General Information

- Thesis supervisor would be Dr. Venkatraman Renganathan, who is a post-doctoral fellow at the Department of Automatic Control, Lund University.
- Students who wish to do a Masters thesis on any one of the topics given below are strongly encouraged to contact either through email `venkat@control.lth.se` or they can directly come and visit office room number 3N:17 at the Department of Automatic Control, Kemicentrum Building, Lund University.
- The thesis topics are expected to be an exciting research opportunity for students who wish to pursue a future career either in academia (pursue PhD) or industry.
- All of the below thesis projects can be initiated anytime starting from January 2022 and it is estimated that they will be one semester long (3-4 months).
- Towards the end of the thesis presentation, student will be strongly expected to submit their work at a conference proceeding and if possible to a journal venue.

Prerequisites

Should be enrolled as M.Sc. student at Lund University; Interest, enthusiasm and willingness to learn about the research topic is expected; Expect the student to have a good grasp of basic control theory, basic probability theory and linear algebra. Student should be comfortable coding using any of the following languages MATLAB/Python/Julia/C++.

Masters Thesis Ideas

There are totally 3 different thesis project ideas. One-page short description for each of the master thesis ideas are given in this document starting from the next page.

1. Mean Square Compensation of a Stochastic LTI System With Multiplicative Noise Via Wasserstein Kalman Filter
2. Robot Path Planning Via Deep Reinforcement Learning
3. Spatio-temporal Risk Allocation for Path Planning

1. Mean Square Compensation of a Stochastic LTI System With Multiplicative Noise Via Wasserstein Kalman Filter

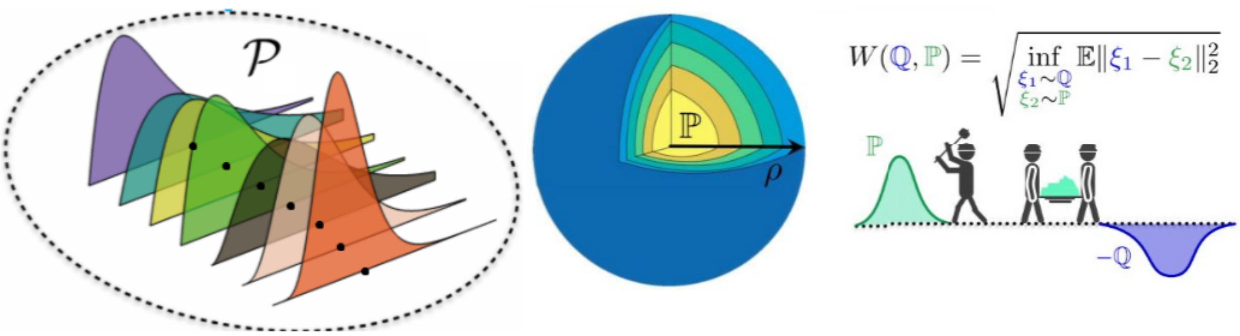
ABSTRACT: Multiplicative noise framework for capturing model uncertainty offers several compelling advantages over the conventional additive noise models. Given a nominal LTI system model, noise multiplies with state, input, & output in this setting and such problem formulation is getting increasingly common nowadays to model networked systems with noisy communication channels, power networks with large penetration of intermittent renewables etc. The challenging part of this problem setting is that the distribution of the system states at any future time t is not guaranteed to be Gaussian even though we start with a LTI system under Gaussian noise at time zero. Hence, the separation principle advocated by Kalman does not apply (we cannot optimally estimate the system state and optimally control it separately like LQG) and we need a different filter for estimation to hedge against the model risk. The main idea of this research will be to find a pair of controller and estimator gain matrices (K, L) that can mean-square compensate [1] a LTI system under such setting. Specifically, we will leverage the estimation technique developed by authors in [3] and additionally design a controller gain such that the pair (K, L) mean-square compensates the whole system.

Problem Statement

Given the following stochastic LTI system

$$
\begin{aligned}
x_{k+1} &= A_k x_k + B_k u_k + w_k, \quad w_k \sim \mathbb{P}_w \in \mathcal{P}^w := \left\{\mathbb{P}_w \mid \mathbb{E}[w_k] = 0, \mathbb{E}[w_k w_k^\top] = \Sigma_w\right\}, \\
y_k &= C_k x_k + v_k, \quad v_k \sim \mathbb{P}_v \in \mathcal{P}^v := \left\{\mathbb{P}_v \mid \mathbb{E}[v_k] = 0, \mathbb{E}[v_k v_k^\top] = \Sigma_v\right\}, \\
A_k &= \left(\bar{A} + \hat{A}_k\right), \; B_k = \left(\bar{B} + \hat{B}_k\right), \; C_k = \left(\bar{C} + \hat{C}_k\right), \\
\hat{A}_k &= \sum_{i=1}^{n_a} \gamma_{ki} \mathcal{A}_i, \; \hat{B}_k = \sum_{j=1}^{n_b} \delta_{kj} \mathcal{B}_j, \; \hat{C}_k = \sum_{l=1}^{n_c} \kappa_{kl} \mathcal{C}_l.
\end{aligned}
$$

where $\bar{A}$, $\bar{B}$, $\bar{C}$ denote the nominal dynamics, control, and output matrices respectively. The multiplicative noise terms are modeled by the i.i.d. across time (white), zero-mean, mutually independent scalar random variables γ_{ki}, δ_{kj}, κ_{kl}, which have variances $\sigma^2_{a,i}$, $\sigma^2_{b,j}$, $\sigma^2_{c,l}$ for $i \in [1 : n_a]$, $j \in [1 : n_b]$, $l \in [1 : n_c]$ respectively with $n_a, n_b, n_c \in \mathbb{Z}_{>0}$. The pattern matrices $\mathcal{A}_i \in \mathbb{R}^{n\times n}$, $\mathcal{B}_j \in \mathbb{R}^{n\times m}$, and $\mathcal{C}_l \in \mathbb{R}^{p\times n}$ specify how each scalar noise term affects the system matrices. Design the pair (K_t, L_t) for all time steps $t \geq 0$ such that they mean-square compensate the above system.



2. Robot Path Planning Via Deep Reinforcement Learning

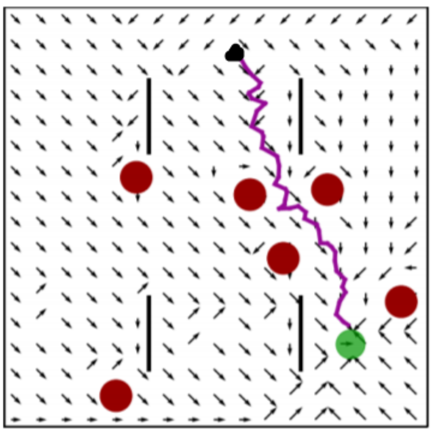
ABSTRACT: The main idea of this project is to find a path in real time for a robot operating in a known environment with both static and dynamic obstacles from the given initial position to the goal position. The resulting path should have provably least risk of collision with any obstacles. The robot is assumed to be a discrete-time LTI system but with an additive process noise belonging to a moment based ambiguity set and the robot is controlled with a finite set of inputs. A MDP formulation as in [2] will be adopted and a solution based on Distributionally Robust Reinforcement Learning framework as in [4] will be developed for the problem. Numerical experiments will be performed for the proposed method using Deep reinforcement learning algorithm. This approach is supposed to return a sub-optimal path at the expense of very less computational time compared to the significantly slower and optimal motion planning algorithms like RRT^*.

Problem Statement

Given an initial state $x_0 \in \mathcal{X} \subset \mathbb{R}^n$ and a set of final goal locations $\mathcal{X}_{goal} \subset \mathcal{X}$, find a measurable control policy $\pi = [\pi_0, \dots \pi_{T-1}]$ with $u_t = \pi_t(x_t)$ that maximizes the finite-horizon expected and discounted reward function subject to constraints:

$$
\begin{aligned}
\underset{\pi}{\text{maximize}} \quad & \mathbb{E}\left[\sum_{n=t}^{N-1} \gamma^{n-t} r(x_n, u_n, x_{n+1})\right] \\
\text{subject to} \quad & x_{t+1} = Ax_t + Bu_t + w_t, \\
& x_0 \sim \mathbb{P}_{x_0} \in \mathcal{P}^x := \{\mathbb{P}_{x_0} \mid \mathbb{E}[x_0] = \bar{x}_0, \mathbb{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top] = \Sigma_{x_0}\}, \\
& w_t \sim \mathbb{P}_w \in \mathcal{P}^w := \{\mathbb{P}_w \mid \mathbb{E}[w_t] = 0, \mathbb{E}[w_t w_t^\top] = \Sigma_w\}, \\
& u_t \in \mathcal{U}, \\
& \mathcal{X}_t^{\texttt{free}} = \mathcal{X} \setminus \bigcup_{i \in \mathcal{B}} \underbrace{\{x_t \mid A_i x_t \le b_{it}\}}_{\mathcal{O}_{it}}, \\
& \sup_{\mathbb{P}_{x_t} \in \mathcal{P}^x} \mathbb{P}_{x_t}(x_t \notin \mathcal{X}_t^{\texttt{free}}) \le \alpha_t, \quad \forall t \ge 0,
\end{aligned}
\tag{1}
$$

where $\mathcal{P}^x$ is an ambiguity set of marginal state distributions and $\alpha_t \in (0, 0.5]$ is a stage risk budget parameter such that $\sum_{t=0}^{T} \alpha_t \le \alpha$ with α being the user-prescribed total risk budget. The task for the robot is to learn a control policy such that risk of obstacle collision is minimum.



3. Spatio-temporal Risk Allocation for Path Planning

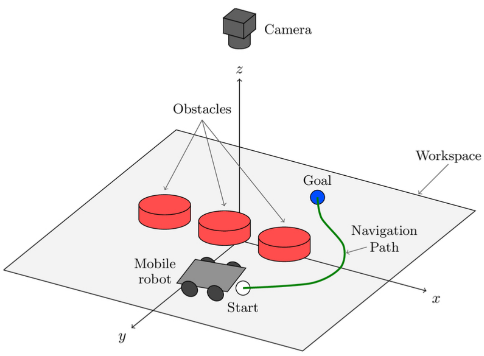
ABSTRACT: The main idea of this project is to find a path in real time for a robot operating in an environment with obstacles from the given initial position to the goal position. The resulting path should be constructed incrementally such that given a total risk budget for avoiding the obstacles at all time steps, the budget should be allocated for all obstacles across the entire time period in a spatio-temporal fashion. For instance, all time steps can have equal share of the risk budget and further at each time step, all obstacles (located at different locations) can have an equal share of the per-time step risk budget. However, such a simpler spatio-temporal risk allocation is known to result in both over-conservative and computationally expensive path planning procedure. The challenge is to allocate the risk in a non-uniform way by taking a spatial-temporal (space & time based) approach so that it results in minimally conservative path with reduced computational expense.

Problem Statement

Given an initial state $x_0 \in \mathcal{X} \subset \mathbb{R}^n$, a set of obstacles $\mathcal{B}$ and a set of final goal locations $\mathcal{X}_{goal} \subset \mathcal{X}$, find a measurable control policy $\pi = [\pi_0, \ldots \pi_{T-1}]$ with $u_t = \pi_t(x_t)$ that minimizes the finite-horizon expected cost function subject to constraints:

$$
\begin{aligned}
\underset{\pi}{\text{minimize}} \quad & \mathbb{E}\left[J_N(x_N) + \sum_{n=t}^{N-1} J_n(x_n, u_n) \right] \\
\text{subject to} \quad & x_{t+1} = A x_t + B u_t + w_t, \\
& x_0 \sim \mathbb{P}_{x_0} \in \mathcal{P}^x := \{\mathbb{P}_{x_0} \mid \mathbb{E}[x_0] = \bar{x}_0, \mathbb{E}[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top] = \Sigma_{x_0}\}, \\
& w_t \sim \mathbb{P}_w \in \mathcal{P}^w := \{\mathbb{P}_w \mid \mathbb{E}[w_t] = 0, \mathbb{E}[w_t w_t^\top] = \Sigma_w\}, \\
& u_t \in \mathcal{U}, \\
& \mathcal{X}_t^{\texttt{free}} = \mathcal{X} \setminus \bigcup_{i \in \mathcal{B}} \underbrace{\{x_t \mid A_i x_t \le b_{it}\}}_{\mathcal{O}_{it}}, \\
& \sup_{\mathbb{P}_{x_t} \in \mathcal{P}^x} \mathbb{P}_{x_t}(x_t \notin \mathcal{X}_t^{\texttt{free}}) \le \alpha_t, \quad \forall t \ge 0,
\end{aligned} \tag{2}
$$

where $\alpha_n \in (0, 0.5]$ is a stage risk budget parameter at time n such that $\sum_{n=t}^{N} \alpha_n \le \alpha$ with α being the user-prescribed total risk budget. The task is to satisfy α by allocating the risk $\forall i \in \mathcal{B}$ in a non-uniform way $\forall n$ and use it design a control policy to construct a safe path.



References

[1] Willem L De Koning. “Compensatability and optimal compensation of systems with white parameters”. In: *IEEE Transactions on Automatic Control* 37.5 (1992), pp. 579–588.

[2] Jeevan Raajan et al. “Real Time Path Planning of Robot using Deep Reinforcement Learning”. In: *IFAC-PapersOnLine* 53.2 (2020), pp. 15602–15607.

[3] Soroosh Shafieezadeh-Abadeh et al. “Wasserstein distributionally robust Kalman filtering”. In: *arXiv preprint arXiv:1809.08830* (2018).

[4] Elena Smirnova, Elvis Dohmatob, and Jérémie Mary. “Distributionally robust reinforcement learning”. In: *arXiv preprint arXiv:1902.08708* (2019).