Risk Sensitive Path Integral Control

Bart van den Broek Wim Wiegerinck Bert Kappen

SNN, Radboud University Nijmegen
6525 EZ Nijmegen, The Netherlands

1 Introduction

The objective in conventional stochastic optimal control is to minimize an expected cost-to-go. Risk sensitive optimal control generalizes this objective by minimizing an expected exponentiated cost-to-go. Depending on its risk parameter θ, expected exponentiated cost-to-go puts more emphasis on the mode of the distribution of the cost-to-go, or on its tail, and in that way allows for a modelling of more risk seeking $(\theta < 0)$ or risk averse $(\theta > 0)$ behaviour. The conventional optimal control can be viewed as a special case of risk sensitive optimal control with a risk neutral parameter $\theta = 0$.

Risk sensitive control was first considered in continuous space in the LEQG problem [1], which is the risk sensitive analogue of the Linear Quadratic Gaussian (LQG) problem. Relations with other fields such as differential games and robust control have initiated a lot of interest for risk sensitive control.

The dynamic programming (DP) principle provides a well-known approach to a global solution in stochastic optimal control. In the continuous time and state setting that we will consider, it follows from the DP principle that the solution to the control problem satisfies the so-called Hamilton-Jacobi-Bellman (HJB) equation, which is a second order nonlinear partial differential equation. If the dynamics is linear and the cost is quadratic in both state and control, the HJB equation can be solved exactly, both for LQG and LEQG.

Recently, a path integral formalism has been developed to solve the HJB equation. This formalism is applicable if (1) both the noise and the control are additive to the (nonlinear) dynamics, (2) the cost is quadratic in the control (but arbitrary in the state), and (3) the noise satisfies certain additional conditions. Under these conditions the nonlinear HJB equation can be transformed into a linear one, which can be solved by forward integration of a diffusion process [2]. This formalism contains LQG control as a special case.

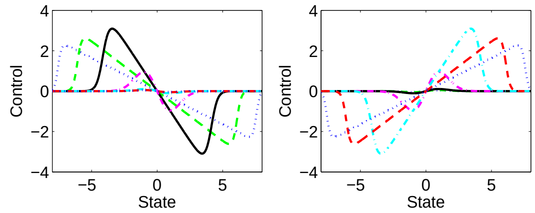
In our full paper [3] we show how path integral control generalizes to risk sensitive control problems. The required conditions to apply path integral control in the risk sensitive case are the same as those in the risk neutral setting. As a consequence, characteristics of path integral control, such as superposition of controls, symmetry breaking and approximate inference, carry over to the setting of risk sensitive control.

2 Risk Sensitive Behaviour

We make use of the path integral solutions to obtain insight in risk sensitive control, and in particular interpret risk sensitive optimal control as emergent behaviour of an agent with an optimistic or pessimistic attitude.

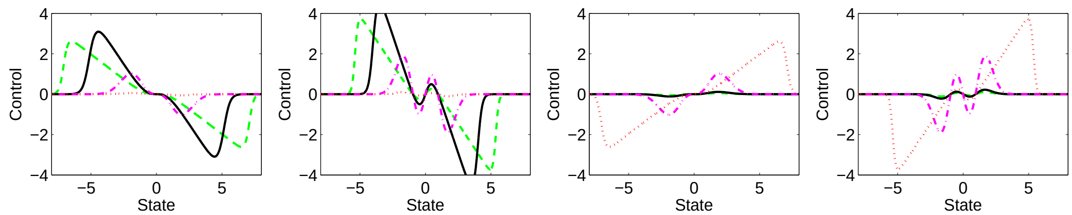
We first illustrate this in the case where the agent has to reach a target or avoid a threat, see Figure 1. The figure shows controls as a function of the state at a fixed time, for risk seeking $(\lambda_\theta < 1)$, neutral $(\lambda_\theta = 1)$ and averse $(\lambda_\theta > 1)$ behaviour. The parameter λ_θ arises naturally in risk sensitive path integral control and is related to θ by $\lambda_\theta^{-1} = \lambda_0^{-1} - \theta$, where λ_0 is a constant. In case of a target (threat), when $\lambda_\theta < 0$ $(\lambda_\theta > 0)$ the control is about zero, which can be explained by the fact that in that case the control is less sensitive to targets (threats). We see that the nonzero control has a bounded support, that increases if $|\lambda_\theta^{-1}|$ increases. This can be understood from the fact that the control is nonzero only when the agent gets sufficiently close to the target or threat. If not, control cost towards the target is too expensive compared to the reward, or the probability to hit the threat is so small that a significant control is not needed.

The second example considers the phenomenon of symmetry breaking. Figure 2 shows the control at two different times t in case of two targets or two threats that are located at -1 and $+1$. In case of two targets,



(a) Target (b) Threat

Figure 1: The control as a function of the state in case of a target (a) or threat (b) located at 0, with $\lambda_\theta = -\frac{1}{3}$ ($\cdots$), $-\frac{1}{2}$ ($--$), -1 ($-\cdot-$), $\pm\infty$ ($-\cdot-$), 1 ($-$), $\frac{1}{2}$ ($--$), $\frac{1}{3}$ ($\cdots$).



(a) Two targets, $t = 0$ (b) Two targets, $t = 0.5$ (c) Two threats, $t = 0$ (d) Two threats, $t = 0.5$

Figure 2: The control as a function of the state in case of two targets (a,b) or threats (c,d), with $\lambda_\theta = -\frac{1}{2}$ ($\cdots$), $\pm\infty$ ($-\cdot-$), 1 ($-$), $\frac{1}{2}$ ($--$). There is no symmetry breaking at time $t = 0$ but there is at time $t = 0.5$.

at an early time $(t = 0)$ the optimal control is towards the middle, whereas at a later time $(t = 0.5)$ there is a symmetry breaking, and the agent makes a choice towards which target it steers. In case of two threats a similar but reversed phenomenon occurs. The risk sensitivity does not influence the point of symmetry breaking, but it does influence the magnitude of the control, just as in the single target or threat case.

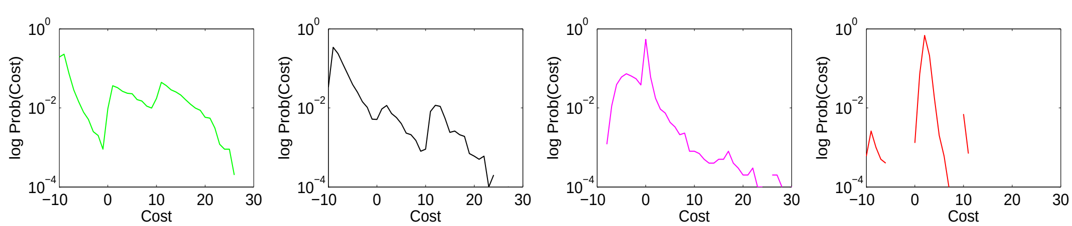
A third case that we consider is a target and an adjacent threat. Figure 3 shows histograms of the log-probability of the cost. We see that with larger θ, the mode of the distribution shifts to higher costs. On the other hand, the tails of the distribution at the high cost end are thinner with larger θ. This is what is to be expected: small θ is more greedy, aiming at low cost, however at the expense of some outliers with high costs. A larger θ is more cautious, reducing the probability of costly outliers.

References

[1] D.H. Jacobson. Optimal stochastic linear systems with exponential performance criteria and relation to deterministic differential games. *IEEE Trans Autom Control*, 18:124–131, 1973.

[2] H. J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of statistical mechanics: theory and experiment*, page 11011, November 2005.

[3] Bart van den Broek, Wim Wiegerinck, and Bert Kappen. Risk sensitive path integral control. In *Proceedings of UAI*, 2010.



(a) $\theta = -1$ (b) $\theta = 0$ (c) $\theta = 1$ (d) $\theta = 3$

Figure 3: The log-probability of the cost in the case of a target and an adjacent threat.