Statistical Learning Notes

December 2019

1 Lecture 1

intro and housekeeping

2 Lecture 2

2.1 Markov Process

2.1.1 Markov Property

A given state S, is markovian iff $\mathbb{P}(S_{t+1}|S_t) = \mathbb{P}(S_{t+1}|S_1, S_2, ...S_t)$

2.1.2 State Transitions

$$\mathcal{P}_{ss'} = \mathbb{P}(S_{t+1} = s'|S_t = s)$$

$$\mathcal{P} = \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} & \dots & \mathcal{P}_{1n} \\ \vdots & \ddots & & \vdots \\ \mathcal{P}_{n1} & \mathcal{P}_{n2} & \dots & \mathcal{P}_{nn} \end{bmatrix}$$

A Markov Process (or Markov Chain) is a tuple $M = (\mathcal{S}, \mathcal{P})$

1. $\mathcal{S}$ is a finite set of states
2. $\mathcal{P}$ is a state transition probability matrix $\mathcal{P}_{ss'} = \mathbb{P}(S_{t+1} = s'|S_t = s)$.

2.2 Markov Reward Process

A Markov Reward Process is a tuple $M = (\mathcal{S}, \mathcal{P}, \mathcal{R}, \gamma)$

1. $\mathcal{S}$ is a finite set of states
2. $\mathcal{P}$ is a state transition probability matrix $\mathcal{P}_{ss'} = \mathbb{P}(S_{t+1} = s'|S_t = s)$.
3. $\mathcal{R}$ is a reward function, $\mathcal{R}_s = \mathbb{E}[R_{t+1}|S_t = s]$
4. $\gamma \in [0, 1]$ is the discount factor

2.2.1 Return

The return at a timestamp t $G_t = R_{t+1} + \gamma T_{t+2} + ... = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$
A closely related concept is the value function for a given state $v(s) = \mathbb{E}[G_t|S_t = s]$

2.2.2 Bellman Equations for MRPs

$$\begin{aligned}
v(s) &= \mathbb{E}[G_t|S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + ...|S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma(R_{t+2} + \gamma R_{t+3} + ...)|S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma G_{t+1}|S_t = s] \\
&= \mathbb{E}[R_{t+1} + \gamma v(S_{t+1})|S_t = s]
\end{aligned}$$

$$v(s) = \mathcal{R}s + \gamma \sum_{s_0 \in S} \mathcal{P}_{ss_0} v(s_0)$$

or equivalently using matrices

$$v = \mathcal{R} + \gamma \mathcal{P} v$$
$$v = (I - \gamma \mathcal{P})^{-1} \mathcal{R}$$

2.3 Markov Decision Process

A Markov Decision Process is a tuple $M = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$

1. $\mathcal{S}$ is a finite set of states
2. $\mathcal{A}$ is a finite set of actions
3. $\mathcal{P}$ is a state transition probability matrix $\mathcal{P}^a_{ss'} = \mathbb{P}(S_{t+1} = s'|S_t = s, A_t = a)$.
4. $\mathcal{R}$ is a reward function, $\mathcal{R}^a_s = \mathbb{E}[R_{t+1}|S_t = s, A_t = a]$
5. $\gamma \in [0, 1]$ is the discount factor

2.3.1 Policies

A policy π is a distribution over actions given states

$$\pi(a|s) = \mathbb{P}[A_t = a|S_t = s]$$

Note that MDP policies depend only on the current state (again history independent / memory-less).

We also observe that an MDP $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$ and a policy π together define a Markov Process $(\mathcal{S}, \mathcal{P}^\pi$, and a Markov Reward Process $(\mathcal{S}, \mathcal{P}^\pi, \mathcal{R}^\pi, \gamma$ where

$$\mathcal{P}^\pi = \sum_{a \in \mathscr{A}} \pi(a|s), \mathcal{P}^a_{ss'}$$

$$\mathcal{R}^\pi = \sum_{a \in \mathscr{A}} \pi(a|s), \mathcal{R}^a_s$$

2.3.2 Value Function

The 'state value function' $v_\pi(s)$ is the expected return starting from state s and following policy π

$$v_\pi(s) = \mathbb{E}[G_t|S_t = s]$$

The 'action value function' $q_\pi(s, a)$ is the expected return starting from state s, taking action a then following policy p.

$$q_\pi(s, a) = \mathbb{E}[G_t|S_t = s, A_t = a]$$

2.3.3 Bellman Equations for MDPs

We start by decomposing the equations for the state and action value functions

$$v_\pi(s) = \mathbb{E}_\pi[R_{t+1} + \gamma v_\pi(S_{t+1})|S_t = s]$$
$$q_\pi(s,a) = \mathbb{E}_\pi[R_{t+1} + \gamma q_\pi(S_{t+1}, A_{t+1})|S_t = s, A_t = a]$$

next we note that

$$v_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a|s) q_\pi(s,a)$$

and

$$q_\pi(s,a) = \mathcal{R}_s^a + \gamma \sum_{s'\in\mathcal{S}} \mathcal{P}_{ss'}^a v_\pi(s')$$

Substitution one in the other both ways yields

$$v_\pi(s) = \sum_{a\in\mathcal{A}} \pi(a|s)(\mathcal{R} + \gamma \sum_{s'\in\mathcal{S}} \mathcal{P}_{ss'}^a v_\pi(s'))$$
$$q_\pi(s,a) = \mathcal{R}_s^a + \gamma \sum_{s'\in\mathcal{S}} \mathcal{P}_{ss'}^a (\sum_{a'\in\mathcal{A}} \pi(a'|s') q_\pi(s',a'))$$

We can once again express these functions concisely in matrix form:

$$v_\pi = \mathcal{R}^\pi + \gamma \mathcal{P}^\pi v_\pi$$

2.3.4 Optimal things

We define the optimal state value function $v_*(s)$ as the maximum value function over all possible policies.

$$v_*(s) = \max_\pi v_\pi(s)$$

Similarly, we define the optimal action-value function $q_*(s,a)$ is the maximum action-value function over all policies

$$q_*(s,a) = \max_\pi q_\pi(s,a)$$

The optimal policy is generally what we are searching for in an MDP, we consider an MDP 'solved' when we know the optimal value function.
We define a partial ordering on all policies π s.t.

$$\pi_1 \geq \pi_2 \; if \; v_{\pi_1}(s) \geq v_{\pi_2}(s) \forall s \in S$$

Theorem: For any markov decision process

- There exists an optimal policy π_* s.t. $\pi_* \geq \pi$ for any valid policy π
- all optimal policies achieve the same value function $v_{\pi_{*1}}(s) = v_{\pi_{*2}}(s) \forall s \in S$
- all optimal policies achieve the same action-value function $q_{\pi_{*1}}(s|a) = q_{\pi_{*2}}(s|a) \forall s \in S$

Given an optimal action value function or value function, we can easily find an optimal policy:

$$\pi_*(a|s) = \begin{cases} 1 & a = \underset{a\in A}{argmax} \; q_*(s,a) \\ 0 & otherwise \end{cases}$$

Lecture 3, Planning By Dynamic Programming

2.4 Iterative Policy Evaluation

Goal: evaluate a given plicy π for a MDP $M =< \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma >$. We iterative apply the Bellman expectation backup $v_1 \rightarrow v_2 \rightarrow ... \rightarrow v_p i$.

$$v_{k+1}(s) = \sum_{a \in \mathcal{A}} \pi(a|s)(\mathcal{R}_s^a + \gamma \sum_{s \ \in \mathcal{S}} \mathcal{P}_{ss'}^a v_k(s'))$$

$$\boldsymbol{v}^{k+1} = \boldsymbol{\mathcal{R}}^\pi + \gamma \boldsymbol{\mathcal{P}}^\pi \boldsymbol{v}^v$$

The convergence of this iterative algorithm is based on a fixed point argument. Nameley that the point $\boldsymbol{v}_{pi}$ is a fixed point up this update equation, (which should be clear from the bellman equations), and moreover that this update function is contraction mapping. Therefore, convergence to $\boldsymbol{v}_\pi$ is guaranteed for *any* starting point v_0.

2.5 Iterative Policy Improvement

Now that we have a way to evaluate a given policy, we would like to be able to improve it as well. we do this in a two step process
Given a policy π

1. evaluate the policy π_k via iterative policy evaluation
   $v_1 \rightarrow v_2 \rightarrow ... \rightarrow v_p i$.

2. improve the policy by acting greedily with respect to v_π

$$\pi_{k+1}(a|s) = arg \max_{a \in \mathcal{A}} q_\pi(s, a)$$

3. iterate

Again, this procedure will always converge to the optimal policy π_*, with the proof for this again being based on fixed point / contraction mapping arguments.

2.6 Value Iteration

It turns out we can actually shortcut the policy iteration portion of our iterative improvement algorithm. This should intuitively make sense as given any value function, the optimal policy is determined using a greedy strategy.

$$v_{k+1}(s) = \max_{a \in \mathcal{A}} (\mathcal{R}_s^a + \gamma \sum_{s' \in \mathcal{S}} \mathcal{P}_{ss'}^a v_k(s'))$$

$$\boldsymbol{v}_{k+1}(s) = \max_{a \in \mathcal{A}} \boldsymbol{\mathcal{R}}^{\boldsymbol{a}} + \gamma \boldsymbol{\mathcal{P}}^{\boldsymbol{a}} \boldsymbol{v}_k$$

3 Why This Works: Contraction Mappings and Fixed Points

Define the *Bellman Expectation Backup* operator T^π as :

$$T^\pi(v) = \mathcal{R} + \gamma \mathcal{P}^\pi v$$

This function is a γ contraction i.e.

$$\begin{aligned}
||T^{\pi}(u) - T^{\pi}(u)||_{\infty} &= ||\mathcal{R} + \gamma \mathcal{P}^{\pi} u - (\mathcal{R} + \gamma \mathcal{P}^{\pi} v)||_{\infty} \\
&= ||\gamma \mathcal{P}^{\pi}(u - v)||_{\infty} \\
&\leq \gamma ||\mathcal{P}^{\pi}||_{\infty} ||(u - v)||_{\infty} \\
&\leq \gamma ||(u - v)||_{\infty}
\end{aligned}$$

Similarly, if we define

$$T^{*}(v) = \max_{a \in \mathcal{A}} \mathcal{R}^{a} + \gamma \mathcal{P}^{a} v$$

we have

$$\begin{aligned}
||T^{*}(u) - T^{*}(u)||_{\infty} &= ||\max_{a \in \mathcal{A}} (\mathcal{R}^{a} + \gamma \mathcal{P}^{a} v) - (\max_{a \in \mathcal{A}} (\mathcal{R}^{a} + \gamma \mathcal{P}^{a} v))||_{\infty} \\
&\leq \gamma ||(u - v)||_{\infty}
\end{aligned}$$

Thus convergence is shown.

4 Risk Aversion through Utility Theory

4.1 Definitions

- Utility of consumption $U(x)$
  - x represents the uncertain outcome being consumed
  - $U(\cdot)$ is a concave function, therefore $\mathbb{E}[U(x)] \leq U(\mathbb{E}[x])$
- **Certainty-Equivalent Value** $x_{CE} = U^{-1}(\mathbb{E}(U(x))$
- **Absolute Risk-Premium** $\pi_A = \mathbb{E}[x] - x_{CE}$
- **Relative Risk-Premium** $\pi_R = \frac{\pi_A}{\mathbb{E}[x]} = \frac{\mathbb{E}[x] - x_{CE}}{\mathbb{E}[x]} = 1 - \frac{x_{CE}}{\mathbb{E}[x]}$

4.2 Calculating Risk-Premium

From here on we will call $\mathbb{E}[x] = \bar{x}$ and $Var(x) = \sigma_x^2$
First we take the second order taylor expansion of $U(x)$ around $\bar{x}$:

$$U(x) \approx U(\bar{x}) + U'(\bar{x})(x - \bar{x}) + \frac{1}{2} U''(\bar{x})(x - \bar{x})^2$$

next we take the first order taylor expansion of $U(x_{CE})$ around $\bar{x}$:

$$U(x_{CE}) \approx U(\bar{x}) + U'(\bar{x})(x_{CE} - \bar{x})$$

Taking the expectation of $U(x)$ we get

$$\mathbb{E}[U(x)] \approx U(\bar{x}) + \frac{1}{2} U''(\bar{x}) \sigma_x^2$$

Noting that $\mathbb{E}[U(x)] = U(x_{CE})$ we have

$$U'(\bar{x}(x_{CE} - x) \approx \frac{1}{2} U''(\bar{x}) \cdot \sigma_x^2$$

From this, we can get expressions for the Absolute and Relative Risk Aversion

$$\pi_a = \bar{x} - x_{CE} \approx \frac{1}{2} \cdot \frac{U''(\bar{x})}{U'(\bar{x}} \cdot \sigma_x^2$$

$$\pi_r = \frac{\pi_a}{\bar{x}} \approx \frac{1}{2} \cdot \frac{U''(\bar{x}) \cdot \bar{x}}{U'(\bar{x}} \cdot \frac{\sigma_x^2}{\bar{x}^2} = 1\frac{1}{2} \cdot \frac{U''(\bar{x}) \cdot \bar{x}}{U'(\bar{x}} \cdot \sigma_{\frac{x}{\bar{x}}}^2$$

Define:

1. **Absolute Risk-Aversion** $A(\bar{x}) = -\frac{U''(\bar{x})}{U'(\bar{x})}$
2. **Relative Risk-Aversion** $R(\bar{x}) = -\frac{U''(\bar{x}) \cdot \bar{x}}{U'(\bar{x})}$

4.3 CARA and Applications

4.3.1 CARA definition

Allow $U(x) = \frac{-e^{-ax}}{a}$ for $a \neq 0$

$$A(x) = \frac{-U''(x)}{U'(x)} = a$$

a is called the coefficient of Constant Absolute Risk Aversion (CARA) if we allow the randome outcome $x \sim \mathcal{N}(mu, \sigma^2)$, then

$$\mathbb{E}[U(x)] = \begin{cases} \frac{-e^{-a\mu + \frac{a^2\sigma^2}{2}}}{a} & a \neq 0 \\ \mu & a = 0 \end{cases}$$

$$x_{CE} = \mu - \frac{a\sigma^2}{2}$$

Therefore, $\pi_a = \frac{a\sigma^2}{2}$

4.3.2 CARA applied to portfolio allocation

consider the following scenario

- We are given \$1 to invest and hold for a horizon of 1 year
- Investment choices are 1 risky asset and 1 riskless asset
  1. riskless asset annual return $\sim r$
  2. risky asset annual return $\sim \mathcal{N}(\mu, \sigma^2)$
- we want to determine the optimal (unconstrained) π to allocate to risky asset $(1 - \pi)$ is allocated to riskless asset to maximize utility of wealth in 1 year

We note that portfolio wealth $\mathcal{N}(1 + r + \pi(\mu - r), \pi^2\sigma^2$ so we optimize this, i.e. differentiate and set equal to zero yielding

$$\pi^* = \frac{\mu - r}{a\sigma^2}$$

4.4 CRRA and Applications

4.4.1 CRRA definition

Allow $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ for $\gamma \neq 1$

Relative Risk Aversion: $R(x) = \gamma$, γ is the *Coefficient of Constant Relative Risk Aversion*, note for $\gamma = 1,\ U(x) = log(x)$

If the randome outcome x is lognormal, $log(x) \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{E}[U(x)] = \begin{cases} \frac{e^{\mu(1-\gamma)+\frac{\sigma^2}{2}(1-\gamma)^2}}{1-\gamma} & gamma \neq 1 \\ \mu & \gamma = 1 \end{cases}$$

Relative Risk Premium, $\pi_R = s - \frac{x_{CE}}{\bar{x}} = 1 - e^{-\frac{\sigma^2 \gamma}{2}}$

4.4.2 CRRA application, Portfolio Construction (Merton 1969)

Problem Definition, Merton's 1969 Portfolio Problem

- 1 risky asset, 1 riskless asset
- Riskless asset $dR_y = r \cdot R_t \cdot dt$
- Risky asset $dR_y = \mu S_t \cdot dt + \sigma \cdot S_t \cdot dz_t$, (Geometric Brownian, more on this later)
- Given \$1 to invest, continuous rebalancing
- Determine π, fraction of W_t to allocate to risky asset to maximize expected utility of wealth $W = W_1$

The process for wealth with this construction is:

$$dW_t = (r + \pi(\mu - r)) \cdot W_t \cdot dt + \pi \cdot \sigma \cdot W_t \cdot dz_t$$

Solve with CRRA Utility $U(W) = \frac{W^{1-\gamma}}{1-\gamma}$, $\gamma \in (0, 1)$

Applying Ito's Lemma on $log(W_t)$ gives:

$$log(W_t) = \int_0^t (r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}) \cdot du + \int_0^t \pi \cdot \sigma \cdot dz_u$$
$$\rightarrow log(W) \sim \mathcal{N}(r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}, \pi^2\sigma^2)$$

From the previous section, we need to maximize :

$$r + \pi(\mu - r) - \frac{\pi^2\sigma^2\gamma}{2}$$

therefore

$$\pi^* = \frac{\mu - r}{\gamma\sigma^2}$$

5 Stochastic Calc Primer

Taking a step back to do some math to understand what's happening...

5.1 Intro

Let's begin by looking at a basic random process. Consider a series of coin tosses. At each toss, a heads means a win of $1, a tails means a loss of $1. Let R_i be the outcome of the ith toss.
Clearly, R_i is a random variable:

$$\mathbb{E}[R_i] = 0,\ \mathbb{E}[R_i^2] = 1,\ \mathbb{E}[R_i R_j] = 0$$

Now, let $S_i = \sum_{j=1}^{i} R_j$, S_j is an example of a random walk.

$\mathbb{E}[S_i] = 0$, $\mathbb{E}[S_i^2] = \mathbb{E}[R_1^2 + 2R_iR_j + ...] = i$ importantly, however, $\mathbb{E}[S_i|S_1...S_j] = \mathbb{E}[S_i|S_j] = S_j$.

This process exhibits both the markov property and the martingale property.

5.2 Brownian Motion Properties

- Continuity: the paths are continuous
- Markov: $X(t|X(1)....X(\tau) = X(t|X(\tau)$ for $\tau \leq t$
- Martingale: Given information up until $\tau < t$ $\mathbb{E}[X(t)] = X(\tau)$
- Sample paths have "Quadratic Variation" i.e.

$$\lim_{h \to 0} \sum_{i=m}^{n-1} (z_{(i+1)h} - z_{ih})^2 = h(n-m)$$

or equivalently

$$\int_S^T (dz_t)^2 = T - S$$

5.3 Stochastic calculus stuff

Define the *stochastic integral of f with respect to the brownian motion X* by:

$$W(t) = \int_0^t f(\tau) dX(\tau) := \lim_{n \to \infty} \sum_{j=1}^{n} f(t_{j-1}(X(t_j) - X(t_{j-w}))$$

where $t_j = \frac{jt}{n}$
This leads to *Stochastic differential euqations*

$$dW = f(t)dX$$

5.3.1 Mean Square Limit

Looking at

$$E[(\sum_{j=1}^{n}(X(t_j) - X(t_i))^2 - t)^2]$$

where $t_j = \frac{jt}{n}$
we note that $X(t_j) - X(t_i) \sim \mathcal{N}(0, t/n)$, therefore the above expectation is $O(\frac{1}{n}$, therefore we say

$$\sum_{j=1}^{n}(X(t_j) - X(y_i))^2 = t$$

in the *mean square limit*, we often write this as

$$\int_0^t (dX)^2 = t$$

5.3.2 Functions of stochastic variables

Stochastic functions do not behave as they do in normal calculus i.e. if $F = X^2$, it is not generally true that $dF = 2XdX$, for this we need Itô's Lemma
For the setup, allow $F(X)$ to be an arbitrary function, where $X(t)$ is a Brownian motion. If we consider a very small time scale $h = \frac{\delta t}{n}$ s.t. $F(X(t+h))$ can be approximated with a taylor series:

$$\begin{aligned} &F(X(t+h)) - F(X(t)) = \\ &\quad ((X(t+h) - C(t))\frac{dF}{dX}(X(t)) + \frac{1}{2}(X(t+h) - X(t))^2 \frac{d^2F}{dX^2}(X(t)) + ... = \end{aligned}$$

from this we have that

$$\begin{aligned} ((F(X(t+h)) - F(X(t))) + ((F(X(t+2h)) - F(X(t+h))) + ... \\ + ((F(X(t+nh)) - F(X(t+(n-1)h))) = \\ \sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)h))\frac{dF}{dX}(X(t+(j-i)h)) + \\ \frac{1}{2}\frac{d^2F}{dX^2}(X(t))\sum_{j=1}^{n}(X(t+jh) - X(t+(j-1)h))^2 + ... \end{aligned}$$

We note that the rhs is simply $F(X(t+nh)) - F(X(t)) = F(X(t+\delta t)) - F(X(t))$ and that the lhs is just $\int_t^{t+\delta t}\frac{dF}{dX}dX + \int_t^{t+\delta t}\frac{1}{2}\frac{d^2F}{dX^2}(X(t))\delta t$ This leaves us :

$$F(X(t+\delta t)) - F(X(t)) = \int_t^{t+\delta t}\frac{dF}{dX}(X(\tau))dX(\tau) + \int_t^{t+\delta t}\frac{1}{2}\frac{d^2F}{dX^2}(X(\tau))d\tau$$

Extending this over arbitrary time scales, we get the integral version of Itô's Lemma

$$F(X(t)) = F(X(0)) + \int_t^{t+\delta t}\frac{dF}{dX}(X(\tau))dX(\tau) + \int_t^{t+\delta t}\frac{1}{2}\frac{d^2F}{dX^2}(X(\tau))d\tau$$

the differental form of this is:

$$dF = \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dt$$

Now looking at $F = X^2$ we see that if we apply Itô's Lemma we get

$$dF = 2XdX + dt$$

5.3.3 Black-Scholes

Some baseline definitions:

A stock S is usually modelled as $dS = \mu S dt + \sigma S dX$, where μ is the drift and σ is the volatility

Allowing $F(S) - \log(S)$, we use Itô's Lemma to get

$$dF = \frac{dF}{dS}dS + \frac{1}{2}\sigma^2 S^2 \frac{d^2F}{dS^2}df = (\mu - \frac{1}{2}\sigma^2)dt + \sigma dX$$

so

$$S(t) = S(0)e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma(X(t) - X(0))}$$

So $S(t)$ is not exactly a random walk, but rather it is a lognormal random walk

We would like to value a call option that gives us the right to buy a specific asset for an agreed amount at a specified time in the future. At expiration, the value of this option is clearly $\max(S - E, 0)$, to uncover how much we should pay for it now,we consider a portfolio of a position of ΔS assets

$$\Pi = V(S, t) - \Delta S$$

We note that the change in the portfolio from t to $t + dt$ is

$$d\Pi = dV - \Delta dS$$

Ito's lemma tells us that V must satisfy

$$dV = \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}dt$$

hence, the change in value of our portfolio is

$$d\Pi = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt + (\frac{\partial V}{\partial S} - \Delta)dS$$

If we were to choose $\Delta = \frac{\partial V}{\partial S}$, we would remove all randomness from our portfolio, which is called *delta hedging*, a dynamic hedging strategy. After choosing Δ to be this, our portfolio changes by the amount

$$d\Pi = (\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt$$

This is a "riskless" change. The no arbitrage principle states then this change must be the same as the growth we would get putting an equivalent amount of cash into a risk free interest bearing asset:

$$d\Pi = r\Pi dt$$

therefore

$$(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})dt = r\Pi dt$$

Therefore, as $\Pi = V - \Delta S = V - \frac{\partial V}{\partial S}S$

$$\frac{\partial V}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

6 HJB Equation and Merton's Portfolio Problem

Problem Statement:

- You will live for T more years

- $W_0 > 0$ n risky assets, 1 riskless asset
- long and short positions allowed
- each asset as a normal distribution of treturns
- trading in continuous time, with no transaction costs
- continuous consumption of any fraction of wealth at any time

Problem Notation

- Riskless asset: $dR_t = r \cdot R_t \cdot dt$
- Risky assets: $dS_t = \mu \cdot S_) t \cdot dt + \sigma \cdot S_t \cdot dz_t$ (geometric brownian)
- $\mu > r > 0$, $\sigma > 0$ (for n assets, you need a covariance matrix)
- $W_t > 0$ wealth at time t
- $\pi(t, W_t) \in \mathbb{R}$, the fraction of wealth to be allocated to risky asset, (pi_t)
- $((1 - \pi(t, W_t)) \in \mathbb{R}$, the fraction of wealth to be allocated to riskless asset
- $c(t, W_t) \in \mathbb{R} \cap [0, W_t]$, wealth consumption per unit time, (c_t)
- utility of consumption $U(x) = \frac{x^{1-\gamma}}{1-\gamma}, \ \gamma \in (0,1)$
- utility of consumption $U(x) = log(x), \ \gamma = 1$
- γ (constant) Relative Risk-Aversion

We have the following constraint for our wealth at any timestep

$$dW_t = ((\pi_t \cdot (\mu - r) + r)W_t - c_t) + \pi_t \cdot \sigma \cdot W_t \cdot dz_t$$

At each timestep, we would like to determine the optimal $[pi_t, c_t]$ to maximize the expectation of future utility

$$E[\int_t^T \frac{e^{\rho(s-t)} \cdot c_s^{1-\gamma}}{1-\gamma} ds + \frac{e^{\rho(T-t)} \cdot B(T) \cdot W_T^{1-\gamma}}{1-\gamma} | W_t]$$

This is essentially a continuous time stochastic control problem

1. State, (t, W_T)
2. Action, $[\pi_t, c_t]$
3. Reward, $U(c_t)$
4. Return, accumulated discounted reward
5. find optimal Policy: $(t, W_t) \rightarrow [pi_t, c_t]$ to maximize expected return
6. we note that π_t is unconstrained however $c_t \geq 0$

We solve for the *Optimal Discounted Value Function*, the value function discounted to time zero

$$V^*(t, W_t) = \max_{\pi_t, c_t} \mathbb{E}[\int_t^T \frac{e^{-\rho s} \cdot c^{1-\gamma}}{1-\gamma} ds + \frac{e^{-\rho T} e^{\gamma} W_T^{1-\gamma}}{1-\gamma}]$$

We note that $V^*(t, W_t)$ satisfies the following recursive relation for all $t \leq t_1 < T$:

$$V^*(t, W_t) = \max_{\pi_t, c_t} \mathbb{E}[\int_t^{t_1} \frac{e^{-\rho s} \cdot c^{1-\gamma}}{1-\gamma} ds + V^*(t_1, W_{t_1})]$$

We can rewrite this in SDE form, and it gives us the HJB formulation:

$$\max_{\pi_t, c_t} \mathbb{E}[dV^*(t, w_t) + \frac{e^{-\rho t} \cdot c^{1-\gamma}}{1-\gamma} \cdot dt]$$

We use Ito's lemma to expand dV^*

$$dV^* = FIXME$$

removing the dz_t term as it is a matringale and we are taking the expectation, we get:

$$\max_{\pi_t, c_t}[\frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial W_t}((\pi_t(\mu - r) + r)W_t - c_t) + \frac{\partial^2 V^*}{\partial W_t^2} \frac{\pi_t^2 \sigma^2 W_t^2}{2} + \frac{e^{-\rho t} \cdot c^{1-\gamma}}{1-\gamma}] = 0$$

We can call the LHS of this $\Phi(t, w_t; \pi_t, c_t)$, so we have

$$\max_{\pi_t, c_t} \Phi(t, w_t; \pi_t, c_t) = 0$$

To find optimal π_t, c_t we can simply differentiate $\Phi(t, w_t; \pi_t, c_t)$ with respsect to π_t and c_t and sept equal to zero.

$$\begin{aligned} \frac{\partial}{\partial \pi_t} \Phi(t, w_t; \pi_t, c_t) &= (\mu - r) \cdot \frac{\partial V^*}{\partial W_t} + \frac{\partial^2 V^*}{\partial W_t^2} \cdot \pi_t \cdot \sigma^2 \cdot W_t \\ &= 0 \\ &\downarrow \\ \pi_t^* &= \frac{-\frac{\partial V^*}{\partial W_t} \cdot (\mu - r)}{\frac{\partial^2 V^*}{\partial W_t^2} \cdot \sigma^2 \cdot W_t} \end{aligned}$$

Similarly we have

$$\begin{aligned} \frac{\partial}{\partial c_t} \Phi(t, w_t; \pi_t, c_t) &= -\frac{\partial V^*}{\partial W_t} + e^{\rho t} \cdot (c_t^*)^{\gamma} = 0 \\ &\downarrow \\ c_t^* &= (\frac{\partial V^*}{\partial W_t} \cdot e^{\rho t})^{-\frac{1}{\gamma}} \end{aligned}$$

We can substitute these back into our equation for $\Phi(t, w_t; \pi_t, c_t)$ giving us

$$\frac{\partial V^*}{\partial t} - \frac{(\mu - r)^2}{2\sigma^2} \cdot \frac{(\frac{\partial V^*}{\partial W_t})^2}{\frac{\partial^2 V^*}{\partial W_t^2}} + \frac{\partial V^*}{\partial W_t} \cdot r \cdot W_t + \frac{\gamma}{1-\gamma} \cdot e^{\frac{\rho t}{\gamma}} \cdot (\frac{\partial V^*}{\partial W_t})^{\frac{\gamma - 1}{\gamma}} = 0$$

This is just a second order PDE, with a boundary condition of

$$V^*(T, W_T) = e^{-\rho T} \cdot \epsilon^{\gamma} \cdot \frac{W_T^{1-\gamma}}{1-\gamma}$$

whatever we have we consume right before we die
We guess that the solution has the form

$$V^*(t, W_t) = f(t)^\gamma \cdot e^{-\rho t} \cdot \frac{W_t^{1-\gamma}}{1-\gamma}$$

then we would have

$$\frac{\partial V^*}{\partial t} = (\gamma \cdot f(t)^{\gamma-1} f'(t) - \rho f(t)^\gamma) \cdot e^{-\rho t} \cdot \frac{W_t^{1-\gamma}}{1-\gamma}$$

$$\frac{\partial V^*}{\partial W_t} = f(t)^\gamma \cdot e^{-\rho t} \cdot W_t^{-\gamma}$$

$$\frac{\partial^2 V^*}{\partial W_t^2} = -f(t)^\gamma \cdot e^{-\rho t} \cdot \gamma \cdot W_t^{-\gamma-1}$$

We substitude this back into our PDE, yielding

$$f'(t) = (\frac{\rho - (1-\gamma) \cdot (\frac{(\mu-r)^2}{2\sigma^2\gamma} + r)}{\gamma}) f(t) - 1$$

$f(T) = \epsilon$ if we allow $v = \frac{\rho-(1-\gamma)\cdot(\frac{(\mu-r)^2}{2\sigma^2\gamma}+r)}{\gamma}$
the ODE has a solution of

$$f(t) = \begin{cases} \frac{1+(v\epsilon)\cdot e^{-v(T-t)}}{v} \text{ for } v \neq 0 \\ T - t + \epsilon \text{ for } v = 0 \end{cases}$$

this yields to solution:

$$\pi^*(t, W_t) = \frac{\mu - r}{\sigma^2 \gamma}$$

$$c^*(t, W_t) = \frac{W_t}{f(t)} = \begin{cases} \frac{v \cdot W_t}{1+(v\epsilon)\cdot e^{-v(T-t)}} \text{ for } v \neq 0 \\ \frac{W_t}{T-t+\epsilon} \text{ for } v = 0 \end{cases}$$

$$V^*(t, W_t) = \begin{cases} e^{\rho t} \cdot \frac{(1+(v\epsilon-1)\cdot e^{-v(T-t)}\gamma}{c^\gamma} \cdot \frac{W_t^{1-\gamma}}{1-\gamma} \text{ for } v \neq 0 \\ e^{\rho t} \cdot \frac{(T-t+\epsilon)^\gamma W_t^{1-\gamma}}{1-\gamma} \text{ for } v \neq 0 \end{cases}$$

References

- http://web.stanford.edu/class/cme241