

Stanford CME 241 (Winter 2023) - Assignment 4

Due: 2/6 @ 11:59pm. Solve 1, 2 and choose between either of 3 or 4.

1. Implement *Approximate Policy Iteration*, generalization of the tabular Policy Iteration we covered in the previous class. In order to implement Approximate Policy Iteration, first review the interface and implementation of *Approximate Policy Evaluation* and *Approximate Value Iteration* (in file [rl/approximate_dynamic_programming.py](#)), then design the interface of *Approximate Policy Iteration* to be the same as that of *Approximate Value Iteration*. Note that your implementation of *Approximate Policy Iteration* would need to invoke *Approximate Policy Evaluation* since Policy Evaluation is a component of Policy Iteration. Test that your implementation is correct in two ways:
 - Ensure that *Approximate Policy Iteration* gives the same Optimal Value Function/Optimal Policy as that obtained by *Approximate Value Iteration*.
 - Ensure that *Approximate Policy Iteration* produces the same result as our prior implementation of Policy Iteration (in file [rl/dynamic_programming.py](#)). For this you need to pass to your implementation of *Approximate Policy Iteration* a FiniteMarkovDecisionProcess input and a Tabular instance for the FunctionApprox input.
2. Assume the Utility function is $U(x) = x - \frac{\alpha x^2}{2}$. Assuming $x \sim \mathcal{N}(\mu, \sigma^2)$, calculate:
 - Expected Utility $\mathbb{E}[U(x)]$
 - Certainty-Equivalent Value x_{CE}
 - Absolute Risk-Premium π_A

Assume you have a million dollars to invest for a year and you are allowed to invest z dollars in a risky asset whose annual return on investment is $\mathcal{N}(\mu, \sigma^2)$ and the remaining (a million minus z dollars) would need to be invested in a riskless asset with fixed annual return on investment of r . You are not allowed to adjust the quantities invested in the risky and riskless assets after your initial investment decision at time $t = 0$ (static asset allocation problem). If your risk-aversion is based on this Utility function, how much would you invest in the risky asset? In other words, what is the optimal value for z , given your level of risk-aversion (determined by a fixed value of α)?

Plot how the optimal value of z varies with α .

3. Assume you are playing a casino game where at every turn, if you bet a quantity x , you will be returned $x \cdot (1 + \alpha)$ with probability p and returned $x \cdot (1 - \beta)$ with probability $q = 1 - p$ for $\alpha, \beta \in \mathbb{R}^+$ (i.e., the return on bet is α with probability p and $-\beta$ with probability $q = 1 - p$). The problem is to identify a betting strategy that will maximize one's expected wealth over the long run. The optimal solution to this problem is known as the Kelly criterion, which involves betting a constant fraction of one's wealth at each turn (let us denote this optimal fraction as f^*).

It is known that the Kelly criterion (formula for f^*) is equivalent to maximizing the Expected Utility of Wealth after a single bet, with the Utility function defined as: $U(W) = \log(W)$. Denote your wealth before placing the single bet as W_0 . Let f be the fraction (to be solved for) of W_0 that you will bet. Therefore, your bet is $f \cdot W_0$.

- Write down the two outcomes for wealth W at the end of your single bet of $f \cdot W_0$.

- Write down the two outcomes for $\log(\text{Utility})$ of W .
 - Write down $\mathbb{E}[\log(W)]$.
 - Take the derivative of $\mathbb{E}[\log(W)]$ with respect to f .
 - Set this derivative to 0 to solve for f^* . Verify that this is indeed a maxima by evaluating the second derivative at f^* . This formula for f^* is known as the Kelly Criterion.
 - Convince yourself that this formula for f^* makes intuitive sense (in terms of its dependency on α , β and p).
4. Derive the solution to Merton's Portfolio problem for the case of the $\log(\cdot)$ Utility function. Note that the derivation in the textbook is for CRRA Utility function with $\gamma \neq 1$ and the case of the $\log(\cdot)$ Utility function was left as an exercise to the reader.