Supplemental Notes

February 20, 2019

1 Deriving the HJB equation

Let

$$J(t, X_t; c, \pi) := \mathbb{E}_t \left[\int_t^T e^{-\beta(s-t)} u_1(c_s) ds + e^{-\beta(T-t)} u_2(X_T) \right].$$

Step 1 is to show that

$$e^{-\beta t} J(t, X_t; c, \pi) + \int_0^t e^{-\beta s} u_1(c_s) ds = \mathbb{E}_t \left[\int_0^T e^{-\beta s} u_1(c_s) ds + e^{-\beta T} u_2(X_T) \right]$$

is a martingale.

The key tool is the law of iterated expectations.

Theorem 1. *If $\mathcal{G} \subset \mathcal{F}$ are $\sigma-$algebras, and Y is a random variable adapted to them, then*

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{F}]|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}].$$

Using this, we can prove the following proposition.

Proposition 2. *Let Y be a r.v. and $(\mathcal{F}_t, t \geq 0)$ be a filtration on $(\Omega, \mathcal{F}, P)$. Let $Z_t := \mathbb{E}[Y|\mathcal{F}_t]$, then $(Z_t, t \geq 0)$ is a martingale w.r.t. $(\mathcal{F}_t, t \geq 0)$.*

Proof. By the definition of conditional expectation, we have $\sigma(Z_t) \subset \mathcal{F}_t$, i.e., Z_t is adapted to $\mathcal{F}_t$. Since $(\mathcal{F}_t : t \geq 0)$ is a filtration, for all $0 \leq s < t$, we have $\mathcal{F}_s \subset \mathcal{F}_t$.

In addition, $\mathbb{E}[Z_t|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[Y|\mathcal{F}_s] = Z_s$. The first and third equalities are directly from the definition of Z_t. The second equality holds because of the law of iterated expectations. □

Let $\mathcal{F}_t$ be the information set ($\sigma-$algebra) at time t, then

$$\left(Z_t := \mathbb{E}_t \left[\int_0^T e^{-\beta s} u_1(c_s) ds + e^{-\beta T} u_2(X_T) \right], t \geq 0 \right)$$

is a stochastic process adapted to the filtration $(\mathcal{F}_t : t \geq 0)$. We then conclude $e^{-\beta t} J(t, X_t; c, \pi) + \int_0^t e^{-\beta s} u_1(c_s) ds$ is a martingale by applying proposition 2.

Step 2 is to apply Itô's Lemma. Recall Itô's Lemma: If $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$,

$$df(t, X_t) = \left[\mu(t) f_x + f_t + \frac{1}{2} \sigma^2(t) f_{xx} \right] dt + \sigma(t) f_x dW_t.$$

And recall the dynamic constraint:

$$dX_t = \underbrace{[rX_t + \pi_t(\mu - r) - c_t]}_{:=\mu(t)} dt + \underbrace{\pi_t \sigma}_{:=\sigma(t)} dW_t.$$

Let $f(t, X_t) := e^{-\beta t} J(t, X_t; c, \pi) + \int_0^t e^{-\beta s} u_1(c_s) ds$. We have

$$\begin{aligned} f_x &= e^{-\beta t} J_x, \\ f_t &= -\beta e^{-\beta t} J + e^{-\beta t} J_t + e^{-\beta t} u_1(c_t), \\ f_{xx} &= e^{-\beta t} J_{xx}. \end{aligned}$$

By Itô's Lemma,

$$\begin{aligned} df(t, X_t) &= \left[\mu(t) f_x + f_t + \frac{1}{2}\sigma^2(t) f_{xx}\right] dt + \sigma(t) f_x dW_t \\ &= e^{-\beta t}\Bigg[J_t - \beta J + \frac{1}{2}\sigma^2 \pi_t^2 J_{xx} + rxJ_x \\ &\quad + \pi_t(\mu - r) J_x - c_t J_x + u_1(c_t)\Bigg] dt + e^{-\beta t} \pi_t \sigma J_x dW_t. \end{aligned}$$

We have proved that $f(t, X_t)$ is a martingale, hence

$$J_t - \beta J + \frac{1}{2}\sigma^2 \pi^2 J_{xx} + rxJ_x + \pi(\mu - r) J_x - cJ_x + u_1(c) = 0.$$

Let $V(t, x) := \max_{c,\pi} J(t, x; c, \pi)$. Then

$$\begin{aligned} \beta V &= \beta \max_{c,\pi} J \\ &= \max_{c,\pi} \left[V_t + \frac{1}{2}\sigma^2 \pi^2 V_{xx} + rxV_x + \pi(\mu - r) V_x - cV_x + u_1(c)\right]. \end{aligned}$$

Finally, we get the so-called Hamilton-Jacobi-Bellman (HJB) equation:

$$\beta V = \max_{c,\pi} \left\{ V_t + [rx + \pi(\mu - r) - c] V_x + \frac{1}{2}\sigma^2 \pi^2 V_{xx} + u_1(c) \right\}.$$

$J(T, X_T; c, \pi) = \mathbb{E}_T \left[\int_T^T e^{-\beta(s-t)} u_1(c_s) ds + e^{-\beta(T-T)} u_2(X_T)\right] = u_2(X_T) \Rightarrow V(T, x) = u_2(x)$, which is the terminal condition.

2 The dual approach

2.1 Deriving the risk-neutral SPD

The building block is the fundamental theorem of asset pricing (FTAP).

Definition 3. *A stochastic process $(\bar{Z}_t, t \geq 0)$ is a risk-neutral state price density (SPD) if (i) $\bar{Z}_0 = 1$, (ii) $\bar{Z}_t > 0, \forall t \geq 0$, and (iii) $\forall s < t, E_s[\bar{Z}_t P_t] = \bar{Z}_s P_s$, where $(P_t, t \geq 0)$ is any self-financing wealth process.*

Theorem 4. *If the market is complete and there is no arbitrage, then there exists a unique risk-neutral SPD* $(\bar{Z}_t, t \geq 0)$.

I claim:

Proposition 5. *In the environment of chapter 2, the unique SPD is*

$$\bar{Z}_t = e^{-(r+\theta^2/2)t-\theta W_t}.$$

Proof. The FTAP implies there exists a unique SPD $(\bar{Z}_t, t \geq 0)$. By the martingale representation theorem, we can assume

$$\frac{d\bar{Z}_t}{\bar{Z}_t} = \mu(t)dt + \sigma(t)dW_t. \tag{1}$$

Let B_t be the price of the risk-free bond at time t and define $X_t := \bar{Z}_t B_t$. Note that $dB_t/B_t = rdt$, then by Itô's lemma, $dX_t/X_t = [\mu(t) + r]dt + \sigma(t)dW_t$. By the definition of SPD, $(X_t, t \geq 0)$ is a local martingale. Hence, $\mu(t) = -r$. Define $Y_t := \bar{Z}_t S_t$, then use the same trick (left as an exercise), we have $\sigma(t) = -(\mu - r)/\sigma := -\theta$. With the boundary condition $\bar{Z}_0 = 1$, we can solve (1), which yields $\bar{Z}_t = e^{-(r+\theta^2/2)t-\theta W_t}$ (left as another exercise). □

2.2 Deriving the static budget constraint

Itô's lemma implies

$$d\left[\int_0^t \bar{Z}_s c_s ds + Z_t X_t\right] = \bar{Z}_t(\pi_t \sigma - \theta X_t)dW_t.$$

Hence, $\left(\int_0^t \bar{Z}_s c_s ds + \bar{Z}_t X_t, t \geq 0\right)$ is a martingale and we have

$$\int_0^t \bar{Z}_s c_s ds + \bar{Z}_t X_t = \mathbb{E}_t\left[\int_0^T \bar{Z}_s c_s ds + \bar{Z}_T X_T\right].$$

Set $t = 0$, then we get the static budget constraint

$$\mathbb{E}\left[\int_0^T \bar{Z}_s c_s ds + \bar{Z}_T X_T\right] = X_0.$$

2.3 The envelope theorem

Theorem 6. *Let f be a smooth enough function and consider the following optimization problem:*

$$\max_x f(x, p).$$

Let $x^*(p)$ be a solution to this problem and define

$$f^*(p) := f(x^*(p), p).$$

We fix $p = p_0$. If $x^*(p)$ is differentiable at p_0, then

$$\left.\frac{\partial f^*}{\partial p}\right|_{p=p_0} = \left.\frac{\partial f}{\partial p}\right|_{(x,p)=(x^*(p_0),p_0)}.$$

Proof. Left as an exercise. (Hint: Use the chain rule.) $\square$

Here we give an example of the application of the envelope theorem. Let $f(x,p) = \sqrt{p} - \frac{1}{4}(x-p)^4$. It is easy to verify

$$\begin{aligned} x^*(p) &= p, \\ f^*(p) &= \sqrt{p}. \end{aligned}$$

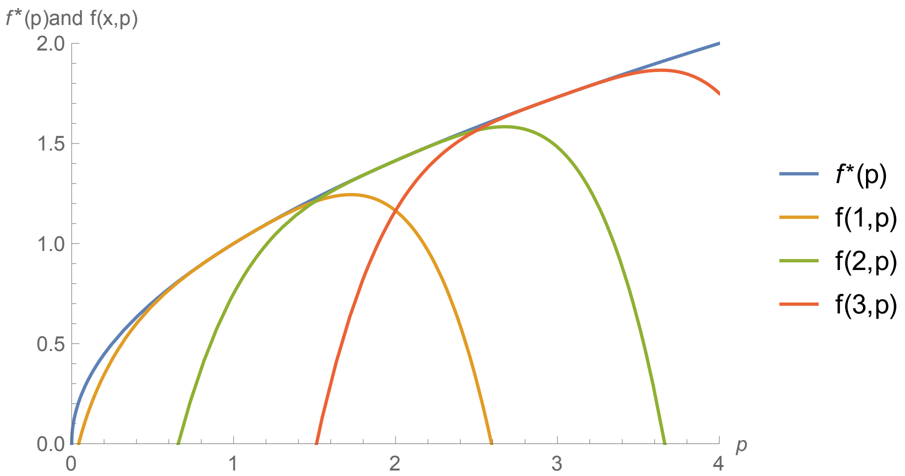
Hence,

$$\frac{\partial f^*}{\partial p} = \frac{1}{2\sqrt{p}}.$$

But we can directly apply the envelope theorem:

$$\left.\frac{\partial f}{\partial p}\right|_{(x,p)=(x^*(p),p)} = \frac{1}{2\sqrt{p}} + (x^*(p)-p)^3 = \frac{1}{2\sqrt{p}}.$$

Why is it called the **envelope** theorem? The following graph should give you some intuition.



2.4 Lagrange duality

Consider a standard optimization problem (P):

$$\max_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$$

such that

$$h(\mathbf{x}) = 0.$$

Suppose $X := \{\mathbf{x}|h(\mathbf{x}) = 0\} \neq \emptyset$. Let $\mathbf{x}^*$ be a solution to P and define $P^* := f(\mathbf{x}^*)$. The Lagrangian to P is

$$L(\mathbf{x},\lambda) := f(\mathbf{x}) + \lambda h(\mathbf{x}),$$

and we call λ the Lagrange multiplier.

Now we can define the Lagrangian dual function

$$g(\lambda) := \sup_{\mathbf{x}} L(\mathbf{x},\lambda).$$

We claim $g(\lambda) \geq P^*$, i.e., $g(\lambda)$ is a upper bound on optimal value. To see this, consider an arbitrary $\mathbf{x}' \in X$. Then

$$L(\mathbf{x}', \lambda) = f(\mathbf{x}') + \lambda h(\mathbf{x}') = f(\mathbf{x}').$$

Hence, $\forall \mathbf{x}' \in X$,

$$g(\lambda) = \sup_{\mathbf{x}} L(\mathbf{x}, \lambda) \geq L(\mathbf{x}', \lambda) = f(\mathbf{x}').$$

Therefore, $g(\lambda) \geq P^*$.

The Lagrange dual problem (D) to P is

$$\min_{\lambda} g(\lambda).$$

D gives us the best upper bound. Let $D^* := \min_\lambda g(\lambda)$. By the upper bound property above, we have

$$D^* = \min_{\lambda} g(\lambda) \geq P^*,$$

which is the so-called weak duality. If

$$D^* = P^*,$$

we say the strong duality holds, which means we can solve primal by solving dual. However, the strong duality does not always hold. But in this class we are fine...

2.5 Finding the optimal portfolio

Now we combine all the pieces together. The investor's problem at time t:

$$\max_{(c_s, \pi_s)_{s \geq t}} \mathbb{E}_t \left[\int_t^T e^{-\beta(s-t)} u_1(c_s) ds + e^{-\beta(T-t)} u_2(X_T) \right]$$

such that

$$\mathbb{E}_t \left[\int_t^T \bar{Z}_s c_s ds + \bar{Z}_T X_T \right] = \bar{Z}_t x_t. \tag{2}$$

A fact is that instead of choosing $(c_s, \pi_s)_{s \geq t}$, the investor can directly choose $(c_s)_{s \geq t}$ and X_T subject to (2) to maximize the objective function.

Set up the Lagrangian

$$\begin{aligned} \mathfrak{L}(c_s, X_T, \lambda) = \mathbb{E}_t & \left[\int_t^T e^{-\beta(s-t)} u_1(c_s) ds + e^{-\beta(T-t)} u_2(X_T) \right] \\ & + \lambda \left\{ \bar{Z}_t x_t - \mathbb{E}_t \left[\int_t^T \bar{Z}_s c_s ds + \bar{Z}_T X_T \right] \right\}. \end{aligned}$$

The first-order conditions (FOCs) are

$$\begin{aligned} \frac{\partial \mathfrak{L}}{\partial c_s} &= 0 \Rightarrow u_1'(\hat{c}_s) = \lambda e^{\beta(s-t)} \bar{Z}_s, \\ \frac{\partial \mathfrak{L}}{\partial X_T} &= 0 \Rightarrow u_2'(\hat{X}_T) = \lambda e^{\beta(T-t)} \bar{Z}_T. \end{aligned}$$

Hence,

$$\hat{c}_s = (u_1')^{-1}(\lambda e^{\beta(s-t)}\bar{Z}_s),$$
$$\hat{X}_T = (u_2')^{-1}(\lambda e^{\beta(T-t)}\bar{Z}_T).$$

Then we can use (2) to solve $\lambda, (\hat{c}_s)_{s\geq t}$ and $\hat{X}_T$. Done?! But...

The Lagrange dual function is $\max_{c_s,X_T} \mathfrak{L}(c_s, X_T, \lambda) = \mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda)$. The Lagrange dual problem is

$$\min_\lambda \mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda).$$

If the strong duality holds, then

$$\min_\lambda \mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda) = V(t, x).$$

Note that

$$\mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda) = \mathbb{E}_t\bigg\{ \int_t^T e^{-\beta(s-t)}[u_1(\hat{c}_s) - \lambda e^{\beta(s-t)}\bar{Z}_s\hat{c}_s]ds$$
$$+e^{-\beta(T-t)}[u_2(\hat{X}_T) - \lambda e^{\beta(T-t)}\bar{Z}_T\hat{X}_T]\bigg\} + \lambda\bar{Z}_tX_t.$$

We define

$$\hat{Z}_s := \lambda e^{\beta(s-t)}\bar{Z}_s,$$
$$\tilde{u}_1(z) := \max_c\{u_1(c) - zc\},$$
$$\tilde{u}_2(z) := \max_x\{u_2(c) - zx\}.$$

Then

$$\begin{aligned}\mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda) &= \mathbb{E}_t\left[\int_t^T e^{-\beta(s-t)}\tilde{u}_1(\hat{Z}_s)ds + e^{-\beta(T-t)}\tilde{u}_2(\hat{Z}_T)\right] + \lambda\bar{Z}_tX_t \\ &= \tilde{V}(t, z) + \lambda\bar{Z}_tX_t,\end{aligned}$$

where

$$\tilde{V}(t, z) := \mathbb{E}_t\left[\int_t^T e^{-\beta(s-t)}\tilde{u}_1(\hat{Z}_s)ds + e^{-\beta(T-t)}\tilde{u}_2(\hat{Z}_T)\middle|\hat{Z}_t = z\right].$$

Again $e^{-\beta t}\tilde{V}(t, z) + \int_0^t e^{-\beta s}\tilde{u}_1(\hat{Z}_s)ds$ is a martingale. Itô's lemma implies $d\hat{Z}_t = -(r-\beta)\hat{Z}_tdt - \theta\hat{Z}_tdW_t$. Then (see HW!)

$$\begin{aligned}d\left[e^{-\beta t}\tilde{V}(t, z) + \int_0^t e^{-\beta s}\tilde{u}_1(\hat{Z}_s)ds\right] =& e^{-\beta t}[\tilde{V}_t - \beta\tilde{V} + \tilde{u}_1(z) + \theta^2z^2\tilde{V}_{zz}/2 \\ &- (r-\beta)z\tilde{V}_z]dt - e^{-\beta t}\theta z\tilde{V}_zdW_t.\end{aligned}$$

The dual HJB equation is

$$\tilde{V}_t + \frac{1}{2}\theta^2z^2\tilde{V}_{zz} - (r-\beta)z\tilde{V}_z - \beta\tilde{V} + \tilde{u}_1(z) = 0,$$

with the terminal condition $\tilde{V}(T, z) = \tilde{u}_2(z)$.

Recall the strong duality, we have

$$\begin{aligned} V(t, x) &= \min_{\lambda} \mathfrak{L}(\hat{c}_s, \hat{X}_T, \lambda) \\ &= \min_{\lambda} [\tilde{V}(t, z) + \lambda \bar{Z}_t X_t] \\ &= \min_{z} [\tilde{V}(t, z) + zx]. \end{aligned}$$

The FOC of $\min_z[\tilde{V}(t, z) + zx]$ is $\tilde{V}_z + x = 0 \Rightarrow$

$$\hat{X}_t = -\tilde{V}_z(t, \hat{Z}_t).$$

By Itô's lemma,

$$d\hat{X}_t = \left[(r - \beta)\hat{Z}_t \tilde{V}_{zz} - \tilde{V}_{zt} - \frac{1}{2}\theta^2 \hat{Z}_t^2 \tilde{V}_{zzz} \right] dt + \theta \hat{Z}_t \tilde{V}_{zz} dW_t.$$

Recall the dynamic budget constraint

$$dX_t = [rX_t + \pi_t(\mu - r) - c_t]dt + \pi_t \sigma dW_t.$$

By comparing the diffusion terms, then we have

$$\hat{\pi}_t = \frac{\theta \hat{Z}_t \tilde{V}_{zz}}{\sigma}.$$

2.6 Non-Markovian Case

Proposition 7. *If $dS_t = \mu_t S_t dt + \sigma_t S_t dW_t$ and the interest rate at time t is r_t, the unique SPD is*

$$\bar{Z}_t = e^{-\int_0^t (r_s + \theta_s^2/2)ds - \int_0^t \theta_s dW_s}.$$

Proof. Left as an exercise. (Hint: The trick is almost identical.) □

What do we have now?

$$\begin{aligned} d\bar{Z}_t &= -r_t \bar{Z}_t dt - \theta_t \bar{Z}_t dW_t, \\ dX_t &= r_t X_t dt + \pi_t(\mu_t - r_t)dt + \pi_t \sigma_t dW_t - c_t dt. \end{aligned}$$

By the two-dimensional Itô's lemma,

$$\begin{aligned} d\left(\int_0^t \bar{Z}_s c_s ds + \bar{Z}_t X_t \right) &= \bar{Z}_t c_t dt + X_t d\bar{Z}_t + \bar{Z}_t dX_t + dX_t d\bar{Z}_t \\ &= \bar{Z}_t c_t dt + X_t(-r_t \bar{Z}_t)dt + X_t(-\theta_t \bar{Z}_t)dW_t \\ &\quad + \bar{Z}_t[r_t X_t + \pi_t(\mu_t - r_t) - c_t]dt + \bar{Z}_t \pi_t \sigma_t dW_t \\ &\quad + \pi_t \sigma_t(-\theta_t \bar{Z}_t)dt \\ &= \bar{Z}_t(\pi_t \sigma_t - \theta_t X_t)dW_t. \end{aligned}$$

Therefore, the LHS is a martingale. By the definition of martingale,

$$\mathbb{E}_t \left[\int_0^T \bar{Z}_s c_s ds + \bar{Z}_T X_T \right] = \int_0^t \bar{Z}_s c_s ds + \bar{Z}_t X_t.$$

The above equation implies

$$\mathbb{E}_t\left[\int_0^T \bar{Z}_s c_s ds + \bar{Z}_T X_T\right] - \int_0^t \bar{Z}_s c_s ds = \bar{Z}_t X_t,$$

which is equivalent to

$$\mathbb{E}_t\left[\int_t^T \bar{Z}_s c_s ds + \bar{Z}_T X_T\right] = \bar{Z}_t X_t.$$

To deal with the non-Markovian case, most procedures are the same as the Markovian case. But we can no longer write the value function as a function of current states. To recover the optimal consumption and the investment policy, we have to rely on the static budget constraint now.