Last Time:

- Friction
- Iterative Learning Control

Today:

- Finish ILC
- Stochastic Optimal Control
- LQG

---

* ILC Algorithm:

- Given a nominal trajectory $\bar{x}$, $\bar{u}$ computed with a nominal model, use rollouts on the real system to find a $\Delta u$ correction to $\bar{u}$ such that the real system tracks $\bar{x}$ as closely as possible.

do:

note: can be different from $\bar{u}$ due to tracking controller

$$x_{1:N}, u_{1:N-1} \leftarrow \underbrace{\text{rollout}(\bar{x}_0, \bar{u})}_{\text{on real system possibly with tracking controller}}$$

This is a QP

$$\Delta x, \Delta u \leftarrow \left\{ \begin{array}{l} \text{argmin } J(\Delta x, \Delta u) \\ \text{s.t. } \Delta x_{n+1} = A_n \Delta x_n + B_n \Delta u_n \\ \qquad u_{min} \leq u_n \leq u_{max} \end{array} \right.$$

$$\bar{u} \leftarrow \bar{u} + \Delta u$$

while $\underbrace{\|x_N - \bar{x}_N\|}_{\text{(many options)}} \geq \text{tol}$

* Why Should ILC Work?

- We've already seen approximations in Newton's method: Gauss-Newton + regularization

- In general, these are called "inexact" and/or "quasi-Newton" methods. Many variants (BFGS, Newton-CG), with well-developed theory.

- For a generic root-finding problem:

$$f(x+\Delta x) \approx \underbrace{f(x) + \frac{\partial f}{\partial x}\Delta x = 0}_{\text{exact Newton step}}$$

- As long as $\Delta x$ satisfies

$$\|f(x) + J\Delta x\| \leq \eta \|f(x)\| \text{ for some } \eta < 1,$$

an inexact Newton method will converge.

- Convergence is slower than exact Newton

- This means we can use $J \approx \frac{\partial f}{\partial x}$ to compute $\Delta x$

* Example:

- Cartpole swing-up with model mismatch

# Stochastic Control:

- So far we have assumed we know the system's state perfectly.
- What happens when all we have are noisy measurements of quantities related to the state?

$$\underset{\text{measurement}}{y} = \underbrace{g(x)}_{\text{"measurement model"}}$$

| Deterministic | | Stochastic |
|---|---|---|
| $x$ | $\longrightarrow$ | $\underbrace{p(x \mid y)}_{\text{PDF of the state conditioned on the measurements}}$ |

* Stochastic Optimal Control Problem

$$\min_u E[J(x,u)]$$

- In principle, can solve with DP
- Very hard in general

# * LQG

- Special case we can solve in closed form

  **L**inear dynamics

  **Q**uadratic cost

  **G**aussian noise

- Dynamics

$$x_{n+1} = Ax_n + Bu_n + \underbrace{w_n}_{\text{"process noise"}}$$

$$y_n = Cx_n + \underbrace{v_n}_{\text{"measurement noise"}}$$

$$w_n \sim N(0, W) \quad , \quad v_n \sim N(0, V)$$

"drawn from"

Normal distribution (a.k.a. Gaussian)

mean

covariance

* Multivariate Gassian

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det(P)}} \exp\left(-\frac{1}{2}(x-\mu)^T S^{-1}(x-\mu)\right)$$

mean : $\mu = \hat{x} = E[x] \in \mathbb{R}^n$

covariance : $S = E[(x-\mu)(x-\mu)^T] \in S^n_{++}$

$$E[f(x)] = \int_{\text{(all space)}} f(x)\, p(x)\, dx$$

"Uncorrelated" $\Rightarrow$ $E[(x-\hat{x})(y-\hat{y})^T] = 0$

---

- Cost Function

$$J = E\left[\frac{1}{2}x_N^T Q_N x_N + \frac{1}{2}\sum_{n=1}^{N-1}(x_n^T Q x_n + u_n^T R u_n)\right]$$

- D.P. Recursion

$$V_N(x) = E[x_N^T Q_N x_N] = E[x_N^T P_N x_N]$$

$$V_{N-1}(x) = \min_u E\Big[\frac{1}{2}x_{N-1}^T Q x_{N-1} + \frac{1}{2}u_{N-1}^T R u_{N-1} + \frac{1}{2}(Ax_{N-1} + Bu_{N-1} + w_{N-1})^T P_N (Ax_{N-1} + Bu_{N-1} + w_{N-1})\Big]$$

$$= \min_u \underbrace{E\Big[\frac{1}{2}x_{N-1}^T Q x_{N-1} + \frac{1}{2}u_{N-1}^T R u_{N-1} + (Ax_{N-1} + Bu_{N-1})^T P_N \cdots (Ax_{N-1} + Bu_{N-1})\Big]}_{\text{Standard LQR}}$$

$$+ E\Big[\underbrace{(A\cancel{x_{N-1}} + \cancel{Bu_{N-1}})^T P_N w_{N-1} + w_{N-1}^T P_N (\cancel{Ax_{N-1}} + Bu_{N-1})}_{0} \; \underbrace{+ \; w_{N-1}^T P_N w_{N-1}}_{\text{constant}}\Big]$$

Noise terms

Noise sample drawn at time k has nothing to do with state (or control) of time k. $x_n$ depends on $w_{n-1}$ (and all past $w$) but not on $w_n$ or any future $w$.

$\Rightarrow$ Noise terms have no impact on the controller design!

(you just get a higher cost)

* "Certainty - Equivalence Priciple"

 - The optimal LQG controller is just an LQR controller with $x$ replaced by $E[x]$

* "Separation Principle"

 - For LQG we can design an optimal feedback controller and an optimal estimator seperately and then hook them togethe. The resulting feedback policy is optimal.

* Neither of these holds in general, but are still frequently used in practice to design sub-optimal policies.