Model Predictive Control

Lecturer: Pieter Abbeel *Scribe: Andrew Wan*

1 Lecture outline

- MPC
- SLAM
- Linear Representation of Bellman Backups
- Hierarchical RL

2 Model Predictive Control

Previously: Interested in:

$$\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t, u_t) | \pi\right]$$

This is typically computationally expensive.
We simplify by:

1. Fixing the initial state
2. Using deterministic $\max_{u_0, u_1, \ldots}$ rather than $\max_{\pi}$
3. Using a finite horizon H

New formulation:

$$\max_{u_0, \ldots, u_H} \sum_{t=0}^{H} R(s_t, u_t) \tag{1}$$

$$\begin{aligned} s.t. \forall t \in 0 \ldots H: \; s_{t+1} &= f_t(s_t, u_t) \\ s_0 &= s \\ u_t &\in U_t \end{aligned}$$

2.1 Nonlinear MPC

For $t = 0, 1 \ldots \infty$
  solve (1) with $s_0 = s_t$
  execute u_0

2.2 Linear MPC

Replace (1) by:

$$\max_{u_0,\ldots,u_H} \sum_{t=0}^{H} R_t(s_t, u_t)$$
$$s.t.: \; s_{t+1} = A_t s_t + B_t u_t + b_t$$
$$s_0 = s$$

$u_t \in U_t \;\rightsquigarrow U_t$ convex sets

3 SLAM - Simultaneous Localization and Mapping

Notable papers:
Durant Whyte & John Leonard '91
Smith, Self & Cheeseman '86

3.1 Typical Setup

Wheeled robot with odometry (measurements of velocity, angular rate)
SICK laser range finder (Hokuyo/Velodyne: .25 degrees, ±5min. accuracy)
Robot is released into an unknown environment.

3.2 Process

Using an Extended Kalman Filter, loop:
Odometry change
EKF odometry update
EKF re-observation
EKF new observation
(Laser scan reading → landmark extraction → data association → EKF re-observation/new observation)
State of the EKF:
state of the robot (x, y, θ)
position of extracted landmarks $\{(x_i, y_i)\}$
Landmark extraction from point cloud:
RANSAC: Find lines through data, compare to landmark models
"spikes": Check for angles within the cloud

3.3 Data Association Problem

When we see a new landmark, we store it in a database of landmarks. How do we know whether a landmark is newer than a previous seen one?

EKF stores the position of previously seen landmarks with uncertainty estimates. Use EKF as described before, with observation model:

$$range = \sqrt{(x^{(i)} - x)^2 + (y^{(i)} - y)^2} + v_r$$

$$bearing = \tan^{-1}\left(\frac{y^{(i)} - y}{x^{(i)} - x}\right) - \theta + v_\theta$$

$$s.t.:\ v_r, v_\theta = noise$$

4 Linear Representation of Bellman Backups (Emo Todorov)

$$v(x) = \min_a l(x,a) + E_{y\sim(P.|x,a)}\left[v(y)\right]$$

$$s.t. : l(x,a) = \text{cost of action a}$$

$$E_{y\sim(P.|x,a)}\left[v(y)\right] = \sum_y P(y|x,a)\cdot v(y)$$

$$u(a) = \text{Probability of taking action a}$$

$$v(x) = \min_u \sum_a u(a)\cdot l(x,a) + \sum_y P(y|x,a)\cdot v(y)$$

More generally:

$$u(s) = P(\text{visiting states s at time t+1})$$

$$v(x) = \min_u l(x,u) + \sum_y P(y|x,u)v(y)$$

$$l(x,u) = q(x) \rightarrow q(x) + \sum_y u(y)\log\frac{u(y)}{\bar{P}(y|x)}$$

$$\sum_y u(y)\log\frac{u(y)}{\bar{P}(y|x)} = KL(u\|\bar{P}(y|x)) = \text{"Natural Dynamics"}$$

$$V_u = 0 \rightarrow u(y) = \frac{e^{-v(y)}\cdot\bar{P}(y|x)}{\sum_z e^{-v(z)}\bar{P}(z|x)}$$

(closed form solution to minimization)

substituting this back:

$$v(x) = q(x) - \log\sum_y \bar{P}(y|x)e^{-v(y)}$$

introduce $\Psi(x) = e^{-v(x)}$:

$$v(x) = -\log\Psi(x)$$

$$\text{so: } -\log\Psi(x) = q(x) - \log\sum_y \bar{P}(y|x)\cdot\Psi(y)$$

$$\Psi(x) = e^{-q(x)}\cdot\sum_y \bar{P}(y|x)\cdot\Psi(y)$$