Risk-sensitive LQR problems with exponential noise

Olzhas Shortanbaiuly
Supervisor: Kerem Ugurlu

Nazarbayev University

olzhas.shortanbaiuly@nu.edu.kz

26 April, 2024

Overview

Markov Decision Process

Definition

Definition 1. A Markov decision process is a 4-tuple (X, A, P_a, R_a) such that:

- State space X - set of all possible states
- Action space A - set of all possible actions
- $P_a(x_1, x_2) = P(x_{t+1} = x_2 | x_t = x_1, a_t = a)$ is a probability that using action a, we move from state x_t at time t to the state x_{t+1} at time $t + 1$.
- $R_a(x_t, x_{t+1})$ is an immediately received reward by moving from state x_t to x_{t+1} by performing an action a

Lemma 1. A decision process (X, A, P_a, R_a) is Markovian if and only if

$$P_a(x_{t+1}|x_t) = P_a(x_{t+1}|x_t, x_{t-1}, x_{t-2}, ..., x_0)$$

Dynamic Programming formulation

- set of times $t = 0, 1, 2, \ldots T$
- the states of the dynamic program $x_t \in X$ with the initial state x_0
- policy $\pi(s, x) = a_s$ for actions $a_s \in A$ and $s = t, ..., T$
  $\pi = (\pi_t : t = 0, 1, 2, \ldots T)$ with controls(actions) π_t at each time $t = 0, 1, 2, \ldots T$
- costs for taking action a_t at state x_t given by $c_t(a_t, x_t)$
- the sequence of states defined as: $x_{t+1} = f(x_t, \pi_t)$ and transition given by Markov Decision Process [3]

With this given information, we solve the following optimization problem through iterating backward in time:

$$\begin{aligned} \underset{\pi}{\text{minimize}} \quad & C(x_0, \pi) := \sum_{t=0}^{T-1} c(x_t, \pi_t) + c_T(x_T) \\ \text{subject to} \quad & x_{t+1} = f(x_t, \pi_t),\ \pi_t \in A \qquad\qquad t = 0, 1, 2, \ldots, T. \end{aligned}$$

LQR problem

For the Linear Quadratic Regulator (LQR) Problem, the following conditions are given:

LQR problem formulation

- Transition (linear) function: $x_{t+1} = x_t + a_t + \xi_t$ for the **random noise** ξ_t
- $\xi_t \sim Bernoulli(p)$ or $\xi_t \sim Exponential(\lambda)$
- Running (quadratic) cost: $c_t(x, a, t) = x_t^2 + a_t^2$
- Total cost: $\sum_{s=t}^{s=T} c(x_s, a_s)$, where $\pi(s, x) = a_s$, for $s = t, ..., T$.

Average-Value-at-Risk

Definition

Definition 2. For real-valued random variable $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ defined on measurable space $(\Omega, \mathcal{A}, \mathbb{P})$ with finite mean (stating integrability with $E|X| < \infty$), at given risk level $\alpha \in (0, 1)$ the following can be defined [3]:

- Value-at-Risk at risk level α:

$$VaR_\alpha(X) = inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq \alpha\}$$

- Average-Value-at-Risk at risk level α:

$$AVaR_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_t(X) dt$$

$AVaR_\alpha(X)$ is an average of Value-at-Risk's which are larger than the Value-at-Risk at risk level α. $AVaR_\alpha(X)$ gives the value for the losses greater than the given $VaR_\alpha(X)$ level. Compute it by averaging (integral).

Average-Value-at-Risk

Definition (Bauerle, 2011)

Lemma 2. For real-valued random variable $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ defined on measurable space $(\Omega, \mathcal{A}, \mathbb{P})$ with finite mean (stating integrability with $E|X| < \infty$), at given risk level $\alpha \in (0,1)$, the Average-Value-at-Risk can be defined as [1]:

$$AVaR_\alpha(X) = \min_{s \in \mathbb{R}} \{s + \frac{1}{1-\alpha} E[(X-s)^+]\}$$

with the minimum point $s^* = VaR_\alpha(X)$.

Average-Value-at-Risk

Coherent risk measure

Definition 3. For real-valued random variable $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$ defined on measurable space $(\Omega, \mathcal{A}, \mathbb{P})$, a **coherent risk measure** is defined to be a mapping $\rho : L^1 \to \mathbb{R}$ such that the following axioms hold [3]:

- Convexity:
  $\rho(\gamma X + (1-\gamma)Y) \leq \gamma\rho(X) + (1-\gamma)\rho(Y) \quad \forall\gamma \in (0,1), X, Y \in L^1$
- Monotonicity: if $X \leq Y$ $\mathbb{P}$-a.s. then $\rho(X) \leq \rho(Y) \quad \forall X, Y \in L^1$
- Translational invariance: $\rho(c + X) = c + \rho(X) \quad \forall c \in \mathbb{R}, X \in L^1$
- Homogeneity: $\rho(\beta X) = \beta\rho(X) \quad \forall X \in L^1, \beta \geq 0$

Average-Value-at-Risk is a coherent risk measure.

Remark on end behavior

Remark 1. For the Average-Value-at-Risk $AVaR_\alpha(X)$ defined on real-valued random variable $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, the following end behavior holds [3]:

- $\lim\limits_{\alpha \to 0} \mathrm{AVaR}_\alpha(X) = \mathbb{E}[X]$
- $\lim\limits_{\alpha \to 1} \mathrm{AVaR}_\alpha(X) = \text{ess sup} X \leq \infty$

Probability distributions

Bernoulli random variable

- Bernoulli distribution is a special case of Binomial distribution corresponding to a single trial. It can be thought of as a result of a "yes-no" experiment with two outcomes. For $X \sim Bernoulli(p)$, $P(X = 1) = p$ and $P(X = -1) = q = 1 - p$.
- The probability mass function would be the following:

$$f(n,p) = \begin{cases} p & \text{if } k = 1, \\ q = 1 - p & \text{if } k = -1 \end{cases}$$

- $E[X] = 2p - 1$ and $E[X^2] = 1^2 \cdot p + (-1)^2 \cdot (1 - p) = 1$.

Probability distributions

Exponential random variable

- The exponential distribution is a probability distribution corresponding to a Poisson point process and this process' event distance. The Poisson point process can be understood as a process with an average constant rate, where the events are independent and continuous.
- Rate of events λ can only take positive values so $\lambda > 0$ and for $X \sim \mathit{Exponential}(\lambda)$, $X \geq 0$ respectively.
  The probability mass function would be the following:
  $$f(x, \lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0, \\ 0 \quad \text{if } x < 0 \end{cases}$$
- $E[X] = \frac{1}{\lambda}$ and $E[X^2] = \frac{1}{\lambda^2}$.

Quantile computation

The computation of the quantile s^* is done using the *numpy.quantile* function. The linear interpolation method is used.

Quantile formula

For finding the new virtual index $i+g$ of an element $x \in [i, i+1]$, the following the quantile $0 \leq q \leq 1$ with a sorted list of n elements is given as:

$$i + g = q \cdot (n - \alpha - \beta + 1) + \alpha$$

For the linear interpolation, $\alpha = 1$ and $\beta = 1$. So, the formula becomes

$$i + g = q \cdot (n - 1) + 1$$

Example: a list $[1, 2, 3, 4]$ and $q = 0.25$.
$i + g = 0.25(4 - 1) + 1 = 1.75$
The 1.75th element lies between 1st element of 1 and 2nd element of 2.
Finding the value of the 1.75th element:
$1 + (2 - 1) * 0.75 = 1.75.$

Hamiltonian-Jacobi-Bellman equation

Definition

Definition 4. The optimal value function is given through the following HJB equation [1] for (t, x) where $0 \leq t \leq T$:

$$
\begin{aligned}
Q^\pi(t, x_t) &\triangleq c_t(x_t, a_t, t) + Q(t+1, x_{t+1}) \\
Q^\pi(T, x) &\triangleq g(x) \\
V(t, x) &= \inf_\pi Q^\pi(t, x)
\end{aligned}
$$

Through the iterations, each value function is obtained. Using the value function, the optimization problem will be as follows:

$$
\min_a AVaR_\alpha(Q^\pi(t, x) | \mathcal{F}_T)
$$

for some history or information given by σ-algebra $\mathcal{F}_T$ and terminal cost $g(t)$ at terminal time T.

Approximate dynamic programming algorithm - motivation

- **the state space** X for the problem may be too large (difficult to evaluate the value function $V_t\left(x_t\right)$ for all states within a reasonable time);
- **the decision space** A may be too large (difficult to find the optimal decision for all states within a reasonable time);
- computing the expectation of 'future' costs may be intractable when **the outcome space** (set of all possible states in time period $t+1$, given the state and decision in time period t) is large;

Literature review

Most existing literature focuses on the LQR problem in a matrix formulation, where the cost and transition functions are given in terms of matrix equations. However, they are given in a continuous time setting. Other than that, most literature focuses on solving LQR problems using neural networks or Q-learning. Initially, the thesis attempted to follow these ideas, however, it was concluded that approximate dynamic programming is a better approach.
Bauerle [1] and Ugurlu's [3] papers were most important for theoretical background, while Mes's work on approximate dynamic programming gave a motivation for the algorithm.

Solved problems

- LQR Problem with Bernoulli Noise
- LQR Problem with Theoretical Exponential Noise at risk level $\alpha = 0$
- LQR Problem with Sampled Exponential Noise at risk level $\alpha = 0$
- LQR Problem with Sampled Exponential Noise at risk level $\alpha = 0.25$
- LQR Problem with Sampled Exponential Noise at risk levela $\alpha = 0.5, 0.75, 0.99$

LQR problem with Bernoulli noise

Baseline setting

- Transition (linear) function: $x_{t+1} = x_t + a_t + \xi_t$ for the **random noise** ξ_t
- Running (quadratic) cost: $c_t(x, a, t) = x_t^2 + a_t^2$
- $\xi_t \sim Bernoulli(p = 1/2)$
- $\alpha \in \{0, 0.25, 0.5, 0.75, 0.99\}$
- $a_t \in \{-1, -0.5, 0, 0.5, 1\}$
- $x_0 = 1, t_0 = 0$
- $T = 2.$

LQR problem with Bernoulli noise

We start the computations at time $t = 2$, then go backwards in time in 1 time unit steps.

Step 1. t = 2

$J(2, x_2) = \inf_{a_2} E[x_2^2 + a_2^2 | x_2, a_2] = x_2^2.$

Therefore, the minimizing action $a_2 = 0$.

Step 2. t = 1

$J(1, x_1) =$

$2x_1^2 + 2x_1 E[\xi_1 | x_1, a_1] + E[\xi_1^2 | x_1, a_1] + \inf_{a_1} \{2a_1^2 + 2x_1 a_1 + 2a_1 E[\xi_1 | x_1, a_1]\}$

Computing expected values:

$E[\xi_1 | x_1, a_1] = 1(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0$

$E[\xi_1^2 | x_1, a_1] = 1^2(\frac{1}{2}) + (-1)^2(\frac{1}{2}) = \frac{1}{2} + \frac{1}{2} = 1.$

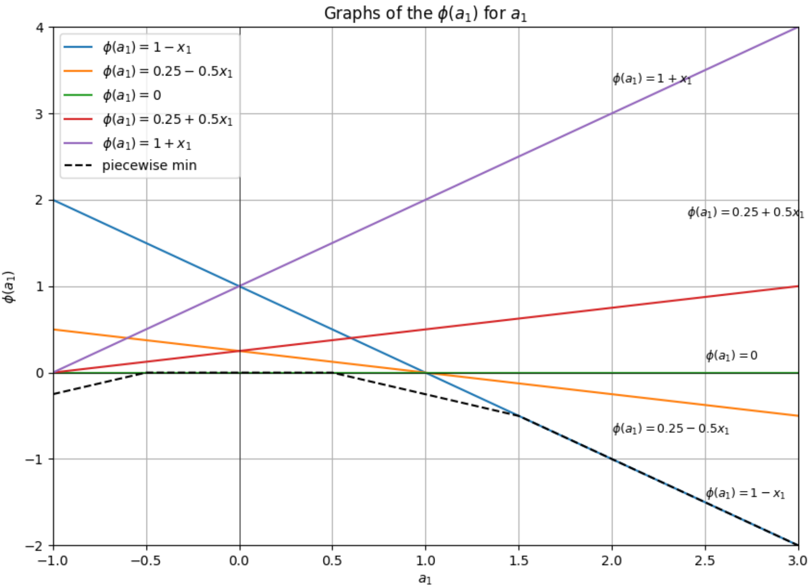
Then,

$J(1, x_1) = 2x_1^2 + 1 + 2\inf_{a_1} \{a_1^2 + x_1 a_1\}.$

LQR problem with Bernoulli noise

We have that $\phi(a_1) = a_1^2 + x_1 a_1$
$a_1 = -1 : \phi(-1) = 1 - x_1$
$a_1 = -0.5 : \phi(-0.5) = 0.25 - 0.5x_1$
$a_1 = 0 : \phi(0) = 0$
$a_1 = 0.5 : \phi(0.5) = 0.25 + 0.5x_1$
$a_1 = 1 : \phi(1) = 1 + x_1$

LQR problem with Bernoulli noise



LQR problem with Bernoulli noise

So, the piecewise function giving the minimum value at each interval is

$$\phi(a_1) = \begin{cases} 0.25 + 0.5x_1, & x_1 \in [-1, -0.5) \\ 0, & x_1 \in [-0.5, 0.5) \\ 0.25 - 0.5x_1, & x_1 \in [0.5, 1.5) \\ 1 - x_1 & x_1 \in [1.5, 3] \end{cases}$$

Step 3. $\mathbf{t = 0}$

We obtain the equations for $J(1, x_1)$ for some given x_1 in certain interval.
Compute $J(0, x_0, a_0)$ for $a_0 \in \{-1, -0.5, 0, 0.5, 1\}$.

$$J(0, x_0) = \inf_{a_0 \in \{-1,0,1\}} \{J(0, x_0, a_0 = -1), J(0, x_0, a_0 = -0.5), \\ J(0, x_0, a_0 = 0), J(0, x_0, a_0 = 0.5), J(0, x_0, a_0 = 1)\}$$

$$J(0, x_0) = 4.25$$

LQR problem with Theoretical Exponential noise

Problem setting

- Transition (linear) function: $x_{t+1} = x_t + a_t + \xi_t$ for the **random noise** ξ_t
- Running (quadratic) cost: $c_t(x, a, t) = x_t^2 + a_t^2$
- $\xi_t \sim \mathit{Exponential}(\lambda)$
- $\alpha \in \{0, 0.25, 0.5, 0.75, 0.99\}$
- $a_t \in \{-1, 0, 1\}$
- $x_0 = 1, t_0 = 0$
- $T = 2.$

The action set is changed to be $a_t \in \{-1, 0, 1\}$ due to the dimensionality problem, to decrease the number of possible state x_t values for $t \geq 1$.

LQR problem with Theoretical Exponential noise

We start the computations at time $t = 2$, then go backwards in time in 1 time unit steps.

Step 1. t = 2

$J(2, x_2) = \inf_{a_2} E[x_2^2 + a_2^2 | x_2, a_2] = x_2^2.$

Therefore, the minimizing action $a_2 = 0$.

Step 2. t = 1

$J(1, x_1) =$

$2x_1^2 + 2x_1 E[\xi_1 | x_1, a_1] + E[\xi_1^2 | x_1, a_1] + \inf_{a_1} \{2a_1^2 + 2x_1 a_1 + 2a_1 E[\xi_1 | x_1, a_1]\}$

Computing expected values:

$E[\xi_1 | x_1, a_1] = \frac{1}{\lambda}$

$E[\xi_1^2 | x_1, a_1] = Var[\xi_1 | x_1, a_1] + E^2[\xi_1 | x_1, a_1] = \frac{1}{\lambda^2} + \frac{1}{\lambda^2} = \frac{2}{\lambda^2}.$

Then,

$J(1, x_1) = 2x_1^2 + \frac{2}{\lambda} x_1 + \frac{2}{\lambda^2} + 2 \inf_{a_1} \{a_1^2 + (x_1 + \frac{1}{\lambda}) a_1\}.$

LQR problem with Theoretical Exponential noise

$\phi(a_1) = a_1^2 + (x_1 + \frac{1}{\lambda})a_1$
$a_1 = -1 : \phi(-1) = 1 - x_1 - \frac{1}{\lambda}$
$a_1 = 0 : \phi(0) = 0$
$a_1 = 1 : \phi(1) = 1 + x_1 + \frac{1}{\lambda}$
The behavior of $\phi(a_1)$ was checked for several λ values graphically. Different from Bernoulli noise, only two cases were identified for further computations. Here we use the fact that $E[\xi_1|x_1, a_1] = \frac{1}{\lambda}$

Case 1. $\lambda < 1$ OR
$\lambda \geq 1$ and $x \geq 1 - E[\xi_1|x_1, a_1]$.

Case 2. $\lambda \geq 1$ and $x \in [0, 1 - E[\xi_1|x_1, a_1])$

For both cases, the optimizing action $\mathbf{a_1} = \mathbf{0}$.

Case 1. $\lambda < 1$ OR

$\lambda \geq 1$ and $x \geq 1 - E[\xi_1 | x_1, a_1]$.

The optimizing action $\mathbf{a_1} = -\mathbf{1}$.

$$J(1, x_1) = 2x_1^2 + \frac{2}{\lambda}x_1 + \frac{2}{\lambda^2} + 2(1 - x_1 - \frac{1}{\lambda}).$$
$$= 2x_1^2 + 2(\frac{1}{\lambda} - 1)x_1 + 2(1 - \frac{1}{\lambda} + \frac{1}{\lambda^2}).$$
$$J(0, x_0) = \inf_{a_0} E[x_0^2 + a_0^2 + J(1, x_1) | x_0, a_0]$$
$$= \inf_{a_0} E[x_0^2 + a_0^2 + 2(x_0 + a_0 + \xi_0)^2 + 2(\frac{1}{\lambda} - 1)(x_0 + a_0 + \xi_0) + 2(1 - \frac{1}{\lambda} + \frac{1}{\lambda^2}) | x_0, a_0]$$
$$= 3 + 2(\frac{1}{\lambda} - 1) + 2(1 - \frac{1}{\lambda} + \frac{1}{\lambda^2}) + 2(\frac{1}{\lambda} + 1)E[\xi_0 | a_0] + 2E[\xi_0^2 | a_0] + \inf_{a_0}\{3a_0^2$$
$$+2(\frac{1}{\lambda} + 1)a_0 + 4a_0 E[\xi_0 | a_0]\}$$

LQR problem with Theoretical Exponential noise

Use $x_0 = 1$, $E[\xi_0|x_0, a_0] = \frac{1}{\lambda}$ and $E[\xi_0^2|x_0, a_0] = \frac{2}{\lambda^2}$:
$J(0, x_0) = 3 + \frac{2}{\lambda} + \frac{8}{\lambda^2} + \inf_{a_0}\{3a_0^2 + 2(\frac{3}{\lambda} + 1)a_0\}$.
$\phi(a_0) = 3a_0^2 + 2(\frac{3}{\lambda} + 1)a_0$
$a_1 = -1 : \phi(-1) = 3 - 2(\frac{3}{\lambda} + 1) = 1 - \frac{6}{\lambda}$
$a_1 = 0 : \phi(0) = 0$
$a_1 = 1 : \phi(1) = 3 + 2(\frac{3}{\lambda} + 1) = 4 + \frac{6}{\lambda}$
Minimizing action: $\mathbf{a_0 = -1}$, $\mathbf{J(0, x_0) = 4 - \frac{4}{\lambda} + \frac{8}{\lambda^2}}$.

LQR problem with Theoretical Exponential noise

Case 2. $\lambda \geq 1$ and $x \in [0, 1 - E[\xi_1|x_1, a_1])$.
The optimizing action $\mathbf{a_1} = \mathbf{0}$.

$$J(1, x_1) = 2x_1^2 + \frac{2}{\lambda}x_1 + \frac{2}{\lambda^2} + 2 \times 0$$
$$= 2x_1^2 + \frac{2}{\lambda}x_1 + \frac{2}{\lambda^2}$$
$$J(0, x_0) = \inf_{a_0} E[x_0^2 + a_0^2 + J(1, x_1)|x_0, a_0]$$
$$= \inf_{a_0} E[x_0^2 + a_0^2 + 2(x_0 + a_0 + \xi_0)^2 + \frac{2}{\lambda}(x_0 + a_0 + \xi_0) + \frac{2}{\lambda^2}|x_0, a_0]$$
$$= 3 + \frac{2}{\lambda} + \frac{2}{\lambda^2} + 2(2 + \frac{1}{\lambda})E[\xi_0|a_0] + 2E[\xi_1^2|a_0] + \inf_{a_0}\{3a_0^2 + 2(2 + \frac{1}{\lambda})a_0\}$$

Use $x_0 = 1$, $E[\xi_0|x_0, a_0] = \frac{1}{\lambda}$ and $E[\xi_0^2|x_0, a_0] = \frac{2}{\lambda^2}$:

$$J(0, x_0) = 3 + \frac{6}{\lambda} + \frac{8}{\lambda^2} + \inf_{a_0}\{3a_0^2 + 2(2 + \frac{3}{\lambda})a_0\}$$

$$\phi(a_0) = 3a_0^2 + 2(\frac{3}{\lambda} + 2)a_0$$
$$a_1 = -1 : \phi(-1) = 3 - 2(\frac{3}{\lambda} + 2) = -1 - \frac{6}{\lambda}$$
$$a_1 = 0 : \phi(0) = 0$$
$$a_1 = 1 : \phi(1) = 3 + 2(\frac{3}{\lambda} + 2) = 7 + \frac{6}{\lambda}$$

The minimizing action: $\mathbf{a_0} = -\mathbf{1}$. $\mathbf{J(0, x_0)} = \mathbf{2} + \frac{\mathbf{8}}{\lambda^\mathbf{2}}$.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

Law of Large Numbers: for a sample generated from an experiment, we should maximize the sample size for the sample's distribution to converge to the original distribution.

Having 3 samples for noise distribution and having 3 possible actions, at each time t, there would be 9^t possible states.
Take 6 samples for noise term. 18 possible x_1, 324 possible x_2 values.
Use *numpy.random.exponential* to take 6 random samples from the exponential distribution with $\lambda = 1.0$ and some fixed random seed of 41 for the reproducibility of the experiment.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

Problem setting

- Transition (linear) function: $x_{t+1} = x_t + a_t + \xi_t$ for the **random noise** ξ_t
- Running (quadratic) cost: $c_t(x, a, t) = x_t^2 + a_t^2$
- $\xi_t \in \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\}$
- $\alpha \in \{0, 0.25, 0.5, 0.75, 0.99\}$
- $a_t \in \{-1, 0, 1\}$
- $x_0 = 1, t_0 = 0$
- $T = 2.$

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

Step 2. t = 1

$$
\begin{aligned}
J(1, x_1) &= \inf_{a_1} E[x_1^2 + a_1^2 + J(2, x_2) | x_1, a_1] \\
&= \inf_{a_1} E[x_1^2 + a_1^2 + (x_1 + a_1 + \xi_1)^2 | x_1, a_1] \\
&= \inf_{a_1} E[x_1^2 + a_1^2 + x_1^2 + a_1^2 + \xi_1^2 + 2x_1 a_1 + 2x_1\xi_1 + 2a_1\xi_1 | x_1, a_1] \\
&= \inf_{a_1} E[2x_1^2 + 2a_1^2 + \xi_1^2 + 2x_1 a_1 + 2x_1\xi_1 + 2a_1\xi_1 | x_1, a_1] \\
&= 2x_1^2 + \inf_{a_1} \{2a_1^2 + 2x_1 a_1 + E[\xi_1^2 + 2x_1\xi_1 + 2a_1\xi_1 | x_1, a_1]\}
\end{aligned}
$$

$$= 2x_1^2 + 2x_1 E[\xi_1 | x_1, a_1] + E[\xi_1^2 | x_1, a_1] + 2\inf_{a_1} \{a_1^2 + x_1 a_1 + a_1 E[\xi_1 | x_1, a_1]\}$$

Now that we have the samples from an exponential distribution, the expected value is taken to be the sample mean and the sample mean squared replaces the second moment. So,

$$E[\xi_1 | x_1, a_1] = \tfrac{1}{6}(0.04 + 0.05 + 0.12 + 0.29 + 0.93 + 1.13) = 0.43$$
$$E[\xi_1^2 | x_1, a_1] = \tfrac{1}{6}(0.04^2 + 0.05^2 + 0.12^2 + 0.29^2 + 0.93^2 + 1.13^2) = 0.37.$$

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

$J(1, x_1) = 2x_1^2 + 0.86x_1 + 0.37 + 2\inf_{a_1}\{a_1^2 + (x_1 + 0.43)a_1\}$
Meaning that
$\phi(a_1) = a_1^2 + (x_1 + 0.43)a_1$
$a_1 = -1 : \phi(-1) = 1 - x_1 - 0.43 = -x_1 + 0.57$
$a_1 = 0 : \phi(0) = 0$
$a_1 = 1 : \phi(1) = 1 + x_1 + 0.43 = x_1 + 1.43$
From the theoretical noise example, we know that there are two cases given that $\lambda = 1.0$. Note that $E[\xi_1 | x_1, a_1] = 0.43$. This will be used to obtain the equations for $J(1, x_1)$ for some given x_1.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

Case 1. $\mathbf{x_1 \geq 0.57}$.
In this case, the optimizing action $a_1 = -1$.
So,

$$J(1, x_1) = 2x_1^2 + 0.86x_1 + 0.37 + 2(-x_1 + 0.57).$$
$$= 2x_1^2 - 1.14x_1 + 1.51.$$

Case 2. $\mathbf{x_1 < 0.57}$.
In this case, the optimizing action $a_1 = 0$.
So,

$$J(1, x_1) = 2x_1^2 + 0.86x_1 + 0.37 + 2 \times 0.$$
$$= 2x_1^2 + 0.86x_1 + 0.37.$$

For further computations for $J(0, x_0)$, it has to be noted that each noise was sampled randomly, so we know that each of the noises has equal probabilities.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0$

Knowing that

- $x_1 = x_0 + a_0 + \xi_0$
- $x_0 = 1$
- $a_0 \in \{-1, 0, 1\}$
- $\xi_t \in \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\}$

different cases should be considered. Each a_0 case will be considered separately given that $x_0 = 1$ is fixed.

Case a. $\mathbf{a_0 = -1}$

$$J(0, x_0, a_0 = -1) = \inf_{a_0} E[x_0^2 + a_0^2 + J(1, x_1)|x_0 = 1, a_0 = -1]$$
$$= \inf_{a_0} E[x_0^2 + a_0^2 + J(1, x_0 + a_0 + \xi_0)|x_0 = 1, a_0 = -1]$$
$$= x_0^2 + a_0^2 + (1/6)(J(1, 1-1+0.04) + J(1, 1-1+0.05) + J(1, 1-1+0.12) + J(1, 1-1+0.29) + J(1, 1-1+0.93) + J(1, 1-1+1.13)$$
$$= x_0^2 + a_0^2 + (1/6)(J(1, 0.04) + J(1, 0.05) + J(1, 0.12) + J(1, 0.29) + J(1, 0.93) + J(1, 1.13))]$$
$$= 1^2 + (-1)^2 + (1/6)(0.41 + 0.42 + 0.50 + 0.79 + 2.18 + 2.78)$$
$$= 3.18$$

Cases b., c., d., e. with other a_0 are computed similarly.

After considering all these cases for a_0 values, we deduce the final $J(0, x_0)$.

$$\begin{aligned} J(0, x_0) &= \inf_{a_0 \in \{-1,0,1\}} J(0, x_0) \\ &= \inf_{a_0 \in \{-1,0,1\}} \{J(0, x_0, a_0 = -1), J(0, x_0, a_0 = 0), J(0, x_0, a_0 = 1)\} \\ &= \inf\{3.18, 5.34, 12.91\} \\ &= 3.18 \end{aligned}$$

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

From **Remark 1.**, we know that with a risk level $\alpha = 0$, $AVaR_\alpha(X) = \mathbb{E}[X]$. The behavior of $AVaR_\alpha(X)$ for $\alpha \neq 0$ should be investigated next. For this, we take an example $\alpha = 0.25$.
For $\alpha = 0.25$, we are solving the similar LQR problem with

Problem setting

- Transition (linear) function: $x_{t+1} = x_t + a_t + \xi_t$ for the **random noise** ξ_t
- Running (quadratic) cost: $c_t(x, a, t) = x_t^2 + a_t^2$
- $\xi_t \in \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\}$
- $a_t \in \{-1, 0, 1\}$
- $x_0 = 1, t_0 = 0$
- $T = 2$.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

Step 1. t = 2
$J(2, x_2) = \inf_{a_2} AVaR_{0.25}[x_2^2 + a_2^2 | x_2, a_2] = \inf_{a_2} AVaR_{0.25}[x_2^2 + a_2^2] = x_2^2.$
Therefore, the minimizing action $a_2 = 0$.
Step 2. t = 1

$$\begin{aligned} J(1, x_1) &= \inf_{a_1} AVaR_{0.25}[x_1^2 + a_1^2 + J(2, x_2) | x_1, a_1] \\ &= \inf_{a_1} AVaR_{0.25}[x_1^2 + a_1^2 + (x_1 + a_1 + \xi_1)^2 | x_1, a_1] \end{aligned}$$

Then, we use the **Theorem 1**.

$$\begin{aligned} J(1, x_1) &= \inf_{s \in \mathbb{R}} \{s + \tfrac{1}{1-0.25} \inf_{a_1} E[(x_1^2 + a_1^2 + (x_1 + a_1 + \xi_1)^2 - s)^+ | x_1, a_1]\} \\ &= x_1^2 + a_1^2 + \inf_{s \in \mathbb{R}} \{s + \tfrac{1}{1-0.25} \inf_{a_1} E[((x_1 + a_1 + \xi_1)^2 - s)^+ | x_1, a_1]\} \\ &= x_1^2 + a_1^2 + s^* + \tfrac{1}{1-0.25} \inf_{a_1} E[((x_1 + a_1 + \xi_1)^2 - s^*)^+ | x_1, a_1] \end{aligned}$$

We have to find the quantile s^*.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

- We know that
  $s^* \triangleq VaR_\alpha((x_1+a_1+\xi_1)^2)) = inf\{x \in \mathbb{R} : \mathbb{P}((x_1+a_1+\xi_1)^2 \leq x) \geq \alpha\}$.
- Compute all possible x_1 values are computed to consider each case of possible x_1 and a_1 value pairs.
- For each case of x_1 and a_1 value pair compute quantile from $(x_1 + a_1 + \xi_1)^2$ for $\xi_1 \in \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\}$ and getting the quantile using these 6 values by the help of *numpy.quantile*.
- For set of random noises $\Xi = \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\}$:
  $J(1, x_1) = x_1^2 + a_1^2 + s^* + \frac{4}{3} \sum_{\xi_1 \in \Xi} (\frac{1}{6})((x_1 + a_1 + \xi_1)^2 - s^*)^+)$
- Automize 54 computations by using the Python code looping through all x_1, a_1 combinations, giving the s^* and $J(1, x_1, a_1)$ values. In the table, $\pi^*(1, x_1)$ stands for the optimal action for the given $t = 1$ and x_1 value.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

x_1	a_1	s^*	$J(1, x_1, a_1)$	$J(1, x_1)$	$\pi^*(1, x_1)$
0.04	-1	0.13	1.64	0.55	0
	0	0.01	0.55		
	1	1.23	3.74		
0.05	-1	0.13	1.64	0.56	0
	0	0.01	0.56		
	1	1.25	3.77		
0.12	-1	0.13	1.54	0.66	0
	0	0.04	0.66		
	1	1.41	4.01		
0.29	-1	0.18	1.42	0.99	0
	0	0.13	0.99		
	1	1.84	4.68		
0.93	-1	0.00	2.29	2.29	-1
	0	1.00	3.26		
	1	3.99	8.24		
1.13	-1	0.04	2.94	2.94	-1
	0	1.43	4.31		
	1	4.83	9.68		

Table 1. $J(1, x_1, a_1)$ *values*

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

Step 3. t = 0

$J(0, x_0) = \inf_{a_0} AVaR_{0.25}[x_0^2 + a_0^2 + J(1, x_1)|x_0, a_0]$

Knowing that $x_0 = 1$,

$J(0, x_0) = \inf_{a_0} AVaR_{0.25}[1^2 + a_0^2 + J(1, x_1)|a_0]$

Using **Theorem 1**,

$$
\begin{aligned}
J(0, x_0) &= \inf_{s\in\mathbb{R}}\{s + \tfrac{1}{1-0.25}\inf_{a_0} E[(1 + a_0^2 + J(1, x_1) - s)^+|a_0]\} \\
&= 1 + a_0^2 + \inf_{s\in\mathbb{R}}\{s + \tfrac{1}{1-0.25}\inf_{a_0} E[(J(1, x_1) - s)^+|a_0]\} \\
&= 1 + a_0^2 + s^* + \tfrac{1}{1-0.25}\inf_{a_0} E[(J(1, x_1) - s)^+|a_0] \\
&= 1 + a_0^2 + s^* + \tfrac{4}{3}\textstyle\sum_{\xi_1\in\Xi}(\tfrac{1}{6})(J(1, x_1) - s^*)^+)
\end{aligned}
$$

As we know, $x_1 = x_0 + a_0 + \xi_0$ and $x_0 = 1$, $a_0 \in \{-1, 0, 1\}$,

$\xi_t \in \{0.04, 0.05, 0.12, 0.29, 0.93, 1.13\} = \Xi$.

So, each a_0 case will be considered separately given that $x_0 = 1$ is fixed.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

Case a. $\mathbf{a_0 = -1}$

$s^* = VaR_{0.25}(J(1, 1-1+\xi_0)) = VaR_{0.25}(J(\xi_0)) = 0.59$

Therefore,

$$
\begin{aligned}
J(0, x_0, a_0 = -1) &= 1 + (-1)^2 + s^* + \tfrac{4}{3}\textstyle\sum_{\xi_1 \in \Xi}(\tfrac{1}{6})(J(1, x_1) - s^*)^+) \\
&= 1 + (-1)^2 + s^* + \tfrac{4}{3}\textstyle\sum_{\xi_1 \in \Xi}(\tfrac{1}{6})(J(1, 1-1+\xi_1) - s^*)^+) \\
&= 1 + 1 + 0.59 + \tfrac{4}{3}\textstyle\sum_{\xi_1 \in \Xi}(\tfrac{1}{6})(J(1, \xi_1) - 0.59)^+) \\
&= 2.92 + \tfrac{2}{9}[(J(1, 0.04) - 0.59)^+ + (J(1, 0.05) - 0.59)^+ \\
&\quad +(J(1, 0.12) - 0.59)^+ + (J(1, 0.29) - 0.59)^+ \\
&\quad +(J(1, 0.93) - 0.59)^+ + (J(1, 1.13) - 0.59)^+] \\
&= 2.59 + \tfrac{2}{9}[0 + 0.03 + 0.07 + 0.4 + 1.7 + 2.35] \\
&= 3.59
\end{aligned}
$$

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.25$

Cases b., c., d., e. with other a_0 are computed similarly.
After considering all these cases for a_0 values, we deduce the final $J(0, x_0)$.

$$\begin{aligned} J(0, x_0) &= \inf_{a_0 \in \{-1,0,1\}} J(0, x_0) \\ &= \inf_{a_0 \in \{-1,0,1\}} \{J(0, x_0, a_0 = -1), J(0, x_0, a_0 = 0), J(0, x_0, a_0 = 1)\} \\ &= \inf\{3.59, 6.23, 14.56\} \\ &= 3.59. \end{aligned}$$

Also, minimizing action is $a_0 = -1$.

LQR problem with Sampled Exponential noise at risk level $\alpha = 0.5, 0.75, 0.99$

The calculation with nonzero risk α is done similarly to the $\alpha = 0.25$ case. The results of the work are collected in a table.

α	$J(0, x_0)$
0	3.18
0.25	3.59
0.5	4.37
0.75	5.61
0.99	6.81

Table 2. $J(0, x_0)$ values versus risk level α

Plots for LQR Problem with Sampled Exponential Noise at risk level $\alpha = 0$

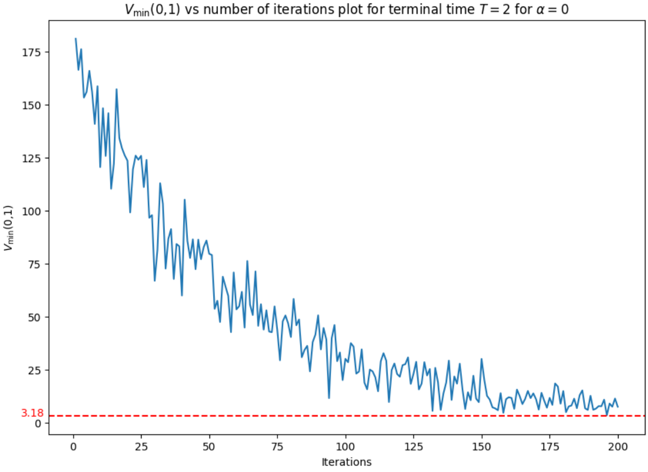


Figure 2. $V_{\min}(0,1)$ vs number of iterations plot for terminal time $T = 2$ for $\alpha = 0$

Plots for LQR Problem with Sampled Exponential Noise at risk level $\alpha = 0$

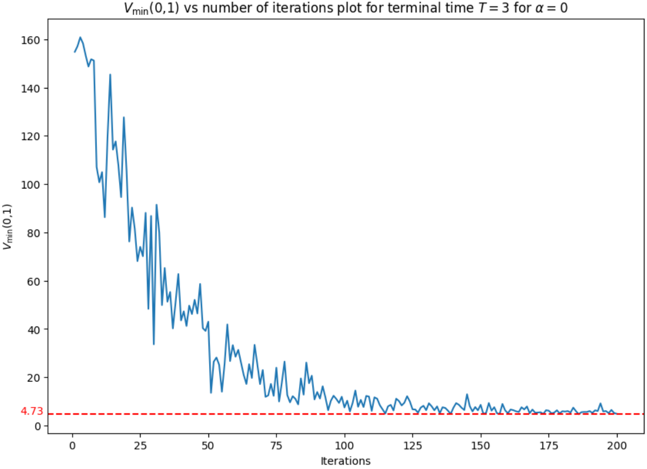


Figure 2. $V_{\min}(0,1)$ *vs number of iterations plot for terminal time* $T = 3$ *for* $\alpha = 0$

Approximate Dynamic Programming algorithm

- Select and fix the number of iterations N.
- Set the iteration counter n $= 1$, set the initial parameters for state(x_0), initial time t_0, terminal time T.
- Set the action space A (so that $a_t \in A$) and take a random noise samples ξ_t (so that $\xi_t \sim \mathit{SelectedDistribution}$)
- Initialize an initial approximation $\overline{V}_t^0$, $\forall t \in \{1, ..., T\}$

Approximate Dynamic Programming algorithm

- **Forward pass**: For each $t \in \{1, ..., T\}$ create a random path by randomly choosing (a_t, ξ_t).
- **Backward pass**: For each $t \in \{1, ..., T\}$ compute following using the the selected learning rate α and the decision $\hat{a}_t^n$ obtained from forward pass:

$$\hat{v}_t^n = c(x_t^n, \hat{a}_t^n) + \hat{v}_t^n, \quad \text{with } \hat{v}_{T+1}^n = 0$$

$$\overline{V}_{t-1}^n(x_{t-1}^{a,n}) = U^V\left(\overline{V}_{t-1}^{n-1}(x_{t-1}^{a,n}), x_{t-1}^{a,n}, \hat{v}_t^n\right) = (1-\alpha)\overline{V}_{t-1}^{n-1} + \alpha\hat{v}_t^n$$

- Increment n until the iteration number $n > N$.
- Return the value functions $\overline{V}_t^N(x_t^{a,n}) \quad \forall t \in \{1, ..., T\}$ and $x_t \in X$.

Evaluation

To evaluate the performance of the code on the LQR problem of this thesis, the check was done with all the given risk levels $\alpha \in \{0, 0.25, 0.5, 0.75, 0.99\}$.

α	$J_{theor}(0, x_0)$	$J_{code}(0, x_0)$	error
0	3.18	3.18	0
0.25	3.59	3.59	0
0.5	4.37	4.37	0
0.75	5.61	5.61	0
0.99	6.81	6.81	0

Table 3. $J_{theor}(0, x_0)$ values comparison to $J_{code}(0, x_0)$ values versus risk level α

Results

LQR Problem with Theoretical Exponential Noise at a risk level $\alpha = 0$ gave a pattern regarding the optimal policy. It considered two cases:

- **Case 1.** $\lambda < \mathbf{1}$ **OR**
  $\lambda \geq \mathbf{1}$ and $\mathbf{x} \geq \mathbf{1} - \mathbf{E}[\xi_{\mathbf{1}} | \mathbf{x_1}, \mathbf{a_1}]$.
- **Case 2.** $\lambda \geq \mathbf{1}$ and $\mathbf{x} \in [\mathbf{0}, \mathbf{1} - \mathbf{E}[\xi_{\mathbf{1}} | \mathbf{x_1}, \mathbf{a_1}])$

So, for Case 1, the optimal policy is $a_0 = -1, a_1 = 0, a_2 = 0$.
For Case 2, the optimal policy is $a_0 = -1, a_1 = -1, a_2 = 0$.

Results

The LQR Problem with Sampled Exponential Noise at a risk level $\alpha = 0.25$ justifies this observation. With the $\lambda = 1$ and $E[\xi_1|x_1, a_1] = 0.43$, we know that $1 - E[\xi_1|x_1, a_1] = 1 - 0.43 = 0.57$. As seen in Table 3.2, for all $x_1 < 0.57$, the optimal action $a_1 = 0$, while the optimal action $a_1 = -1$ for $x_1 \geq 0.57$.

Both for theoretical and exponential noise, $a_0 = -1$ is satisfied for any case. Intuitively, $c_t(x_t, a_t) = x_t^2 + a_t^2$, so $a_0 = 0$ may be assumed to be a minimizing action.
However, since $c_1(x_1, a_1) = x_1^2 + a_1^2$ and $x_1 = x_0 + a_0 + \xi_0$ for $x_0 = 1$ and $\xi_0 > 0$, x_1 is minimized by $a_0 = -1$ due to the transition function. So, compared with Bernoulli noise which can take a value of -1, in the case of the exponential noise, **the action value has to be minimized even further as exponential noise punishes action a_t more**, giving a higher weight to it. So, this point is verified both theoretically and in experiments.

Results

The LQR Problem with Sampled Exponential Noise at a risk level $\alpha \neq 0$ also proves the **Remark 1.**, where

$$\lim_{\alpha \to 1} AVaR_\alpha(X) = \text{ess sup} X \leq \infty$$

This is proved by the pattern that the value function is directly proportional to the risk level α. So, for higher risks, higher-value functions are expected as stated in the Lemma.

References

Nicole Bäuerle, Jonathan Ott (2011)
Markov Decision Processes with Average-Value-at-Risk criteria
Mathematical Methods of Operations Research.. 74. 361-379.

Martijn Mes, Arturo Perez Rivera (2017)
Approximate Dynamic Programming by Practical Examples
Beta Working Paper series 495.

Kerem Ugurlu (2017)
Controlled Markov Processes with AVaR Criteria for Unbounded Costs
Journal of Computational and Applied Mathematics. 319. 24-37.

Thank you for your attention