Risk-Sensitive Control via Iterative Linearizations

Vincent Roulet, Maryam Fazel,
Siddhartha Srinivasa, Zaid Harchaoui

SIAM Pacific NorthWest

Control Problem

Dynamics

$$x_0 = \hat{x}_0 \qquad x_{t+1} = \phi_t(x_t, u_t + w_t) \quad \text{for } t \in \{0, \dots \tau - 1\} \quad \text{(Dyn)}$$

with $x_t \in \mathbb{R}^d$ states, $u_t \in \mathbb{R}^p$ controls, $w_t \sim \mathcal{N}(0, \sigma^2 \operatorname{I}_p)$ noises.
Denote $\bar{x} = (x_0; \dots; x_\tau)$, $\bar{u} = (u_0, \dots, u_{\tau-1})$, $\bar{w} = (w_0, \dots, w_{\tau-1})$

Control Problem

Dynamics

$$x_0 = \hat{x}_0 \qquad x_{t+1} = \phi_t(x_t, u_t + w_t) \quad \text{for } t \in \{0, \ldots \tau - 1\} \quad \text{(Dyn)}$$

with $x_t \in \mathbb{R}^d$ states, $u_t \in \mathbb{R}^p$ controls, $w_t \sim \mathcal{N}(0, \sigma^2 \, \mathrm{I}_p)$ noises.
Denote $\bar{x} = (x_0; \ldots; x_\tau)$, $\bar{u} = (u_0, \ldots, u_{\tau-1})$, $\bar{w} = (w_0, \ldots, w_{\tau-1})$

Noiseless Trajectory is a function of $\bar{u}$

$$\tilde{x}(\bar{u}) = (\tilde{x}_1(\bar{u}); \ldots; \tilde{x}_\tau(\bar{u})),$$
$$\tilde{x}_1(\bar{u}) = \phi_0(\hat{x}_0, u_0), \qquad \tilde{x}_{t+1}(\bar{u}) = \phi_t(\tilde{x}_t(\bar{u}), u_t).$$

Control Problem

Dynamics

$$x_0 = \hat{x}_0 \qquad x_{t+1} = \phi_t(x_t, u_t + w_t) \quad \text{for } t \in \{0, \ldots \tau - 1\} \quad \text{(Dyn)}$$

with $x_t \in \mathbb{R}^d$ states, $u_t \in \mathbb{R}^p$ controls, $w_t \sim \mathcal{N}(0, \sigma^2 \operatorname{I}_p)$ noises.
Denote $\bar{x} = (x_0; \ldots; x_\tau)$, $\bar{u} = (u_0, \ldots, u_{\tau-1})$, $\bar{w} = (w_0, \ldots, w_{\tau-1})$

Noiseless Trajectory is a function of $\bar{u}$

$$\tilde{x}(\bar{u}) = (\tilde{x}_1(\bar{u}); \ldots; \tilde{x}_\tau(\bar{u})),$$
$$\tilde{x}_1(\bar{u}) = \phi_0(\hat{x}_0, u_0), \qquad \tilde{x}_{t+1}(\bar{u}) = \phi_t(\tilde{x}_t(\bar{u}), u_t).$$

Costs

$$h(\bar{x}) = \sum_{t=0}^{\tau} h_t(x_t), \qquad g(\bar{u}) = \sum_{t=0}^{\tau-1} g_t(u_t).$$

Control Problem

Dynamics

$$x_0 = \hat{x}_0 \qquad x_{t+1} = \phi_t(x_t, u_t + w_t) \quad \text{for } t \in \{0, \dots \tau - 1\} \quad \text{(Dyn)}$$

with $x_t \in \mathbb{R}^d$ states, $u_t \in \mathbb{R}^p$ controls, $w_t \sim \mathcal{N}(0, \sigma^2 \operatorname{I}_p)$ noises.
Denote $\bar{x} = (x_0; \dots; x_\tau)$, $\bar{u} = (u_0, \dots, u_{\tau-1})$, $\bar{w} = (w_0, \dots, w_{\tau-1})$

Noiseless Trajectory is a function of $\bar{u}$

$$\tilde{x}(\bar{u}) = (\tilde{x}_1(\bar{u}); \dots; \tilde{x}_\tau(\bar{u})),$$
$$\tilde{x}_1(\bar{u}) = \phi_0(\hat{x}_0, u_0), \qquad \tilde{x}_{t+1}(\bar{u}) = \phi_t(\tilde{x}_t(\bar{u}), u_t).$$

Costs

$$h(\bar{x}) = \sum_{t=0}^{\tau} h_t(x_t), \qquad g(\bar{u}) = \sum_{t=0}^{\tau-1} g_t(u_t).$$

Objective

$$\min_{u_0, \dots, u_{\tau-1}} \mathbb{E}_{\bar{w}}[h(\tilde{x}(\bar{u} + \bar{w})] + g(\bar{u})$$

Risk-Sensitive Objective

Risk Sensitive Objective (Whittle 1981)

$$\min_{u_0,\ldots,u_{\tau-1}} f_\theta(\bar{u}) = \left\{ \underbrace{\frac{1}{\theta} \log \mathbb{E}_{\bar{w}}[\exp \theta h(\tilde{x}(\bar{u} + \bar{w}))]}_{\eta_\theta(\bar{u})} + g(\bar{u}) \right\}$$

Risk-Sensitive Objective

Risk Sensitive Objective (Whittle 1981)

$$\min_{u_0,\ldots,u_{\tau-1}} f_\theta(\bar{u}) = \left\{ \underbrace{\frac{1}{\theta} \log \mathbb{E}_{\bar{w}}[\exp \theta h(\tilde{x}(\bar{u} + \bar{w})]}_{\eta_\theta(\bar{u})} + g(\bar{u}) \right\}$$

Interpretation for $\theta \to 0$

$$\eta_\theta(\bar{u}) = \mathbb{E}_{\bar{w}}[h(\tilde{x}(\bar{u} + \bar{w})] + \frac{\theta}{2} \mathbb{V}\mathrm{ar}_{\bar{w}} [h(\tilde{x}(\bar{u} + \bar{w})] + \mathcal{O}(\theta^2),$$

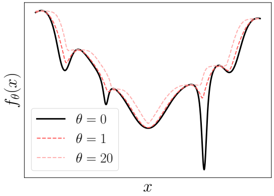
Risk-Sensitive Objective

Risk Sensitive Objective (Whittle 1981)

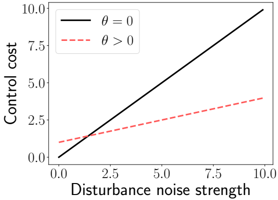
$$\min_{u_0,\ldots,u_{\tau-1}} f_\theta(\bar{u}) = \left\{ \underbrace{\frac{1}{\theta} \log \mathbb{E}_{\bar{w}}[\exp \theta h(\tilde{x}(\bar{u}+\bar{w})]}_{\eta_\theta(\bar{u})} + g(\bar{u}) \right\}$$

Interpretation for $\theta \to 0$

$$\eta_\theta(\bar{u}) = \mathbb{E}_{\bar{w}}[h(\tilde{x}(\bar{u}+\bar{w})] + \frac{\theta}{2} \mathbb{V}\mathrm{ar}_{\bar{w}} [h(\tilde{x}(\bar{u}+\bar{w})] + \mathcal{O}(\theta^2),$$



Effect of θ for
$f_\theta(x) = \frac{1}{\theta} \log \mathbb{E}_{w\sim\mathcal{N}(0,1)} \left[\exp \theta F(x+w)\right]$



Expected behavior of
the risk-sensitive controllers.

Iterative Linear Exponential Quadratic Gaussian algorithm

Linear Exponential Quadratic Gaussian (LEQG) (Whittle 1981)
For *linear* dynamics, *quadratic* costs, *small enough* θ, the risk-sensitive problem can be solved by dynamic programming

Iterative Linear Exponential Quadratic Gaussian algorithm

Linear Exponential Quadratic Gaussian (LEQG) (Whittle 1981)
For *linear* dynamics, *quadratic* costs, *small enough* θ, the risk-sensitive problem can be solved by dynamic programming

Iterative Linear Exponential Quadratic Gaussian (ILEQG) (Farshidian & Buchli 2015)
For non-linear dynamics, convex costs, small enough θ

(i) linearizes dynamics and approx. quad. the objectives around the current command and associated noiseless trajectory,

(ii) solves the LEQG problem to get an update direction,

(iii) moves along the update direction using a line-search.

Iterative Linear Exponential Quadratic Gaussian algorithm

Linear Exponential Quadratic Gaussian (LEQG) (Whittle 1981)
For *linear* dynamics, *quadratic* costs, *small enough* θ, the risk-sensitive problem can be solved by dynamic programming

Iterative Linear Exponential Quadratic Gaussian (ILEQG) (Farshidian & Buchli 2015)
For non-linear dynamics, convex costs, small enough θ

(i) linearizes dynamics and approx. quad. the objectives around the current command and associated noiseless trajectory,
(ii) solves the LEQG problem to get an update direction,
(iii) moves along the update direction using a line-search.

Questions:

- Does this algorithm converge? Under which assumptions?
- How is the line-search implemented? Is there a principled way?

ILEQG as Model Minimization

Approx. Trajectory for a given deviation $\bar{v}$ from current $\bar{u} = \bar{u}^{(k)}$,

$$\tilde{x}(\bar{u} + \bar{v} + \bar{w}) \approx \tilde{x}(\bar{u}) + \nabla \tilde{x}(\bar{u})^\top (\bar{v} + \bar{w})$$

ILEQG as Model Minimization

Approx. Trajectory for a given deviation $\bar{v}$ from current $\bar{u} = \bar{u}^{(k)}$,

$$\tilde{x}(\bar{u} + \bar{v} + \bar{w}) \approx \tilde{x}(\bar{u}) + \nabla \tilde{x}(\bar{u})^\top (\bar{v} + \bar{w})$$

Approx. Objective $f_\theta(\bar{u} + \bar{v}) \approx m_{f_\theta}(\bar{u} + \bar{v}; \bar{u})$ with

$$m_{f_\theta}(\bar{u}+\bar{v}; \bar{u}) \triangleq \frac{1}{\theta} \log \mathbb{E}_{\bar{w}} \exp \theta q_h\big(\tilde{x}(\bar{u})+\nabla\tilde{x}(\bar{u})^\top \bar{v}+\nabla\tilde{x}(\bar{u})^\top \bar{w}; \tilde{x}(\bar{u})\big) \\ + q_g(\bar{u} + \bar{v}; \bar{u}),$$

where $q_h(\bar{x} + \bar{y}; \bar{x}) \triangleq h(\bar{x}) + \nabla h(\bar{x})^\top \bar{y} + \bar{y}^\top \nabla^2 h(\bar{x}) \bar{y}/2.$

ILEQG as Model Minimization

Approx. Trajectory for a given deviation $\bar{v}$ from current $\bar{u} = \bar{u}^{(k)}$,

$$\tilde{x}(\bar{u} + \bar{v} + \bar{w}) \approx \tilde{x}(\bar{u}) + \nabla\tilde{x}(\bar{u})^\top(\bar{v} + \bar{w})$$

Approx. Objective $f_\theta(\bar{u} + \bar{v}) \approx m_{f_\theta}(\bar{u} + \bar{v}; \bar{u})$ with

$$m_{f_\theta}(\bar{u}+\bar{v}; \bar{u}) \triangleq \frac{1}{\theta}\log \mathbb{E}_{\bar{w}} \exp \theta q_h\big(\tilde{x}(\bar{u})+\nabla\tilde{x}(\bar{u})^\top\bar{v}+\nabla\tilde{x}(\bar{u})^\top\bar{w}; \tilde{x}(\bar{u})\big) + q_g(\bar{u} + \bar{v}; \bar{u}),$$

where $q_h(\bar{x} + \bar{y}; \bar{x}) \triangleq h(\bar{x}) + \nabla h(\bar{x})^\top\bar{y} + \bar{y}^\top\nabla^2 h(\bar{x})\bar{y}/2$.

ILEQG

1. Find $\bar{v}^* = \arg\min_{\bar{v}} m_{f_\theta}(\bar{u}+\bar{v}; \bar{u})$ *(LEQG by dyn. prog.)*
2. Find by line-search α s.t. $\bar{u}^{(k+1)} = \bar{u}^{(k)} + \alpha\bar{v}^*$

ILEQG from Optimization Viewpoint

Regularized ILEQG (RegILEQG) with step-size γ_k

$$\bar{u}^{(k+1)} = \arg\min_{\bar{v}} m_{f_\theta}(\bar{u}+\bar{v};\bar{u}) + \frac{1}{2\gamma_k}\|\bar{v}\|_2^2 \quad \textit{(LEQG by dyn. prog.)}$$

$\rightarrow$ how to choose γ_k ?

ILEQG from Optimization Viewpoint

Regularized ILEQG (RegILEQG) with step-size γ_k

$$\bar{u}^{(k+1)} = \arg\min_{\bar{v}} m_{f_\theta}(\bar{u}+\bar{v}; \bar{u}) + \frac{1}{2\gamma_k}\|\bar{v}\|_2^2 \quad \textit{(LEQG by dyn. prog.)}$$

$\rightarrow$ how to choose γ_k ?

Surrogate Risk-Sensitive Cost

$$\hat{f}_\theta(\bar{u}) = \underbrace{\frac{1}{\theta}\log \mathbb{E}_{\bar{w}} \exp[\theta h(\tilde{x}(\bar{u}) + \nabla \tilde{x}(\bar{u})^\top \bar{w})]}_{\hat{\eta}_\theta(\bar{u})} + g(\bar{u}),$$

ILEQG from Optimization Viewpoint

Regularized ILEQG (RegILEQG) with step-size γ_k

$$\bar{u}^{(k+1)} = \arg\min_{\bar{v}} m_{f_\theta}(\bar{u}+\bar{v}; \bar{u}) + \frac{1}{2\gamma_k}\|\bar{v}\|_2^2 \quad \textit{(LEQG by dyn. prog.)}$$

$\rightarrow$ how to choose γ_k ?

Surrogate Risk-Sensitive Cost

$$\hat{f}_\theta(\bar{u}) = \underbrace{\frac{1}{\theta}\log \mathbb{E}_{\bar{w}} \exp[\theta h(\tilde{x}(\bar{u}) + \nabla\tilde{x}(\bar{u})^\top \bar{w})]}_{\hat{\eta}_\theta(\bar{u})} + g(\bar{u}),$$

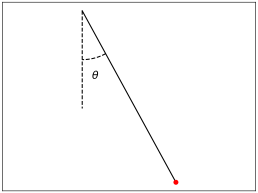
Theoretical Consequences (Roulet et al. 2019)

For h, g quadratics, ϕ_t bounded, Lipschitz, smooth

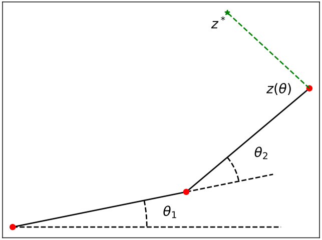
1. Show that RegILEQG minimizes $\hat{f}_\theta(\bar{u})$
2. $\hat{\eta}_\theta(\bar{u})$ can be computed analytically $\rightarrow$ access to line-search
3. Get necessary condition for θ (o.w. LEQG steps not defined)
4. Prove convergence to a near-stationary point of $\hat{f}_\theta$ for small γ_k

Numerical Illustrations

Settings



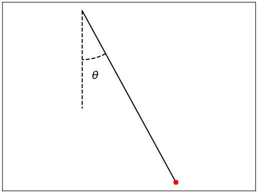
(a) Pendulum param. by θ.



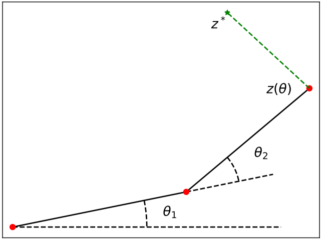
(b) Two-link arm param. by θ_1, θ_2.

Numerical Illustrations

Settings



(a) Pendulum param. by θ.



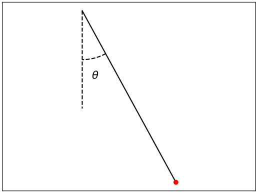
(b) Two-link arm param. by θ_1, θ_2.

Discretized Dynamics from $\ddot{z}(t) = f(z(t), \dot{z}(t), u(t))$

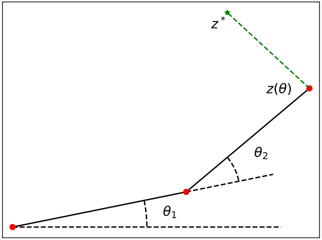
$$\begin{aligned} z_{t+1} &= z_t + \delta \dot{z}_t \\ \dot{z}_{t+1} &= \dot{z}_t + \delta f(z_t, \dot{z}_t, u_t), \end{aligned} \rightarrow (z_t, \dot{z}_t) = \phi_t((z_t, \dot{z}_t), u_t)$$

Numerical Illustrations

Settings



(a) Pendulum param. by θ.



(b) Two-link arm param. by θ_1, θ_2.

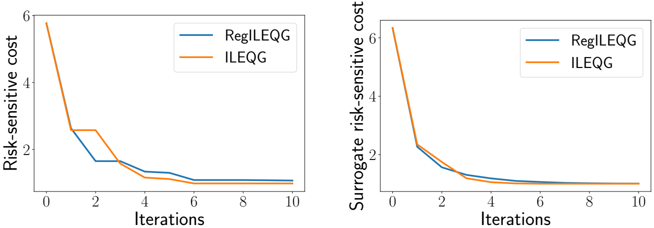
Discretized Dynamics from $\ddot{z}(t) = f(z(t), \dot{z}(t), u(t))$

$$\begin{aligned} z_{t+1} &= z_t + \delta \dot{z}_t \\ \dot{z}_{t+1} &= \dot{z}_t + \delta f(z_t, \dot{z}_t, u_t), \end{aligned} \rightarrow (z_t, \dot{z}_t) = \phi_t((z_t, \dot{z}_t), u_t)$$

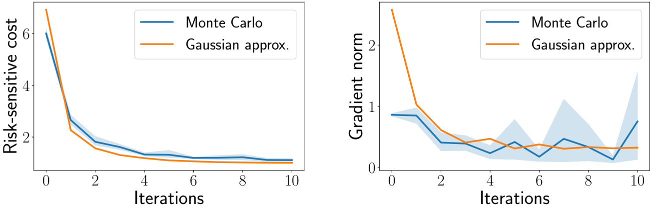
Robustness Test

$$\begin{aligned} z_{t+1} &= z_t + \delta \dot{z}_t \\ \dot{z}_{t+1} &= \dot{z}_t + \delta f(z_t, \dot{z}_t, u_t + \rho \mathbf{1}(t = t_w)), \end{aligned}$$

Numerical Illustrations

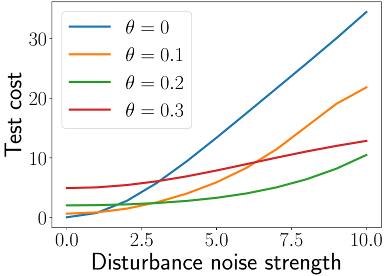


Convergence of ILEQG, RegILEQG

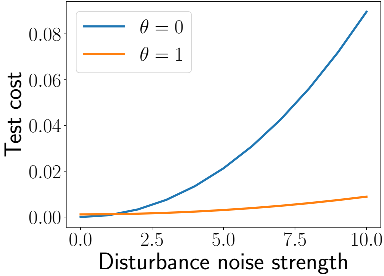


Approximations by surrogate risk-sensitive cost vs Monte-Carlo

Numerical Illustrations



(a) Pendulum.



(b) Two-link arm.

Robustness of controllers against disturbance noise.

Code available at https://github.com/vroulet/ilqc

Conclusion

Outcomes

1. Corrected ILEQG by adding proximal term
2. Clarified line-searches using the surrogate risk-sensitive cost
3. Provided a convergence rate
4. Provided code for testing the framework

Thank you ! Questions ?

Farshidian, F. & Buchli, J. (2015), ‘Risk sensitive, nonlinear optimal control: Iterative linear exponential-quadratic optimal control with Gaussian noise’, *arXiv preprint arXiv:1512.07173* .

Roulet, V., Fazel, M., Srinivasa, S. & Harchaoui, Z. (2019), ‘On the convergence to stationary points of the iterative linear exponential quadratic gaussian algorithm’, *arXiv preprint* .

Whittle, P. (1981), ‘Risk-sensitive linear quadratic Gaussian control’, *Advances in Applied Probability* **13**(4), 764–777.