Mathematics for New Technologies in Finance

Exercise sheet 6

Exercise 6.1 (Portfolio optimization) See notebook 1.

Exercise 6.2 (Optimal portfolio allocation) Consider the Merton's optmization problem outlined in lecture notebook 5. Define the set of admissible controls as $\mathcal{A} = \mathbb{H}_2$, that is to say the set of $\mathbb{F}$-adapted processes such that $\mathbb{E}\left[\int_0^T |\alpha_s|^2 \mathrm{d}s\right] < +\infty$. Fix also the parameters $\{r, \mu, \sigma\} \in \mathbb{R}^3_+$. Define, for $\gamma \in (0,1)$, the the CRRA (Constant Relative Risk Aversion) function $u = \frac{(\cdot)^\gamma - 1}{\gamma}$. The limit case $\gamma = 0$ corresponds to $u = \log$. We aim to solve the optimization problem

$$\tilde{V}(t,x) := \sup_{\alpha \in \mathcal{A}} \mathbb{E}[u(X_T^{\alpha,t,x})],$$

where the process $X^{\alpha,t,x}$ starts at time t with initial value x.

(a) Write down explicitly the dynamics of the wealth process X^α.

(b) Define

$$V(t,x) := \sup_{\alpha \in \mathcal{A}} \mathbb{E}[(X_T^{\alpha,t,x})^\gamma],$$

and notice that $\tilde{V}(t,x) = \frac{V(t,x)-1}{\gamma}$. Verify that the value function V solves the following PDE, known as dynamic programming equation:

$$\begin{aligned} &\frac{\partial w}{\partial t}(t,x) = -\sup_{a \in \mathbb{R}} \left[(r + (\mu - r)a)x\frac{\partial w}{\partial x}(t,x) + \frac{1}{2}\sigma^2 a^2 x^2 \frac{\partial^2 w}{\partial x^2}(t,x)\right] \\ &w(T,x) = x^\gamma \end{aligned} \tag{1}$$

(c) Using the ansatz $V(t,x) = x^\gamma h(t)$, reduce 1 to an ODE and solve it explicitly. Deduce the optimal Merton's ratio $\alpha^\star$, the explicit expression of V and the one of $\tilde{V}$.

Exercise 6.3 (Convergence of norms) Consider a measure space $(\Omega, \mathcal{M}, \sigma)$ and a measurable function $f : \Omega \to \mathbb{R}$, with $f \in \mathcal{L}^p(\Omega) \cap \mathcal{L}^\infty(\Omega)$ for some $0 < p < +\infty$.

(a) Prove that $\lim_{q \to +\infty} \|f\|_q = \|f\|_\infty$.

(b) If we do not assume explicitly that $f \in \mathcal{L}^p(\Omega)$, the statement may not hold anymore. Under which other assumption the statement is true, without directly assuming $f \in \mathcal{L}^p(\Omega)$?

Exercise 6.4 (Bayesian optimization)

(a) Recall the definition of prior, likelihood, posterior, and evidence distributions in Bayesian statistics.

(b) Consider linear model on $\mathbb{R} : Y \sim \theta X + Z, \theta \sim \mathcal{N}(0,1), Z \sim \mathcal{N}(0,1)$ and θ independent with X. Compute $p_\theta(y \mid x)$ and $p(\theta \mid x, y)$. Prove that maximizing the posterior $p(\theta \mid x, y)$ is exactly doing Ridge regression (fix λ here).

(c) Consider Lasso regression, what is the prior under Bayesian perspective? Please calculate the posterior under this prior.

(d) Would you expect a sparser weight or denser weight using Lasso regression instead of Ridge regression.

Exercise 6.5 (Stochastic gradient descent)

(a) Assume that we aim to find the θ^* to maximize the posterior:

$$p(\theta|x_1, \cdots, x_n) = \frac{p(\theta) \prod_{i=1}^n p(x_i|\theta)}{p(x_1, \cdots, x_n)}$$

with stochastic gradient descent method in practice. In each step, do we calculate $\nabla p(\theta|x_1, \cdots, x_n)$? do we calculate $\nabla \log p(\theta|x_1, \cdots, x_n)$? do we calculate $\nabla \log p(\theta)$ or $\nabla \log p(x_i|\theta)$?

(b) If $p(x_1, \cdots, x_n)$ has no closed formula, does it cause a trouble when we do stochastic gradient descent?

(c) Construct a stochastic differential equation with invariant measure to be the posterior distribution $p(\theta|x_1, \cdots, x_n)$.