$$(\mu_t | y_t) \sim \hat{\mu}_t + \xi_{\tau_t}$$

and that we move μ_t using an increment of time ε of the Lévy process to get to μ_{t+1}. Thus

$$(\mu_{t+1} | y_t) \sim \hat{\mu}_t + \xi_{\tau_t + \varepsilon}$$

Now we write

$$\mu_{t+1} - \hat{\mu}_t = \theta (Y_{t+1} - \hat{\mu}_t) + Z$$

where $Z \sim \xi_{\tau_t + \varepsilon - \alpha_t}$, and $\theta (Y_{t+1} - \hat{\mu}_t) \sim \xi_{\alpha_t}$ where we demand as before that

$$\frac{\alpha_t}{\tau_t + \varepsilon - \alpha_t} = \theta \quad \Rightarrow \quad \alpha_t = \frac{\theta(\tau_t + \varepsilon)}{1 + \theta}$$

Therefore

$$\tau_{t+1} = \tau_t + \varepsilon - \alpha_t = \frac{\tau_t + \varepsilon}{1 + \theta}$$

and in the steady state, $\boxed{\tau_t = \varepsilon / \theta}$ So we have in the steady state that

$\alpha = \varepsilon$

$$\boxed{(\mu_t | y_t) \sim \hat{\mu}_t + \xi_{\varepsilon/\theta}}$$

where $\boxed{\hat{\mu}_{t+1} = \theta Y_{t+1} + (1-\theta) \hat{\mu}_t}$

What could be nicer?! Once you've chosen the basic Lévy process, you just have the two parameters ε, θ to select.

(7) How would a Bayesian model comparison look? To understand this we need to keep track of the true likelihood; we can't just ignore the constants out front. So suppose we've got to time t, and the likelihood for $\mu_t = x$ is

$$\lambda_t \, \varphi_{\varepsilon/\theta}(x - \hat{\mu}_t)$$

Then the likelihood that $\mu_{t+1} = x$ will be

$$\lambda_t \int \varphi_{\varepsilon/\theta} (x' - \hat{\mu}_t)\, \varphi_\varepsilon (x - x')\, dx' = \lambda_t\, \varphi_{\varepsilon(1+\theta^{-1})} (x - \hat{\mu}_t)$$

Now there are two random variables $\eta \equiv \theta(Y_{t+1} - \hat{\mu}_t)$ and Z with laws ξ_ε and $\xi_{\varepsilon/\theta}$ which add to $x - \hat{\mu}_t$; so they ~~cannot be~~ taken independently they have to be conditioned on summing to $x - \hat{\mu}_t$. So we get

$$\lambda_t\, \varphi_{\varepsilon(1+\theta^{-1})} (x - \hat{\mu}_t) . \frac{\varphi_\varepsilon(\eta)\, \varphi_{\varepsilon/\theta}(-\eta + (x - \hat{\mu}_t))}{\varphi_{\varepsilon+\varepsilon/\theta}(x - \hat{\mu}_t)}$$

$$= \lambda_t\, \varphi_\varepsilon(\eta)\, \varphi_{\varepsilon/\theta}(x - \hat{\mu}_t - \eta)$$

Hence we deduce that

$$\boxed{\lambda_{t+1} = \lambda_t\, \varphi_\varepsilon\big(\theta(Y_{t+1} - \hat{\mu}_t)\big)}$$

which will be quite easy to do.

(8) However, the issue is this: "<u>What is the implied law of the process (Y_t) which evolves in this way?</u>"

(9) Let's look at the basic linear Gaussian model

$$\begin{cases} X_{t+1} = A X_t + \varepsilon_{t+1} \\ Y_{t+1} = C X_{t+1} + \eta_{t+1} \end{cases}$$

Here we get a Gaussian $N(\hat{X}_t, V_t)$ posterior for X_t given $\mathcal{Y}_t$, and the updating is

$$\hat{X}_{t+1} - A\hat{X}_t = K_{t+1}(Y_{t+1} - CA\hat{X}_t) \equiv K_{t+1}\nu_{t+1}$$

for known K_t (which converge to some steady-state if system is observable). So if we go into the observation filtration, the story is just

$$\begin{cases} \hat{X}_{t+1} = A\hat{X}_t + K_{t+1}\nu_{t+1} \\ Y_{t+1} = CA\hat{X}_t + \nu_{t+1} \end{cases}$$

All I'm really saying here is that we make the dynamics of $(\hat{X}, Y)$ and then pretend there is some $X_t = \hat{X}_t + \xi_t$, where the ξ_t are independent noises. It's a bit flat as a modelling story, but it's <u>equivalent</u> for the Gaussian case, and is computationally feasible else.

Utility from possession again (8/4/11)

(1) Let's look again at the (deterministic) story on pp 35-36. To begin with, notice that by working with discounted w, ξ, c, we may (and shall) suppose that $\boxed{r=0}$. We could also (by considering $p_k \xi_k(t)$) reduce to the case $\boxed{p_k = 1 \text{ for all } k.}$ Let's make these simplifications.

(2) Now let's look at a simpler problem where we may have $c_j < 0$, so long as we keep all the ξ_j positive. The PL approach leads to dual-feasibility conditions

$$\begin{cases} \eta_j(t) = y_0/\lambda_j \\ e^{\rho t} \dfrac{a_j}{\xi_j(t)} (\Pi) = \lambda_j \eta_j(t) = y_0 \end{cases}$$

So we deduce that

$$\boxed{\xi_j(t) = K a_j \, y_0^{-1/(1+A)} e^{-\rho t/(1+A)}} \qquad \left[K \equiv \prod a_j^{-a_j/(1+A)} \right]$$

Thus the ξ_j all decline exponentially at rate $e^{-\rho t/(1+A)}$, as do the c_j and therefore $w(t)$. But notice that

$$\lambda_j c_j(t) = \dot{\xi}_j + \lambda_j \xi_j = \left(\lambda_j - \frac{\rho}{1+A}\right) \xi_j$$

So if $\boxed{\lambda_j < \rho/(1+A)}$ you would choose $c_j(t) < 0$ at all times. This implies that the 'steady state' solution where all the ξ_j decline at the same exponential rate will in general be impossible if we insist $c_j \geq 0$.

(3) If we go back to the original problem with $c_j(t) \geq 0 \; \forall t$, then we expect that all the 'slow' commodities ($\lambda_j \leq \rho/(1+A)$) will never be bought, so the stocks of these simply decay exponentially. The effect of this can be absorbed into a changed ρ, so it seems we may do this and just suppose that $\boxed{\lambda_j > \rho/(1+A)}$ for all j ...? But this cannot be correct: if all the commodities were slow, you would still want to spend some money buying something!

(4) Let's do the stochastic version where we drop the requirement that the transfers to ξ must be non-decreasing. Again supposing that $r=0$, we get the dynamics for wealth and for the ξ^j to be just

$$dw_t = \theta_t (\sigma dW_t + \mu dt) - 1 \cdot dC_t, \qquad d\xi_t = \Lambda (dC_t - \xi_t dt)$$

So now if we try the PL approach with multiplier processes ζ_t, η^j_t with $d\zeta_t = \zeta_t(\alpha_t dW_t + \beta_t dt)$, our Lagrangian objective will be

$$\max E\Big[\int_0^\infty e^{-\rho t} U(\xi_t)dt + \zeta_t(\theta_t \mu dt - 1 \cdot dC_t) + W_t d\zeta_t + d\zeta_t dW_t$$

$$+ \eta_t \cdot \Lambda(dC_t - \xi dt) + \xi \cdot d\eta + d\xi \cdot d\eta\Big] + -$$

$$= \max E\Big[\int_0^\infty e^{-\rho t} U(\xi_t) dt + \zeta_t(\theta_t \mu dt - 1 \cdot dC_t) + W_t \zeta_t \beta_t dt + \zeta_t \alpha_t \theta_t \sigma dt$$

$$+ \eta \cdot \Lambda(dC_t - \xi dt) + \xi \cdot d\eta + d\xi \cdot d\eta\Big] + -$$

so from the coefficient of θ we get $\mu + \sigma\alpha_t = 0$, coefficient of W_t gives $\beta = 0$, coefficient of dC gives $\lambda_j \eta^j_t = \zeta_t$, so ζ is the SPD, and we get

$$= \max E \int_0^\infty \left(e^{-\rho t} U(\xi_t) - \eta_t \cdot \Lambda \xi_t\right) dt + -$$

so the optimality FOC gives $\boxed{e^{-\rho t} \frac{a_j}{\xi^j_t}(\Pi) = \zeta_t}$

and so

$$(\Pi) = \left(\prod_j a_j^{-a_j}\right)\left(e^{-\rho t}(\Pi)/\zeta_t\right)^{-A} \equiv K^{1+A}\left(e^{-\rho t}(\Pi)/\zeta_t\right)^{-A}$$

$$\Rightarrow (\Pi)^{1+A} = K^{1+A}\left(e^{-\rho t}/\zeta_t\right)^{-A}$$

$$\Rightarrow (\Pi) = K\left(e^{\rho t}\zeta_t\right)^{A/(1+A)}$$

$$\Rightarrow \boxed{\xi^j_t = a_j K \left(e^{\rho t}\zeta_t\right)^{-1/(1+A)}}$$

Hedge fund problem again (8/4/11)

(1) The first formulation of this is a bit questionable, in that it fails to account well for the basis level at which the investors come in and go out. In particular, if you think what happens on an excursion down from the max, as the level climbs back towards the max you have to get new assets come in at the current level, so if we're to have a path-independent profile of basis levels, we would need that as the level falls we only have exits at the current level.

(2) One way to achieve this is as follows. Suppose that $\varphi(\bar{w})$ is the total AUM when w first reaches new max $\bar{w}$. Then the profile of basis levels will be

$$f(w) = \varphi'(w) \qquad (w \leq \bar{w})$$

(at least above w_0 - we'll describe what happens there presently)

I propose that if the level falls down to $x < \bar{w}$, then the profile of basis levels should be

$$f(w) = \begin{cases} \varphi'(w) & (w \leq x) \\ p\,\varphi'(w) & (x < w \leq \bar{w}) \end{cases}$$

for some fixed $p \in (0,1)$. Thus a fraction $(1-p)$ of the AUM at a given basis will jump out when the level drops through that basis. Note this will not involve payment of any performance fee - the assets go out with the same value they went in.

Of course, this means that if the level falls to zero, there will still be assets in the fund, but this is OK - when a hedge fund crashes, people lose money.

(3) At the start, let's suppose that there are assets $\varphi(w_0)$ all with basis w_0. If the level falls to $x < w_0$, let's suppose that we reduce the atom at w_0 by $(1-p)(\varphi(w_0) - \varphi(x))$ (the assets which would have exited if the profile f held for all w), and the profile in (x, w_0) is just $p\,\varphi'$. Thus the final profile of basis levels if we end up with max $\bar{w}$, min $\underline{w}$ and terminal value x will be different depending on whether $x > w_0$ or $x < w_0$. We shall have

for $x > w_0$,

$$f(w) = \begin{cases} 0 & w < \underline{w} \\ p\,\varphi'(w) & \underline{w} \leq w < w_0 \\ p\,\varphi(w_0) + q\,\varphi(\underline{w}) & \text{at } w = w_0 \\ \varphi'(w) & w_0 < w \leq x \\ p\,\varphi'(w) & x < w \leq \bar{w} \\ 0 & \bar{w} < w \end{cases}$$

and if $x < w_0$

$$f(w) = \begin{cases} p\,\varphi'(w) & (\underline{w} < w < \bar{w}) \\ + \{p\,\varphi(w_0) + q\,\varphi(\underline{w})\} & \text{at } w_0 . \end{cases}$$

Thus in the first case, the AUM is

$$p\,\varphi(w_0) + q\,\varphi(\underline{w}) + p(\varphi(w_0) - \varphi(\underline{w})) + \varphi(x) - \varphi(w_0)$$
$$+ p(\varphi(\bar{w}) - \varphi(x))$$

$$= (2p-1)(\varphi(w_0) - \varphi(\underline{w})) + p\,\varphi(\bar{w}) + q\,\varphi(x),$$

and the peformance fee will be

$$\alpha \left\{ \int_{\underline{w}}^{w_0} (x-w)\, p\, \varphi'(w)\, dw + (x-w_0)(p\,\varphi(w_0) + q\,\varphi(\underline{w})) \right.$$
$$\left. + \int_{w_0}^{x} (x-w)\, \varphi'(w)\, dw \right\}$$

$$= \alpha \left\{ p\left((x-w_0)\varphi(w_0) - (x-\underline{w})\varphi(\underline{w}) + \int_{\underline{w}}^{w_0} \varphi(y)\, dy \right) \right.$$
$$\left. + (x-w_0)(p\,\varphi(w_0) + q\,\varphi(\underline{w})) + \int_{w_0}^{x} (\varphi(y) - \varphi(w_0))\, dy \right\}$$

For the second AUM will be

$$p(\varphi(\bar{w}) - \varphi(\underline{w})) + p\,\varphi(w_0) + q\,\varphi(\underline{w})$$

and the peformance fee will be

$$\alpha \int_{\underline{w}}^{x} (x-w)\, p\, \varphi'(w)\, dw = \alpha p \int_{\underline{w}}^{x} (\varphi(y) - \varphi(\underline{w}))\, dy$$

Variants of the earlier dynamic contracts example (11/4/11)

(1) Let's look at the problem where the output process Y evolves as

$$dY_t = \sigma\, dX_t + (1-\varepsilon_t)\, dt$$

where ε_t is the agent's control $(0 \le \varepsilon_t \le \beta)$ and the agent has an outside option worth R: when the residual value of the contract falls below R, at time H, the agent jumps out. So the agent's objective is

$$E\left[\int_0^{\tau\wedge H} e^{-\gamma s}(dU_s + \lambda \varepsilon_s\, ds) + e^{-\gamma(\tau\wedge H)} R \right]$$

The principal does not know R, but has some prior density φ for the value of R, with cdf F. Let's suppose that

$$F(x) = \exp\left(-\int_x^\infty h(s)\, ds\right)$$

so that as residual value hits new minima, $h(\cdot)$ is the hazard rate for stopping.

(2) Let's suppose X is a P^0-BM, and the agent's control is expressed in terms of the change-of-measure martingale Λ^ε,

$$d\Lambda^\varepsilon_t = \Lambda^\varepsilon_t\, (-\varepsilon_t/\sigma)\, dX_t .$$

We try to model the residual value process z by

$$dz_t = g(z_t, \underline{z}_t)\, dX_t + b(z_t, \underline{z}_t)\, dt - dL_t$$

where $\underline{z}_t = \inf_{s\le t} z_s$. As before, we represent the agent's objective as

$$E^0\left[\Lambda^\varepsilon_{\tau\wedge H} \left\{ \int_0^{\tau\wedge H} e^{-\gamma s}(dU_s + \lambda\varepsilon_s\, ds) + e^{-\gamma(\tau\wedge H)} R \right\} \right]$$

$$= E^0\left[\Lambda^\varepsilon_{\tau\wedge H} \left\{ \int_0^{\tau\wedge H} e^{-\gamma s}(dU_s - \lambda\sigma\, dX_s) + e^{-\gamma(\tau\wedge H)} R \right\} \right]$$

So the agent's residual value process will be

$$z_t = E^\varepsilon_t\left[\int_t^{\tau\wedge H} e^{-\gamma s}(dU_s - \lambda\sigma\, dX_s) + e^{-\gamma(\tau\wedge H)} R \right]$$

and hence

$$Z_t \equiv \int_0^t e^{-\gamma s}(dU_s - \lambda\sigma\, X_s) + e^{-\gamma t} z_t \quad \text{is a } P^\varepsilon\text{-martingale}$$

under optimal control. Now

$$dz_{t} = g(z_t, \underline{z}_t)\left(dX_t^{\varepsilon} - \frac{\varepsilon_t}{\sigma}dt\right) + b(z_t, \underline{z}_t)\,dt - dL_t$$

so

$$e^{\gamma t} dZ_t \doteq -\gamma z_t\,dt + \left\{b(z_t, \underline{z}_t) - \frac{\varepsilon_t}{\sigma} g(z_t, \underline{z}_t)\right\}dt - dL_t + dU_t - \lambda\sigma\left(dX^{\varepsilon} - \frac{\varepsilon}{\sigma}dt\right)$$

so this tells us that $\boxed{dU_t = dL_t}$ and

$$0 = \sup_{0 \le \varepsilon \le \beta}\left[-\gamma z + b(z, \underline{z}) + \frac{\varepsilon}{\sigma}(\lambda\sigma - g(z, \underline{z}))\right]$$

$$= -\gamma z + b(z, \underline{z}) + \frac{\beta}{\sigma}\left(\lambda\sigma - g(z, \underline{z})\right)^{+}.$$

Therefore

$$\boxed{b(z, \underline{z}) = \gamma z - \frac{\beta}{\sigma}\left(\lambda\sigma - g(z, \underline{z})\right)^{+}}$$

relating the drift to the vol of the process z.

(8) The principal's objective is $E^{\varepsilon}\left[\int_0^{\tau \wedge H} e^{-rs}(dY_s - dU_s) + e^{-r(\tau \wedge H)}\,\ell\right]$ where for the principal the random time H is totally inaccessible with hazard rate $h(\cdot)$. If $v(z_t, \underline{z}_t)$ denotes the principal's value, we get

$$\tilde{Z}_t = e^{-rt}v(z_t, \underline{z}_t) + \int_0^t e^{-rs}(dY_s - dU_s) + \int_0^t h(\underline{z}_s)e^{-rs}(v(z_s, \underline{z}_s) - \ell)\,d\underline{z}_s$$

is a supermartingale up to H, and a martingale under optimal control. (P^{ε}, of course)

So do Itô:

$$e^{rt}d\tilde{Z}_t = -rv\,dt + v_z\left\{g\left(dX^{\varepsilon} - \frac{\varepsilon}{\sigma}\right)dt + b\,dt - dL\right\} + v_{\underline{z}}\,d\underline{z}$$
$$+ \tfrac{1}{2}v_{zz}\,g^2 + \sigma dX^{\varepsilon} + (\mu - \varepsilon)dt - dU + h(\underline{z})(v - \ell)d\underline{z}.$$

So we learn various conditions:

$$\boxed{\begin{aligned} &v_{\underline{z}}(z, \underline{z}) + h(\underline{z})\,(v(\underline{z}, \underline{z}) - \ell) = 0 \\ &v_z(z, \underline{z}) = -1 \quad \text{when } U \text{ grows} \\ &0 = \sup_g\left[-rv + v_z\left(b - \varepsilon g/\sigma\right) + \tfrac{1}{2}g^2 v_{zz} + \mu - \varepsilon\right] \\ &= \sup_g\left[-rv + bv_z + \tfrac{1}{2}g^2 v_{zz} + \mu - \beta I_{\{\lambda\sigma > g\}}(1 + g v_z/\sigma)\right] \end{aligned}}$$

$$= \sup_g \left[-rv + \tfrac{1}{2} g^2 v_{zz} + \mu + \gamma z v_z - I_{\{\lambda_0 > g\}} \beta (1 + \lambda v_z) \right]$$

(4) An alternative story which would keep everything univariate would be the story Takashi was trying out (or something similar) where the agent makes a statement that his outside option R takes value $\tilde{R}$, then the principal builds a contract which would cause the agent to tell the truth. The suggestion is that the principal's contract would involve shutting down the project at rate q in local time A_t at some point a, where a, q can be chosen. The agent's residual value process would look like

$$z_t = z_0 + \int_0^{t \wedge \tau} g(z_s)\, dX_s + \int_0^{t \wedge \tau} b(z_s)\, ds - L_{t \wedge \tau} - (a - R) I_{\{\tau_0 \le t\}}$$

where τ_0 is random time at which the principal shuts down the project. As before, we can do the agent's HJB and we get

$$b(z) = + \gamma z - \beta_\sigma (\lambda\sigma - g(z))^+$$

but with the wages process

$$\boxed{dU_t = dL_t + q (a - R)\, dA_t}$$

Principal gets value martingale

$$\tilde{Z}_t = e^{-r(t \wedge \tau)} v(z_{t \wedge \tau}) + \int_0^{t \wedge \tau} e^{-rs} (dY_s - dU_s) - I_{\{\tau_0 \le t\}} (v(z_{\tau_0 -}) - \ell) e^{-r\tau_0}$$

So we get the HJB equation

$$0 = \sup_g \left[-rv + bv' + \tfrac{1}{2} g^2 v'' + \mu - \varepsilon - \varepsilon g v'/\sigma \right]$$

with the condition $\boxed{\tfrac{1}{2}(v'(a+) - v'(a-)) = q(a - R) + q\,(v(a) - \ell)}$

How would the principal pick q, a?

(5) (27/4/11) Here's another variant of the basic setup.

The manager has type $\mu > 0$, which means that if he is working fully he generates output process

$$dY_t = \sigma\, dX_t + \mu\, dt$$

where X is BM. However, he can choose to divert effort ε_t to searching for a new job, $0 \le \varepsilon_t \le \mu$. He then receives new job offers at rate ε_t, which have values distributed according to density $\varphi(\cdot|\mu)$. Assume that if he walks out on the existing contract he has to pay a penalty $c \geq 0$. Then he will walk out on the contract if the offered value of the new job exceeds $z_t + c$. Write $f(z|\mu) = \int x \, I_{\{x \geq z\}} \varphi(x|\mu)\, dx$. We'll suppose that the principal keeps going with the project until either $z = 0$, or the manager switches to a new job. If the project finishes at $z = 0$, the manager gets residual value R, which we could take to be

$$R = \frac{\mu}{\mu+\gamma} f(0|\mu)$$

if we supposed that the manager takes the first job offered. If we let τ be the time when the manager gets a new job, H_0 the time z hits zero, then the value to the manager will be

$$E^\varepsilon\left[\int_0^{\tau \wedge H_0} e^{-\gamma s} dU_s + e^{-\gamma H_0} I_{\{H_0 < \tau\}} R + e^{-\gamma \tau} I_{\{\tau < H_0\}} Z\right]$$

where Z denotes the value of the job offered to the agent at time τ. The value then is

$$E^\varepsilon\left[\int_0^{\tau \wedge H_0} e^{-\gamma s} dU_s + e^{-\gamma H_0} I_{\{H_0 < \tau\}} R + \int_0^{\tau \wedge H_0} e^{-\gamma s} \varepsilon_s \int_{z_s + c}^{\infty} x \, \varphi(x|\mu)\, dx \, ds\right]$$

$$= E^\varepsilon\left[\int_0^{H_0} \exp(-\gamma s - A_s)\left(dU_s + \varepsilon_s f(z_s + c|\mu)\, ds\right) + e^{-\gamma H_0 - A(H_0)} R\right]$$

where $A(t) = \int_0^t \varepsilon_s \left(\int_{z_s+c}^{\infty} \varphi(x|\mu)\, dx\right) ds \equiv \int_0^t \varepsilon_s \bar{\varphi}(z_s + c|\mu)\, ds$. Now we can define

$$\tilde{\Lambda}_t^\varepsilon \equiv \Lambda_t^\varepsilon e^{-A_t}$$

which solves

$$d\tilde{\Lambda}_t^\varepsilon = -\varepsilon_t \tilde{\Lambda}_t^\varepsilon \left\{\sigma^{-1} dX_t + \bar{\varphi}(z_t + c)\, dt\right\} \equiv -\varepsilon_t \tilde{\Lambda}_t^\varepsilon \sigma^{-1} d\tilde{X}_t$$

and in terms of which the manager's objective is

$$\mathbb{E}^0\left[\int_0^{H_0} e^{-\gamma s} \tilde{\Lambda}_s^\varepsilon \left(dU_s + \varepsilon_s f(z_s + c)\, ds\right) + \tilde{\Lambda}_{H_0}^\varepsilon e^{-\gamma H_0} R\right]$$

Now let's suppose that the principal has set it up so that

$$dz_t = g(z_t)\, dX_t + b(z_t)\, dt - dL_t$$

and let's consider the manager's optimization problem. If his value function

Note: $f(z) - z\,\bar{\varphi}(z) = \int_z^\infty (x - z)\, \varphi(x|\mu)\, dx \geq 0$

is $H(z)$ [of course, $H(z)=z$] then he will have to have that

$$Y_t \equiv \tilde\Lambda^\varepsilon_t H(z_t) e^{-\gamma t} + \int_0^t \tilde\Lambda^\varepsilon_s e^{-\gamma s}(dU_s + \varepsilon_s f(z_s+c)\,ds)$$

is a $\mathbb{P}^0$-supermartingale, and a martingale under optimal control. Thus

$$dY_t \doteq \tilde\Lambda^\varepsilon_t \Big[-\gamma H\,dt + H'dz + \tfrac12 H'' d\langle z\rangle - \varepsilon\bar\varphi(z_t+c) H\,dt - \varepsilon\sigma^{-1} H' dz\,dX$$
$$+ dU + \varepsilon f(z_t+c)\,dt \Big] e^{-\gamma t}$$

$$\doteq \tilde\Lambda^\varepsilon_t e^{-\gamma t}\Big[-\gamma z\,dt + b(z)\,dt - dL - \varepsilon\bar\varphi(z+c) z\,dt - \varepsilon\sigma^{-1} g\,dt$$
$$+ dU + \varepsilon f(z+c)\,dt \Big]$$

so it must be that $\boxed{dU = dL}$ and

$$\boxed{0 = -\gamma z + b(z) + \mu\left(f(z+c) - z\bar\varphi(z+c) - \sigma^{-1} g(z)\right)^+}$$

Thus we know the drift b in terms of the volatility g:

$$\boxed{b(z) = \gamma z - \mu\left(f(z+c) - z\bar\varphi(z+c) - \sigma^{-1} g(z)\right)^+}$$

How does it look for the principal? We know what control ε the manager will select: it's

$$\varepsilon = \mu\, I_{\{ f(z+c) > z\bar\varphi(z+c) + \sigma^{-1} g(z)\}}$$

Also, if $T = H_0 \wedge \tau$, then the intensity describing the manager's departure is $\varepsilon\bar\varphi(z_t+c)$ so if v is the principal's value weight

$$Z_t \equiv e^{-r(t\wedge T)} v(z_{t\wedge T}) + \int_0^{t\wedge T} e^{-rs}(dY_s - dU_s) + I_{\{t\geq T\}} e^{-rT}(q - v(z_T))$$

(where q is the value to P of the project after manager leaves) should be a supermart and a martingale under optimal control. So

$$\boxed{\begin{gathered} 0 = \sup_g \Big[-rv + b_\varepsilon v' + \tfrac12 g^2 v'' + \varepsilon\bar\varphi(z+c)(q-v) + (\mu-\varepsilon) \Big] \\ v'(\bar z) = -1, \qquad v(\underline{z}) = q \end{gathered}}$$

$$\boxed{b_\varepsilon(z) = b(z) - \varepsilon g(z)/\sigma.}$$

Write $\lambda(z) = f(z+c) - z\bar{\varphi}(z+c)$ so that

$$b(z) = \gamma z - \frac{\mu}{\sigma}(\lambda\sigma - g)^+ , \quad \varepsilon = \mu I_{\{g < \lambda\sigma\}}$$

and the terms involving g in the principal's optimization are

$$\frac{1}{2} g^2 v'' + \left(b(z) - \frac{\varepsilon g}{\sigma}\right) v' + \varepsilon\left((q-v)\bar{\varphi}(z+c) - 1\right)$$

$$= \frac{1}{2} g^2 v'' + v'\left(\gamma z - \frac{\mu}{\sigma}(\lambda\sigma - g)^+ - \frac{\varepsilon g}{\sigma}\right) + \mu I_{\{g<\lambda\sigma\}}\left((q-v)\bar{\varphi}(z+c) - 1\right)$$

Must have $v'' \le 0$, and the range of g splits into $(-\infty, \lambda\sigma) \cup [\lambda\sigma, \infty)$.
Note that it could happen that $\lambda(z) < 0$? No - we have $f(z) - z\bar{\varphi}(z) \ge 0$.
So altogether we get HJB in the form

$$\boxed{0 = -rv + \mu + \gamma z v' + \max\left\{ \frac{1}{2}(\lambda\sigma)^2 v'', \ -\lambda\mu v' + \mu\left((q-v)\bar{\varphi} - 1\right)\right\}}$$

$$q = v(\bar{z}), \quad v'(\bar{z}) = -1.$$

Dynamic contracting: an important observation (2/5/11)

(1) If we have some diffusion-style dynamic contracting story to tell, we expect that the value function v of the principal and the value function V of the agent will depend on some state variables x and will satisfy individual HJB equations, firstly optimizing the agent's actions assuming the principal's actions are fixed, then optimizing the principal's choices assuming these optimal responses for the agent. The state variables $x = (x_0, x_1, \dots, x_N) \equiv (z, x_1, \dots, x_N)$ include the residual value z to the agent; so we must have $\boxed{V(x) = z}$ Thus the agent's HJB has a known solution, and can be used typically to extract conditions that have to be satisfied.

(2) Let's look at a simple problem where the output process of some productive asset is

$$dY_t = \sigma\, dX_t + \mu\, dt$$

where X is a standard BM, but the agent hired to run the project may divert $\varepsilon_t \geq 0$ for his own consumption. If the principal pays wages $q_t\, dt$, then for some $\lambda \in (0,1)$ we suppose that the agent's objective is

$$E\left[\int_0^\tau e^{-\gamma s}\, U(q_s + \lambda \varepsilon_s)\, ds\right]$$

and the principal's is

$$E\left[\int_0^\tau e^{-rs}\,(dY_s - q_s\, ds) + e^{-r\tau} R\right]$$

where U is a nice utility, R is residual value of the project.

Suppose the principal wants to offer a contract which ensures that the agent never diverts; what should that contract be?

Suppose that the agent's residual value process z solves

$$dz_t = g(z_t)\, dX_t + b(z_t)\, dt$$

Then we'd have

$$e^{-\gamma t} V(z_t) + \int_0^t e^{-\gamma s}\, U(q_s + \lambda \varepsilon_s)\, ds \quad \text{is supermart etc.}$$

The control ε changes dX to $dX - \varepsilon\, dt/\sigma$, so if we do Itô we get

$$0 = \sup_\varepsilon \left[-\gamma z + b(z) - \frac{\varepsilon\, g(z)}{\sigma} + U(q + \lambda \varepsilon) \right]$$

and if the contract is going to ensure that $\varepsilon = 0$ under optimal control, this says that

$$\boxed{\lambda U'(q) \le g/\sigma}$$

and the drift b satisfies

$$\boxed{0 = -\gamma z + b(z) + U(q(z)).}$$

Now let's turn to the principal's problem. Here we have

$$e^{-rt} v(z_t) + \int_0^t e^{-rs}(dY_s - q_s\, ds) \quad \text{is a supermartingale etc}$$

Therefore

$$0 = \sup\left[-rv + bv' + \tfrac{1}{2} g^2 v'' + \mu - q \right]$$

where the sup is taken over g, q subj to $\lambda U'(q) \le g/\sigma$, and the relation $b(z) = \gamma z - U(q(z))$. So we see

$$0 = \sup\left[-rv + (\gamma z - U(q))v' + \tfrac{1}{2} g^2 v'' + \mu - q \right]$$

Now the constraint to ensure no diversion is $q \ge I(g/\sigma\lambda)$, so if $v' > 0$ then we would have to have $q = I(g/\sigma\lambda)$. However, if $v' < 0$, there is the possibility of a larger wage. The FOC for q is

$$-U'(q)v' = 1 \Rightarrow q = I(-1/v')$$

$$\Rightarrow \quad q = \begin{cases} I\left(\frac{g}{\sigma\lambda} \wedge \left(\frac{-1}{v'}\right)\right) & \text{for } v' > 0 \\ I(g/\sigma\lambda) & \text{for } v' \le 0 \end{cases}$$

If we substitute this back, and now try to optimize over g, then we must have $v'' \le 0$, and in a region where $-1/v' < g/\sigma\lambda$, we could move g smaller and improve. So looks like we have always

$$q = I(g/\sigma\lambda)$$

and the principal's HJB is

$$0 = \sup_g \left[-rv + (\gamma z - (U \circ I)(g/\sigma\lambda))v' + \tfrac{1}{2} g^2 v'' + \mu - I(g/\sigma\lambda) \right]$$

Utility from possession again (2/5/11)

(i) If we relax the requirement that $c_t \geq 0$ in the story on p7 again, then the dynamics we see are

$$\begin{cases} dw_t = rw_t\,dt + \theta_t(\sigma dW_t + (\mu - r)dt) - p\cdot dC_t \\ d\xi^j_t = \lambda_j\,(dC^j_t - \xi^j_t\,dt) \end{cases}$$

where we think of C as a continuous semimartingale. The objective of the agent is to maximise $E\left[\int_0^\infty e^{-\rho t}\, U(\xi_t)\,dt\right]$, $U(\xi) = -\prod_j \xi_j^{-\gamma_j}$.

(ii) If ζ_t is the state-price density then

$$d(\zeta_t w_t) = \zeta_t\left[dw_t - w_t(\kappa dW_t + r\,dt) - \kappa\sigma\theta_t\,dt + \kappa(p\cdot dC)dW\right]$$

$$= \zeta_t\left[\theta_t\sigma\,dW_t - p\cdot dC_t + \kappa p\cdot dC_t\,dW_t\right]$$

so we deduce that

$$\zeta_t w_t + \int_0^t \zeta_s\left(p\cdot dC_s - \kappa p\; d\langle C, W\rangle_s\right) \quad \text{is a martingale}$$

so that the budget constraint is

$$w_0 = E\int_0^\infty \zeta_s\; p\cdot\left(dC_s - \kappa\, d\langle C, W\rangle_s\right).$$

(iii) If we now do the Pontryagin–Lagrange analysis of the problem, we get

$$\max E\int_0^\infty \left\{ e^{-\rho t} U(\xi_t)\,dt + \eta_t\cdot\Lambda(dC_t - \xi_t\,dt) + \xi_t\cdot d\eta_t + d\xi_t\,d\eta_t - y\,p\cdot(dC - \kappa\, dC\,dW)\zeta \right\}$$

$$-\left[\eta\cdot\xi\right]_0^\infty + y w_0 \qquad \left[d\eta = \alpha\,dW + \beta\,dt\right]$$

$$= \max E\int_0^\infty \left\{ e^{-\rho t}U(\xi_t)\,dt + dC\cdot(\Lambda\eta - y p\zeta) - \eta\cdot\Lambda\xi\,dt + \xi\cdot\beta\,dt + \alpha\cdot\Lambda\,dC\,dW \right.$$

$$\left. + \zeta y p\cdot dC\,dW.\kappa \right\} + y w_0 + \eta_0\cdot\xi_0$$

Maxing over dC leads to

$$\boxed{\Lambda\eta_t = y\zeta_t\,p}$$

so that $d\eta^j = \eta^j(-\kappa\,dW - r\,dt)$ $\Rightarrow$ $\alpha^j = -\kappa\eta^j$, $\beta^j = -r\eta^j$ and so the objective is

$$\max E \int_0^\infty \{ e^{-\rho t} U(\xi_t) dt - \xi_t(\Lambda\eta + r\eta) dt \} + y w_0 + \eta_0 \xi_0$$

So optimizing over ξ gives

$$e^{-\rho t} \frac{a_j}{\xi_j}(\Pi) = (r+\lambda_j)\eta_j = (r+\lambda_j) p_j y \zeta_t / \lambda_j$$

$$\Rightarrow \xi_t^j = a_j e^{-\rho t} \frac{\lambda_j}{(r+\lambda_j) p_j y \zeta_t} (\Pi)$$

$$\Rightarrow \Pi = \Pi^{-A} \zeta_t^A \prod (a_j \lambda_j / y(r+\lambda_j) p_j)^{-a_j}$$

$$\Rightarrow \boxed{\Pi = \zeta_t^{A/(1+A)} B} \quad \text{where } B \equiv \prod (a_j \lambda_j / y(r+\lambda_j) p_j)^{-a_j/(1+A)}$$

Thus

$$\boxed{\xi_t^j = a_j \frac{e^{-\rho t} \lambda_j}{(r+\lambda_j) p_j y} \zeta_t^{-1/(1+A)} B}$$

So the ξ^j are held in fixed proportions,

$$\boxed{\xi_t^j = \frac{\lambda_j a_j B}{(r+\lambda_j) p_j y} e^{-\rho t} \zeta_t^{-1/(1+A)}}$$

Hence $d\xi^j = \xi^j \left[\frac{\kappa\, dW}{1+A} + \left(\frac{r}{1+A} - \rho + \frac{2+A}{(1+A)^2} \frac{\kappa^2}{2} \right) dt \right]$

and so $dc^j = \lambda_j^{-1} d\xi^j + \xi^j dt$

$$= \lambda_j^{-1} \xi_t^j \left[\frac{\kappa}{1+A} dW + \left(\frac{r}{1+A} - \rho + \lambda_j + \frac{2+A}{(1+A)^2} \frac{\kappa^2}{2} \right) dt \right]$$

Hedge fund problem again (3/5/11)

(i) Let's see if we can concoct a reasonable story for this. We want the AUM to be $\varphi(\bar{w}).\,h(x)$ where $x = w/\bar{w}$ when current level is w, and current max is $\bar{w}$. Let's now restrict to special forms:

$$\boxed{\varphi(z) = z^{\alpha}, \quad h(x) = x^{\beta} \qquad \text{for some } 0 < \beta < \alpha < 1.}$$

As w falls through an excursion, we take out of the fund at bars values at and above the current w. As the excursion comes back up, we refill at the current level so as to achieve the profile we had at the start of the excursion down. Moritz highlights the problems with what happens below w_0; we can't suppose that the refilling which happens below w_0 will depend on the current value of $\bar{w}$, because this will likely not be the terminal value of $\bar{w}$, and we don't want to have to keep track of anything else.

(ii) So let's suppose that when an excursion falls from $\bar{w}$ to $\underline{w}$, the amount $q(v, \underline{w})$ removed at $v \in (\underline{w}, \bar{w})$ does not depend on $\bar{w}$. This would then give the condition

$$\int_{\underline{w}}^{\bar{w}} q(v, \underline{w})\,dv = \varphi(\bar{w}) - \varphi(\bar{w})\,h(\underline{w}/\bar{w})$$

so differentiation gives $(x \equiv \underline{w}/\bar{w})$

$$q(\bar{w}, \underline{w}) = \varphi'(\bar{w})\,(1 - h(x)) + x\,h'(x)\,\varphi(\bar{w})/\bar{w}$$

which we would of course need to be non-negative

[Check: for the assumed functional form, writing z in place of $\bar{w}$ for clarity,

$$q(z, xz) = \varphi'(z)(1-h(x)) + x h'(x)\,\varphi(z)/z$$

$$= z^{\alpha-1}\left[\alpha(1-x^{\beta}) + \beta x^{\beta}\right] > z^{\alpha-1}\beta > 0 \quad]$$

and also we would need that as $\underline{w}$ decreases, $q(v, \underline{w})$ increases for each v.

[check: $q(z, xz) = z^{\alpha-1}\left[\alpha(1-x^{\beta}) + \beta x^{\beta}\right]$

so

$$\frac{d}{dx}\, q(z, xz) = z^{\alpha-1}(\beta - \alpha)\,\beta x^{\beta-1} < 0 \quad]$$

We also require that what is removed at level v should not exceed the full profile $\varphi'(v)$

[<u>Check</u> : $q(v, xv) = v^{\alpha-1}\left[\alpha(1-x^\beta) + \beta x^\beta\right]$

$\leq \alpha v^{\alpha-1} = \varphi'(v)$.]

So this looks OK.

(iii) So what do things look like initially? We have initial level w_0, initial high-water mark $\bar{w}_0$, and initial AUM $A_0 = \varphi(\bar{w}_0)\, h(w_0/\bar{w}_0)$. Below w_0, the profile is $\varphi'(\cdot)$ so the initial profile of basis levels has nothing in $(0, w_0)$, an atom of size $\varphi(w_0)$ at w_0 (<u>note</u>: $\varphi(w_0) < A_0$). Above w_0 there is some profile which we don't need to specify for reasons that will be explained shortly. Things run until time 1, when we have $\bar{w}_1 \geq \bar{w}_0$, w_1, and AUM $A_1 = \varphi(\bar{w}_1)\, h(w_1/\bar{w}_1)$. Let's look at two cases.

<u>Case 1</u> $\underline{w}_1 < w_0 \leq w_1 \leq \bar{w}_1$. As the path falls to $\underline{w}_1$, we take out from levels $v \geq w_0$ according to the recipe given by q, but we can't (initially) take out from levels $v < w_0$, because there is nothing there. So we will remove the corresponding mass from the atom at w_0. As the excursion climbs back up to w_0, we put in at the current level according to $q(v, \underline{w}_1)$. So at time 1, the profile of basis levels below w_0 will just be

$$q(v, \underline{w}_1)\; I_{(\underline{w}_1, w_0)}(v)\; dv$$

Between w_0 and w_1 we have profile $\varphi'(v)\,dv$; at w_0, the atom has been reduced by $\int_{\underline{w}_1}^{w_0} q(v, \underline{w}_1)\,dv = \varphi(w_0)\left(1 - h(\underline{w}_1/w_0)\right)$. Above w_1, we don't know what the profile is, but it doesn't matter as none of this pays performance fees! We know how much AUM there is at basis below w_1, and how much there is in total, so we can work out the fees.

<u>Case 2</u> $\underline{w}_1 < w_1 \leq w_0 \leq \bar{w}_1$

This time, the profile below w_1 is just $q(v, \underline{w}_1)\, I_{(\underline{w}_1, w_1)}(v)\, dv$ as before, and the AUM is A_1. Only the profile below w_1 contributes to the performance fee, so this is all we need.

Utility from possession: equilibrium story (3/5/11)

(i) How explicitly could we solve the simplest equilibrium tale where there is a single productive asset generating output process $(\delta_t)_{t\geq 0}$, strictly positive and continuous? Assume there is just one commodity that this output can be used for, and there is a stock and bond, price processes S_t and B_t resp. Then we have dynamics

$$\begin{cases} dw_t = \varphi_t \, dB_t + \theta_t (dS_t + \delta_t \, dt) - c_t \, dt \\ d\xi_t = \lambda (\xi_t + c_t) \, dt \end{cases}$$

with the objective

$$\sup E \int_0^\infty e^{-\rho t} U(\xi_t) \, dt$$

where $U'(x) = x^{-R}$ as usual. Try the PL approach:

$$\max E \int_0^\infty \Big\{ e^{-\rho t} U(\xi_t) \, dt + \eta_t \lambda (\xi_t + c_t) \, dt + \xi_t \, d\eta_t + y_t \big(\varphi_t dB_t + \theta_t (dS_t + \delta_t dt) - c_t dt \big)$$
$$+ w \, dy + dw \, dy \Big\} + \xi_0 \eta_0 + y_0 w_0$$

$$= \max E \int_0^\infty \Big\{ e^{-\rho t} U(\xi_t) dt - \lambda \eta_t \xi_t + \xi_t \, d\eta_t + c_t (\lambda \eta_t - y_t) dt + \varphi \, d(yB) + \theta \, d(yS)$$
$$+ \theta y \delta \, dt \Big\} + \xi_0 \eta_0 + y_0 w_0$$

Hence $\boxed{y_t S_t + \int_0^t y_s \delta_s \, ds \text{ is a martingale, as is } y_t B_t}$ Suppose we have $d\eta_t = \eta_t \beta_t \, dt$. Then optimizing over ξ gives us

$$\boxed{e^{-\rho t} \xi_t^{-R} = \eta_t (\lambda - \beta_t)}$$

and complementary slackness for c gives

$$\boxed{\lambda \eta_t = y_t}$$

Now when markets clear, we know that $c_t = \delta_t$ and therefore

$$\xi_t = e^{-\lambda t} \Big(\xi_0 + \int_0^t e^{\lambda s} \delta_s \, ds \Big)$$

is known, and C^1. We also shall have

$$d(e^{-\lambda t} \eta_t) = e^{-\lambda t} (-\lambda \eta_t + \beta_t \eta_t) = -e^{-\rho t} \xi_t^{-R} e^{-\lambda t}$$

Suppose $\delta_t = e^{-\alpha t}\left(\delta_0 + \int_0^t e^{\alpha s}\, dZ_s\right)$ is a shot noise process, Z a subordinator.

Then

$$e^{\lambda u}\xi_u = \left(\xi_0 + \int_0^u e^{\lambda t}\delta_t\, dt\right) = \xi_0 + \int_0^u e^{(\lambda-\alpha)t}\left(\delta_0 + \int_0^t e^{\alpha s}\, dZ_s\right) dt$$

$$= \xi_0 + \delta_0 \frac{e^{(\lambda-\alpha)u} - 1}{\lambda - \alpha} + \int_0^u e^{\alpha s}\left(\int_s^u e^{(\lambda-\alpha)t}\, dt\right) dZ_s$$

$$= \xi_0 + \delta_0 \frac{e^{(\lambda-\alpha)u} - 1}{\lambda-\alpha} + \int_0^u e^{\alpha s}\, \frac{e^{(\lambda-\alpha)u} - e^{(\lambda-\alpha)s}}{\lambda-\alpha}\, dZ_s$$

so if Z has Laplace exponent ψ, and $0 < t < u$, we have

$$E\exp\left(-a\delta_t - b\xi_u\right) = \exp\left\{-a e^{-\alpha t}\delta_0 - b e^{-\lambda u}\left(\xi_0 + \delta_0 \frac{e^{(\lambda-\alpha)u} - 1}{\lambda-\alpha}\right)\right.$$

$$- \int_0^t \psi\left(e^{\alpha s - \alpha t} + \frac{e^{\alpha s - \lambda u}}{\lambda-\alpha}\left(e^{(\lambda-\alpha)u} - e^{(\lambda-\alpha)s}\right)\right) ds$$

$$\left. - \int_t^u \psi\left(\frac{e^{\alpha s - \lambda u}}{\lambda-\alpha}\left(e^{(\lambda-\alpha)u} - e^{(\lambda-\alpha)s}\right)\right) ds\right.$$

Looks quite hopeless......

which gives us

$$\eta_t = e^{\lambda t}\left[\eta_0 - \int_0^t e^{-(\lambda+\rho)s}\,\xi_s^{-R}\,ds\right]$$

The state-price density is proportional to this, which looks pretty hard to deal with! However, when we try to calculate the price of S_t, we are looking at $\int_t^\infty \eta_s \delta_s\,ds$, so in order for this to be finite, you would have to deal with the $e^{\lambda t}$ tendencies of η ....

(ii) I think the resolution of this issue is achieved by supposing that the multiplier process η is C^1 but not adapted. If this is the case, we would have $\dot{\eta} = \beta\eta$, and with $\hat{\beta}_t \equiv E(\beta_t|\mathcal{F}_t)$ the FOCs would be

$$e^{-\rho t}\xi_t^{-R} = E[\eta_t(\lambda-\beta_t)|\mathcal{F}_t] \quad , \; y_t = \lambda E[\eta_t|\mathcal{F}_t]$$

This would imply that

$$\frac{d}{dt}\left(e^{-\lambda t}\eta_t\right) = \eta_t e^{-\lambda t}(\beta_t - \lambda)$$

so assuming $e^{-\lambda t}\eta_t \to 0$ as $t \to \infty$

$$e^{-\lambda t}\eta_t = \int_t^\infty e^{-\lambda s}(\lambda-\beta_s)\,\eta_s\,ds$$

implying that

$$e^{-\lambda t}E_t\eta_t = E_t\int_t^\infty e^{-\lambda s}(\lambda-\beta_s)\,\eta_s\,ds = E_t\left[\int_t^\infty e^{-(\lambda+\rho)s}\,\xi_s^{-R}\,ds\right]$$

so that $y_t = \lambda e^{\lambda t}E_t\left[\int_t^\infty e^{-(\lambda+\rho)s}\,\xi_s^{-R}\,ds\right]$. For calculating the time-t price of the market asset, we would have to face

$$\frac{1}{y_t}E_t\left[\int_t^\infty y_s\delta_s\,ds\right] = \frac{1}{y_t}E_t\int_t^\infty \lambda e^{\lambda s}\delta_s\,E_s\left[\int_s^\infty e^{-(\lambda+\rho)u}\,\xi_u^{-R}\,du\right]ds$$

$$= \frac{1}{y_t}E_t\left[\int_t^\infty \lambda e^{\lambda s}\delta_s\int_s^\infty e^{-(\lambda+\rho)u}\,\xi_u^{-R}\,du\,ds\right]$$

<u>Can we find any process δ so that this could be calculated in closed form?</u>

Multivariate GH distributions (5/5/11)

Going back to Barndorff-Nielsen & Halgreen (ZW 38, 309–311) they show that the distribution with density

$$f(t) = \frac{(a/b)^{\lambda/2}}{2K_\lambda(\sqrt{ab})} \exp\left\{ -\frac{at}{2} - \frac{b}{2t} \right\} \qquad (t>0)$$

is infinitely divisible, and can be used to create some nice ID multivariate dist^ns, the MVGH dist^ns, by subordinating a drifting BM. In more detail, if $X_t = \sqrt{V}\, W_t + ct$ is a drifting BM in d dimensions, then the density of $X(T)$, where T has density f and is independent of X, is obtained by integrating

$$\int_0^\infty f(t) \exp\left\{ -\frac{1}{2t}(x-ct)\cdot V^{-1}(x-ct) \right\} \frac{dt}{(2\pi t)^{d/2}\sqrt{\det V}}$$

$$= \frac{(a/b)^{\lambda/2}}{2K_\lambda(\sqrt{ab})} \frac{(2\pi)^{-d/2}}{\sqrt{|V|}} \int_0^\infty t^{\lambda-\frac{d}{2}-1} \exp\left\{ -\frac{at}{2} - \frac{b}{2t} - \frac{x\cdot V^{-1}x}{2t} + x\cdot V^{-1}c - \frac{c\cdot V^{-1}c}{2}t \right\} dt$$

$$= \frac{e^{x\cdot V^{-1}c}(2\pi)^{-d/2}}{\sqrt{\det V}} \; \frac{(a/b)^{\lambda/2}}{K_\lambda(\sqrt{ab})} \cdot \frac{K_{\lambda'}(\sqrt{a'b'})}{(a'/b')^{\lambda'/2}}$$

where $a' \equiv a + c\cdot V^{-1}c$, $b' \equiv b + x\cdot V^{-1}x$, $\lambda' = \lambda - \frac{d}{2}$.

Similarly, to calculate the MGF we must calculate $(z \in \mathbb{R}^d)$

$$\int_0^\infty f(t) \exp\left((z\cdot c)t + \tfrac{1}{2}(z\cdot Vz)\, t \right) dt$$

$$= \frac{(a/b)^{\lambda/2}}{2K_\lambda(\sqrt{ab})} \int_0^\infty t^{\lambda-1} \exp\left\{ -\frac{t}{2}(a - z\cdot Vz - 2z\cdot c) - \frac{b}{2t} \right\} dt$$

$$= \frac{(a/b)^{\lambda/2}}{K_\lambda(\sqrt{ab})} \cdot \frac{K_\lambda(\sqrt{\tilde{a}b})}{(\tilde{a}/b)^{\lambda/2}}$$

where

$$\tilde{a} \equiv a - z\cdot Vz - 2z\cdot c$$

which must be positive

Preferences with limited look ahead (6/5/11)

(1) The standing assumption that we make decisions based on calculating the consequences for all future time based on certain knowledge of asset dynamics forever is utterly ridiculous! Suppose instead that at time t, what we care about is

$$E_t \left[\int_0^{t+T} U(c_s)\,ds + g(w_{t+T}) \right] \qquad (T>0 \text{ fixed})$$

(2) Now this leads to issues such as Ivar Ekeland, Ali Lazrak + others care about. But a simple special case could be when we imagine that T is exponentially distributed, when we find objective of the form

$$E \int_0^\infty e^{-\rho t} \big(U(c_t) + g(w_t) \big)\,dt.$$

This is amenable to standard HJB approach:

$$0 = \sup \left[-\rho V + U(c) + g(w) + (rw + \theta(\mu - r) - c)V' + \frac{\sigma^2\theta^2}{2} V'' \right]$$

$$= -\rho V + \tilde{U}(V') + g(w) + rwV' - \frac{\kappa^2 (V')^2}{2\sigma^2 V''}$$

We can do the dual variables trick, and obtain

$$\boxed{0 = \tfrac{1}{2}\kappa^2 z^2 J'' + (\rho - r) z J' - \rho J + \tilde{u}(z) + g(-J')}$$

This is second order non-linear, but probably the usual recursive story will get us there.

(3) A special case is where $U'(x) = x^{-R}$, $g(x) = b\,U(x)$ for some $b > 0$. This time we expect from scaling that $V(w) = A^{-R} U(w)$, and when we do the optimization

$$\begin{cases} \theta^* = \pi_M w = \dfrac{\mu - r}{\sigma^2 R} w \\ c^* = A w \end{cases}$$

where now

$$\boxed{0 = -\rho A^{-R} + (1-R) r A^{-R} + \frac{\kappa^2}{2R}(1-R) A^{-R} + b + R A^{1-R}}$$

So we see $\boxed{RA = \rho + (R-1)(r + \kappa^2/2R) - bA^R}$

If $b=0$, this is just the standard Merton solution. Otherwise (assuming the Merton problem is well posed) there is always a solution $A < \gamma_M$, and so the agent invests as with Merton, but consumes more cautiously.

(4) How would an equilibrium with such agents look?

A variation on trading to stops (7/5/11)

If we think that $X_t \equiv \sigma W_t + \mu t$ is actually going up, it might make more sense instead of fixed stops $\{-a, b\}$ to use rising stops:

$$\tilde{\tau} \equiv \inf\{t : X_t \notin (-a+qt, b+qt)\} \equiv \inf\{t : \tilde{X}_t \notin (-a, b)\}$$

where $\tilde{X}_t \equiv X_t - qt$. The objective here would be

$$E\left[e^{-\rho\tilde{\tau}}\left(1 - e^{-\gamma(X_{\tilde{\tau}} - c)}\right)\right] \Big/ \left(1 - E\, e^{-\rho\tilde{\tau}}\right)$$

$$= \frac{E\, e^{-\rho\tilde{\tau}} - E\, e^{-\gamma\tilde{X}_{\tilde{\tau}} - (\rho+\gamma q)\tilde{\tau}}\, e^{\gamma c}}{1 - E\, e^{-\rho\tilde{\tau}}}$$

$$= \frac{L_q(\rho, 0) - e^{\gamma c} L_q(\rho + \gamma q, \gamma)}{1 - L_q(\rho, 0)}$$

Where L_q is the joint laplace transform of $(\tilde{\tau}, \tilde{X}_{\tilde{\tau}})$ calculated as for $L \equiv L_0$ but replacing the drift μ by $\mu - q$.

Quite easy in principle, but the way it's currently coded doesn't perform well on this - recoding is a bit fiddly.

This has been coded up in WORK/SOLO/STOPS/TRY2, but it makes no significant difference to the numerical values obtained whether we allow $q \neq 0$ or not

A question of Peter Ruckdeschel (10/5/11)

(1) Peter is looking at inference for X given an observation Y, where with probability q the observation Y is drawn from density g and is unrelated to X, and with probability $p = 1-q$ the density of Y is $\pi(\cdot|x)$ given $X=x$. Thus if the density of X is f_0, the joint density of (X,Y) will be

$$h(x,y) = \{ q\, g(y) + p\, \pi(y|x) \} f_0(x)$$

The statistician is supposed to know p, π and f_0, but not g and his aim is to choose a kernel $k(x|y)$ which will do best in some appropriate sense. We will write

$$g_0(y) = \int \pi(y|x) f_0(x)\, dx$$

for the marginal density of Y if we have a true observation. In terms of this, the actual marginal of Y is

$$\int h(x,y)\, dx = p\, g_0(y) + q\, g(y).$$

(2) If the statistician proposes kernel k, then the joint law of (X,Y) for his model will be $\eta(x,y) = (q\, g(y) + p\, g_0(y))\, k(x|y)$ and we'd like to max the relative entropy of this over choice of g:

$$\max_g H(\tilde{P} | P) = \max_g E^P \left[\log \frac{d\tilde{P}}{dP} \cdot \frac{d\tilde{P}}{dP} \right]$$

$$= \max_g \iint \log \frac{\eta(x,y)}{h(x,y)} \cdot \eta(x,y)\, dx\, dy$$

$$= \max_g \iint \log \frac{(q\, g(y) + p\, g_0(y))\, k(x|y)}{(q\, g(y) + p\, \pi(y|x))\, f_0(x)} \cdot (q\, g(y) + p\, g_0(y))\, k(x|y)$$

This is the problem of Nature, who tries to make the statistician's choice of k as poor as possible. The statistician's problem (as we shall see) is much easier.

~~Throwing away parts of the objective not affected by g, we have the problem~~

$$\cancel{\max_g \iint \log \frac{q\, g(y) + p\, g_0(y)}{q\, g(y) + p\, \pi(y|x)} \cdot (q\, g(y) + p\, g_0(y))\, k(x|y)\, dx\, dy}$$

In fact, the statistician would always pick for k the conditional density of X given Y, and if he does so the value of H becomes zero!

(3) Seems like we need to formulate another objective. Perhaps the thing to do is to look at the <u>mutual information</u> of (X,Y), defined as the entropy of the joint law of (X,Y) relative to the product of the marginals. This is a nice index of dependence! (for example, for a bivariate Gaussian distribution, it works out as $-\frac{1}{2}\log(1-\rho^2)$). In our current situation, for the original (uncontaminated) data it would be

$$\iint f(x,y)\log\frac{f(x,y)}{f_0(x)\,g_0(y)}\,dx\,dy$$

What we could propose as Nature's problem would be to pick g so as to minimise the mutual information of (X,Y) in the contaminated model:

$$\min_g \iint f_0(x)\big(q\,g(y)+p\,\pi(y|x)\big)\log\frac{q\,g(y)+p\,\pi(y|x)}{q\,g(y)+p\,g_0(y)}\,dx\,dy$$

Should be OK to do numerically for reasonably large discrete examples?

Back to least-action filtering (18/5/11)

(1) The numerics here don't seem to cope well with more complicated examples, and I suspect this is not a little related to having to calculate $\ddot{Y}_t$ which is a place you can expect quite a bit of numerical trouble. But if we suppose some additional structure to the problem (which is actually present in all the examples I've so far investigated numerically) we can get round this problem.

(2) Let's suppose that the state/observation dynamics are

$$\begin{cases} dX_t = a(t, X_t)\, d\tilde{W}_t + b(t, X_t) & \text{in } \mathbb{R}^n \\ Y_t = C X_t + \xi_t & \text{in } \mathbb{R}^s \end{cases}$$

where ξ is an independent OU process

$$d\xi_t = A\, dW_t' - B\xi_t\, dt$$

in s dimensions. We lose no generality in assuming $A = I$, else we rescale Y, and we could for simplicity assume that W' is independent of $\tilde{W}$, but in the end the structure we have is

$$dZ_t \equiv d\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \sigma(t, X_t)\, dW + \mu(t, Z_t)\, dt$$

where $W = [\tilde{W};\, W']$ and $\mu = [b(x);\ Cb(x) - B(Y - Cx)]$. If we write $q \equiv q(t,x) \equiv (\sigma(t,x)\sigma(t,x)^T)^{-1}$ as usual, then the action functional is

$$\frac{1}{2}\int_0^T \begin{pmatrix} \dot{x} - b \\ \dot{Y} - Cb + B(Y - Cx) \end{pmatrix}^T q \begin{pmatrix} \dot{x} - b \\ \dot{Y} - Cb + B(Y - Cx) \end{pmatrix} dt$$

$$= \frac{1}{2}\int_0^T \Big\{ (\dot{x} - b)^T q_{xx} (\dot{x} - b) + 2(\dot{x} - b)^T q_{xy} (\dot{Y} - Cb + B(Y - Cx))$$
$$+ (\dot{Y} - Cb + B(Y - Cx))^T q_{yy} (\dot{Y} - Cb + B(Y - Cx)) \Big\}\, dt$$

$$\stackrel{\circ}{=} \frac{1}{2}\int_0^T \Big\{ (\dot{x} - b)^T q_{xx} (\dot{x} - b) + 2(\dot{x} - b)^T q_{xy} (BY - BCx - Cb) - 2Y^T (q_{yx}(\dot{x} - b))^{\cdot}$$
$$- 2\dot{Y}^T q_{yy} (BCx + Cb) - \cancel{2(BY)^T q_{yy} (BCx + Cb)} + \cancel{(BCx + Cb)^T q_{yy} (BCx + Cb)} \Big\}$$
$$+ \Big[(\dot{x} - b)^T q_{xy} Y \Big]_0^T$$

where "$\stackrel{\circ}{=}$" signifies equality up to some function which depends on Y only (i.e. same whatever x). Note that this requires $\tilde{W}$ independent of W' so that $q_{yy} = (AA^T)^{-1}$.

Write $\boxed{\beta \equiv BCx + Cb}$ for short, a function of t and X. To within a function only of Y, we have that the action functional is

$$\frac{1}{2}\int_0^T \Big\{ (\dot{x}-b)^T q_{xx} (\dot{x}-b) - 2Y^T (q_{yx}(\dot{x}-b))^{\circ} + 2(\dot{x}-b)^T q_{xy} (BY-\beta) + \beta^T q_{yy} \beta$$
$$+ 2Y^T (q_{yy}\beta)^{\circ} - 2\beta^T q_{yy} BY \Big\} dt + \Big[(\dot{x}-b)^T q_{xy} Y - \beta^T q_{yy} Y \Big]_0^T$$

Notice that with these structural assumptions, q_{xy} and q_{yy} are both constant, so that the expression which matters is

$$\frac{1}{2}\int_0^T \Big\{ (\dot{x}-b)^T q_{xx} (\dot{x}-b) - 2Y^T q_{yx}(\ddot{x} - \tfrac{d}{dt} b) + 2(\dot{x}-b)^T q_{xy}(BY-\beta) + \beta^T q_{yy}\beta$$
$$+ 2Y^T q_{yy}\frac{d\beta}{dt} - 2\beta^T q_{yy} BY \Big\} dt + \Big[(\dot{x}-b)^T q_{xy} Y - \beta^T q_{yy} Y \Big]_0^T$$

The hope was to reduce the degree of differentiation demanded of Y, at the price of passing to state variable $[x_t ; \dot{x}_t]$, but it doesn't look like it's going to work

(3) There may be some value in considering the least-squares estimator $\hat{X}_t$ of X_t given Y_t : by standard theory (with $A = I$)

$$\boxed{\hat{X}_t = (C^T C)^{-1} C^T Y_t \equiv M Y_t}$$

Thus if we set

$$\Delta_t \equiv X_t - \hat{X}_t = X_t - MY_t$$

we have that

$$\begin{cases} dY_t = C\,dX_t + dW'_t - B(Y_t - CX_t)dt \\ d\Delta_t = dX_t - M\,dY_t \end{cases}$$

$$= -M\,dW'_t + MB(Y_t - CX_t)\,dt$$

$$= -M\,dW'_t + MB\big(Y_t - C(\Delta_t + MY_t)\big)\,dt.$$

Looking at the action integral, what we care about is the quadratic form

$$\left\{ \begin{pmatrix} \dot{\Delta} + M\dot{Y} \\ \dot{Y} \end{pmatrix} - \begin{pmatrix} b(x) \\ \beta - BY \end{pmatrix} \right\}^T q \,\{ - \}$$

and we don't seem to get away from the problem terms of $\dot{\Delta}(\cdots)\dot{Y}$.. $\boxed{\text{Unless } a \text{ is const}}$

(4) So let's suppose we have the special case

$$\begin{cases} dX_t = dW_t + b(t, X_t)\,dt \\ Y_t = C X_t + \xi_t \end{cases}$$

where $d\xi_t = dW'_t - B\xi_t\,dt$, so that we have together

$$d\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} I & 0 \\ C & I \end{pmatrix}\begin{pmatrix} dW \\ dW' \end{pmatrix} + \begin{pmatrix} b \\ \beta - BY \end{pmatrix} dt$$

and now the idea is to define Δ by

$$\boxed{\Delta_t = X_t - M Y_t}$$

for some M to be determined, and then to optimise the action functional over choice of Δ. In the integral we get

$$\left(\begin{pmatrix} \dot\Delta + M\dot Y \\ \dot Y \end{pmatrix} - \begin{pmatrix} b \\ \beta - BY \end{pmatrix}\right)^T q \left(\begin{pmatrix} \dot\Delta + M\dot Y \\ \dot Y \end{pmatrix} - \begin{pmatrix} b \\ \beta - BY \end{pmatrix}\right)$$

$$= \left(\begin{pmatrix} \dot\Delta \\ \dot Y \end{pmatrix} - \mu\right)^T \begin{pmatrix} I & 0 \\ M^T & I \end{pmatrix} q \begin{pmatrix} I & M \\ 0 & I \end{pmatrix}\left(\begin{pmatrix} \dot\Delta \\ \dot Y \end{pmatrix} - \mu\right)$$

where $\mu \equiv \begin{pmatrix} I & -M \\ 0 & I \end{pmatrix}\begin{pmatrix} b \\ \beta - BY \end{pmatrix}$. We have $q = \begin{pmatrix} I + C^TC & -C^T \\ -C & I \end{pmatrix}$ and in order to diagonalise the quadratic form we have to take

$$\boxed{M = (I + C^TC)^{-1} C^T = C^T (I + CC^T)^{-1}}$$

and then

$$\boxed{\begin{pmatrix} I & 0 \\ M^T & I \end{pmatrix} q \begin{pmatrix} I & M \\ 0 & I \end{pmatrix} = \begin{pmatrix} I + C^TC & 0 \\ 0 & (I + CC^T)^{-1} \end{pmatrix}}$$

after some calculations. If we partition $\mu = [\,\mu^{(\Delta)}\,;\,\mu^{(Y)}\,]$ then the quadratic form is

$$(\dot\Delta - \mu^{(\Delta)})^T (I + C^TC)(\dot\Delta - \mu^{(\Delta)}) + (\dot Y - \mu^{(Y)})^T (I + CC^T)^{-1} (\dot Y - \mu^{(Y)})$$

to be minimised. This has got rid of the troublesome cross ~~terms~~, and so it has got rid of $\ddot Y$!

Different objectives for an agent (25/5/11)

(1) Suppose we have an agent who invests in a standard one-asset market

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

and has the objective to

$$\max E_t \left[\int_t^{t+T} U(s-t, c_s) \, ds + f(w_{t+T}) \right]$$

This is a rolling objective, so we could try the sort of approach of Ekeland, Lazrak, (Björk?) ... but as an alternative approach, what I'd suggest is that we propose some feedback controls $\theta_t = \theta(w_t)$, $c_t = c(w_t)$, calculate the invariant distribution, and then try to maximise the objective assuming that the wealth process has attained stationarity.

(2) For a better notation suppose we want to pick a stationary diffusion

$$dX_t = a(X_t) \, dW_t + b(X_t) \, dt$$

so that

$$\int F(x, a(x), b(x)) \, \pi(dx)$$

is maximized, where F is a given function and π is the invariant density, which as usual has the form

$$\pi(dx)/dx = \tilde{\pi}(x) \cdot \left(\frac{a(x)^2}{2} \right)^{-1}$$

and

$$\tilde{\pi}(x) = \exp \left\{ \int^x \frac{2b(y)}{a(y)^2} \, dy \right\}$$

Thus if we use a Lagrangian to absorb the normalization of the density, we end up with the problem

$$\max_{a,b} \int \left(F(x, a(x), b(x)) - \lambda \right) \frac{2}{a(x)^2} \tilde{\pi}(x) \, dx$$

For ease of working, let's write $\alpha(x) \equiv 1/a(x)$, $G(x, \alpha(x), b(x)) = F(x, a(x), b(x))$ so we want to

$$\max_{\alpha, b} \int \left\{ G(x, \alpha(x), b(x)) - \lambda \right\} \alpha(x)^2 \exp \left(\int^x 2 b(y) \alpha(y)^2 \, dy \right)$$

We can now try a variational approach to this problem: perturb α to $\alpha + \varepsilon$,

b to $b+\eta$ and we obtain FOCs:

$$\alpha^2(G-\lambda)$$

$$0 = \int \tilde{\pi}(x)\Big\{ 2\alpha\varepsilon(G-\lambda) + \varepsilon G_\alpha \alpha^2 + \alpha^2 \eta G_b + \Big[\int^x (2\eta(y)\alpha(y)^2 + 4\varepsilon(y)\alpha(y)\, b(y))\, dy \Big\} dx$$

$$= \int \varepsilon(x)\tilde{\pi}(x)\Big(2\alpha(x)(G(x)-\lambda) + \alpha(x)^2 G_\alpha + \int_x^\infty \tilde{\pi}(z)\alpha(z)^2 (G(z)-\lambda)\, dz \, \frac{4\alpha(x)b(x)}{\tilde{\pi}(x)}\Big)$$

$$+ \int \eta(x)\tilde{\pi}(x)\Big(\alpha^2 G_b + \int_x^\infty \tilde{\pi}(z)\alpha^2(G(z)-\lambda)\, dz \cdot \frac{2\alpha(x)^2}{\tilde{\pi}(x)}\Big)\, dx$$

So this gives us a pair of conditions:

$$\boxed{\begin{aligned} &\frac{\tilde{\pi}(x)\{2\alpha(G-\lambda) + \alpha^2 G_\alpha\}}{4\alpha b} + \int_x^\infty \alpha^2(z)(G(z)-\lambda)\tilde{\pi}(z)\, dz = 0 \\ &\tilde{\pi}(x) G_b + 2\int_x^\infty \tilde{\pi}(z)\alpha(z)^2 (G(z)-\lambda)\, dz = 0 \end{aligned}}$$

We can differentiate these two equations, which then reduce to a pair of coupled non-linear first-order ODEs for (α, b) which should be reasonably OK to solve numerically. However, I suspect this may end up being rather fiddly; the ODE can be expected to need two pieces of information to fix $\alpha(1)$, $b(1)$ ... and there is only one obvious one, namely, that $\tilde{\pi}$ integrates to 1. This suggests that we will need something more slippery, such as minimal positive solution, to nail the solution down..

Macroeconomic story: another try (18/6/11)

This is trying to write down some equations for how a discrete-time macroeconomic model might evolve. Proposal is to have households, firms and banks. At the start of period t, production has just happened. The output of the firms is

$$Y_t = f(K_t, L_t) Z_t$$

where K_t is capital used, L_t is labour used, and Z_t is some economy-wide random shock. There are also firm-specific shocks which we'll come back to. The state at the start of period t is

$$X_t = \begin{pmatrix} x_t \\ Z_t \end{pmatrix} = \begin{pmatrix} K_t \\ L_t \\ b_t \\ \theta_t \\ D_t \\ w_t \\ Z_t \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow \text{bank reserves of cash} \\ \leftarrow \text{households' bank deposits} \\ \leftarrow \text{firms' debt} \\ \leftarrow \text{wage rate} \\ \\ \end{array}$$

The banks are due to receive D_t from the firms, and due to pay θ_t to the households at the start of period t. The firms owe $D_t + w_t L_t$ and due to the idiosyncratic shocks to firms' output, some may be unable to pay, and will default. The firm's assets are: its liabilities:

$$\pi_t (Y_t \xi_i + K_t) \qquad\qquad D_t + w_t L_t$$

where ξ_i is shock to firm i and if liabilities exceed assets, then the firm is liquidated, recovering some fraction $(1-\mu)$ of face value (π_t is the price level at time t) Of the recovered money, debt to banks is paid first, with the rest going on the wages due. Let's write $\tilde{D}_t$, $\widetilde{w_t L_t}$ for the aggregate amount paid to the banks, workers by firms in period t. We also write $\tilde{Y}_t$ for the total amount of output once losses on default have been accounted for. The firms pay dividend d_t, the banks pay dividend δ_t so we have various equations and inequalities:

Households' budget: $\widetilde{w_t L_t} + \tilde{\theta}_t + d_t + \delta_t = \pi_t C_t + \theta'_{t+1}$

Firms' budget: $\widetilde{\pi_t (Y_t + K_t)} = \tilde{D}_t + \widetilde{w_t L_t} + d_t + \pi_t I_t - D'_{t+1}$

Banks' budget: $\tilde{D}_t - \tilde{\theta}_t + \theta'_{t+1} - D'_{t+1} = b_{t+1} - b_t + \delta_t$

output split: $\tilde{Y}_t = I_t + C_t$

where $\theta'_{t+1} \equiv \theta_{t+1}/(1+r_{t+1})$ is the time-t value of the bank deposits which are due to pay θ_{t+1} at time $t+1$, $D'_{t+1} = D_{t+1}/(1+R_{t+1})$, where R_{t+1} is the rate for lending, and I_t is the amount of investment in period t.
We also have

$$K_{t+1} = \delta \tilde{K}_t + I_t$$

$$D'_{t+1} \leq \nu \, b_{t+1}$$ (can't lend more than some multiple of reserves)

$$D'_{t+1} \leq \alpha(\theta'_{t+1} + b_{t+1})$$ (can't lend money you don't have)

$$D'_{t+1} \leq \lambda S_t = \lambda \{ \pi_t (\widetilde{K_t + Y_t}) - \tilde{D}_t - \widetilde{w_t L_t} \}$$ (leverage)

where $\alpha, \lambda \in (0,1)$ are positive constants.

For the objective, I'd suggest we try something like

$$U(C,L) + A_1 (d_t + \delta_t) + \beta E V(X_{t+1})$$

where the form of the value has to be guessed... Notice we have

$$E\, V(X_{t+1}) = E\, V(K_{t+1}, L_{t+1}, b_{t+1}, \theta'_{t+1}(1+r_{t+1}), D'_t(1+R_{t+1}), w_{t+1}, Z_t)$$

and $r_{t+1}, w_{t+1}, R_{t+1}$ appear nowhere in the constraints, nor elsewhere in the objective. Thus when we do the maximization, we'll end up with

$$v(K_{t+1}, L_{t+1}, b_{t+1}, \theta'_{t+1}, D'_{t+1})$$

for this expectation. The variables that we are optimizing over are:

$$(L_{t+1}, c_t, \pi_t, d_t, \delta_t, \theta'_{t+1}, D'_{t+1})$$

... but this hasn't taken into account the marginal information for firms + agents?

Bayesian inference on tails (3/7/11)

Speaking to Genna Samorodnitsky in Braunschweig reminds me of the stuff we did with Natalia Horbenko on Bayes inference for tails. Genna said try to estimate tail exponent for SS(1.7) (so tail density $\sim x^{-(1+\lambda)}$, $\lambda = 1.7$) with 50K data points. Here's one thing you could do:

(i) Take the data and rescale it so that some percentile $X_{(pN)}$ becomes 1. Then suppose the data is from a dist^n whose density is

$$p\, I_{(x<1)} + (1-p)\,\lambda x^{-1-\lambda}\, I_{(x>1)}$$

This allows you to calculate the posterior likelihood for λ quite easily, and it looks OK, but you are stuck with this nasty choice of p

(ii) Alternatively, you could try to let the data choose this for you. One way to do this would be to introduce a positive scaling parameter q so that the density is

$$f(x\,|\,\lambda, q) = \frac{\lambda}{\lambda+1}\, q^{-1} \left(1 \vee \frac{x}{q}\right)^{-\lambda-1}$$

Thus the joint density of $X_1, \dots, X_N$ would be

$$\left(\frac{\lambda}{q(1+\lambda)}\right)^N \left\{\prod_j \left(1 \vee \frac{x_j}{q}\right)\right\}^{-1-\lambda} \equiv f_N(\underline{x}\,|\,\lambda, q)$$

So if we suppose that we've ordered the x_j increasing, we can actually integrate out the q-variable reasonably explicitly (let's suppose the prior makes q uniform on $[0, M]$). We get $(\lambda/(1+\lambda))^N$ times

$$(*)\quad \left(\prod x_i\right)^{-1-\lambda} \frac{x_1^{N\lambda+1}}{N\lambda+1} + \sum_{j=1}^{N-1} \left(\prod_{i=j+1}^{N} x_i\right)^{-1-\lambda} \left[\frac{q^{(1+\lambda)(N-j)-N+1}}{(1+\lambda)(N-j)-N+1}\right]_{x_j}^{x_{j+1}}$$

$$+ \left[\frac{q^{-N+1}}{-N+1}\right]_{x_N}^{M}$$

if $x_N < M$, and if M is somewhere among the x_j, we just replace x_j by $x_j \wedge M$ in this expression. The middle sum reworks to

$$\sum_{j=1}^{N-1} \left(\prod_{i=j+1}^{N} \frac{x_i}{x_{j+1}}\right)^{-1-\lambda} \frac{x_{j+1}^{1-N}}{(N-j)(1+\lambda)+1-N} \left\{1 - \left(\frac{x_j}{x_{j+1}}\right)^{(N-j)(1+\lambda)+1-N}\right\}$$

and if we set $x_0 \equiv 0$, and include $j=0$ in the sum we absorb the initial term in (*) also. Now we set $z_r = x_{N+1-r}$, $r = 1, \dots, N$, which is the sequence x in decreasing order. In terms of this, we can rework (*) to

$$\sum_{r=0}^{N} \left(\prod_{s=1}^{r} \frac{z_s}{z_r} \right)^{-1-\lambda} \frac{z_r^{1-N}}{r(1+\lambda)+1-N} \left\{ 1 - \left(\frac{z_{r+1}}{z_r} \right)^{r(1+\lambda)+1-N} \right\}$$

if $z_0 \equiv x_{N+1} \equiv M$.

Numerically it turns out that $r(1+\lambda)+1-N$ is mainly negative, so then the term $(z_{r+1}/z_r)^{r(1+\lambda)+1-N}$ is huge. We therefore seem to need to do this sum as

$$\sum_{r=0}^{N} \left(\prod_{s=1}^{r} z_s/z_r \right)^{-1-\lambda} \frac{z_{r+1}^{1-N}}{r(1+\lambda)+1-N} \left\{ \left(\frac{z_{r+1}}{z_r} \right)^{N-1} - \left(\frac{z_{r+1}}{z_r} \right)^{r(1+\lambda)} \right\}$$

Care needed over the term $r=N$ because $z_{N+1} = 0$! Even then the numerics are unstable, because for large r, z_r is typically very small, and is being raised to a large negative power. This suggests a better grouping would be to split the sum according to the sign of $r(1+\lambda)+1-N$. So if $-\nu_r \equiv r(1+\lambda)+1-N < 0$, we write the summand as

$$\left(\prod_{s=1}^{r} z_s \right)^{-1-\lambda} z_{r+1}^{-\nu_r} \frac{1}{\nu_r} \left\{ 1 - \left(\frac{z_{r+1}}{z_r} \right)^{\nu_r} \right\}$$

and if $\alpha_r \equiv r(1+\lambda)+1-N > 0$, we write summand as

$$\underbrace{\left(\prod_{s=1}^{r} z_s \right)^{-1-\lambda} z_r^{\alpha_r}}_{\text{handled via the logs.}} \frac{1}{\alpha_r} \left\{ 1 - \left(\frac{z_{r+1}}{z_r} \right)^{\alpha_r} \right\}$$