Optimal and Learning Control
for
Autonomous Robots
Lecture 5

ADRL

Jonas Buchli
Agile & Dexterous Robotics Lab

FNSNF

ETH Zürich

ETH zürich

Help us to improve life at ETH!

→**Look for this email, with link: *studierendenbefragung@ethz.ch***

→**Fill out the online questionnaire.**

→**This is about *your* happiness at ETH!**

Thank you for your help.

The survey will run from 16 March to 6 April 2015.
For detailed information see www.ethz.ch/studentsurvey

Erratum Script

p14 $$\frac{dV^*}{dt} = V_t^* + V_{\mathbf{x}}^{*T}\mathbf{f} + \frac{1}{2}\mathrm{Tr}\left[V_{\mathbf{xx}}^* E[(\mathbf{f}+\mathbf{Bw})(\mathbf{f}+\mathbf{Bw})^T]\Delta t\right]. \qquad (1.55)$$

p28 $$-\mathbf{x}^T\dot{\mathbf{S}}(t)\mathbf{x} = \min_{u\in U}\{\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} + 2\mathbf{u}^T\mathbf{P}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{B}\mathbf{u}\}. \qquad (1.105)$$

Class logistics

Exercise groups

Sign up for the exercises in groups of 2:

https://ethz.doodle.com/c27fqrggtqth2x57

Please avoid single-member groups!

Deadline for inscription: March 20th, 18h

Exercises

Exercises

- 3 programming exercises
- starting L5, 8, 12
- exercises graded pass/fail
- grade boost for passed exercises
  - ➡Ex1: 0.1, Ex 2: 0.05, Ex 3: 0.1
- solutions will be available at end of semester
- topics of exercises will be used for exam

Exercise 1

Today - 16:15

- Submission:
  - Code must be submitted through website form
  - NO EMAIL SUBMISSION!
  - submit by Wed, 15.4.2015
  - USE OFFICE HOURS FOR QUESTIONS!

Office hours:
Thu, 17:30-18:30
Room: ML J37.1

- Interviews:
  - Interviews on Friday, 17.4.2015, all day
  - 10 min session/group
  - explain submitted code and answers
  - pass/fail grade given
  - Doodle link for sign up for interview will be given

Lecture 5 Goals

★ Derivation of ILQC (Part II)
★ LQR

L4 Recap

Solve optimal control problem

$$V^*(n,\mathbf{x}) = \min_{\mathbf{u}_n} \left[L_n(\mathbf{x},\mathbf{u}_n) + \alpha V^*\left(n+1, \mathbf{f}_n\left(\mathbf{x},\mathbf{u}_n\right)\right)\right]$$

1. Principle of optimality: Bellman / HJB Equation
2. Make some assumptions
3. Minimize RHS of Equation
4. ... yields conditions for optimal control
5. substitute back to solve for remaining quantities

Sequential Quadratic Programming (SQP)

'Unsolvable Nonlinear Program':

$$\min_x f(x) \qquad x \in \mathbb{R}^n$$

$$s.t. \qquad f_j(x) \leq 0, \qquad j = 1, \ldots, N$$

$$h_j(x) = 0, \qquad j = 1, \ldots, N$$

Idea: Approximate nonlinear program by a QP, solve iteratively

Sequential Quadratic Programming (SQP)

Idea: Approximate nonlinear program by a QP, solve iteratively

- Initial guess $\tilde{x}_0$
- Approximate $f(x)$ at $\tilde{x}_0$ by 2nd order Taylor series expansion

$$f(x) \approx f(\tilde{x}_0) + (x - \tilde{x}_0)^T \nabla f(\tilde{x}_0) + \frac{1}{2}(x - \tilde{x}_0)^T \nabla^2 f(\tilde{x}_0)(x - \tilde{x}_0)$$

square in x

$$f_j(x) \approx f_j(\tilde{x}_0) + (x - \tilde{x}_0)^T \nabla f_j(\tilde{x}_0)$$

$$h_j(x) \approx h_j(\tilde{x}_0) + (x - \tilde{x}_0)^T \nabla h_j(\tilde{x}_0)$$

constraints: first order

- yields new approximative solution $\tilde{x}_1$
- repeat

$$\lim_{i \to \infty} \tilde{x}_i = x^*$$

if problem convex

Sequential Linear Quadratic Control - SLQ

$$\min_{\mu} \left[\Phi(\mathbf{x}(N)) + \sum_{n=0}^{N-1} L_n(\mathbf{x}(n), \mathbf{u}(n)) \right]$$

$$s.t. \qquad \mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{u}(n)) \qquad \mathbf{x}(0) = \mathbf{x}_0$$

$$\mathbf{u}(n, \mathbf{x}) = \mu(n, \mathbf{x})$$

Idea: Fit simplified subproblem to original problem, solve iteratively

Class of algorithms

value function → optimization target → quadratic

system dynamics → constraints → linear

SQP vs SLQ

SQP

1. Initial guess for parameter
2. Solve sub problem: Approximate original problem with a linear-quadratic problem
3. yields new approximative solution
4. repeat

SLQ

1. Initial guess for policy
2. Solve sub problem: Approximate value function with a linear-quadratic
3. yields new approximative policy
4. repeat

SLQ subproblem in a nutshell

2.1 Forward pass:

integrate to get a state (and controls) trajectory

2.2 Backward pass

Solve simplified optimal control problem around state and control trajectory

3. Adjust guess for optimal control

choice of: approximation, solver ⇒different SLQ algorithms

(examples: DDP, iLQG, ILQC)

ILQC

Overview of derivation

Linearize system dynamics
Quadratize cost

Compute value function
Compute optimal control
Solve for Riccati like equation
Solve Riccati like equation

Linearization of system dynamics

$$\bar{\mathbf{x}}_{n+1} + \delta\mathbf{x}_{n+1} = \mathbf{f}_n(\bar{\mathbf{x}}_n + \delta\mathbf{x}_n, \bar{\mathbf{u}}_n + \delta\mathbf{u}_n)$$

$$\approx \mathbf{f}_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n) + \frac{\partial \mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}}\delta\mathbf{x}_n + \frac{\partial \mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}}\delta\mathbf{u}_n$$

$$\bar{\mathbf{x}}_{n+1} = \mathbf{f}_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)$$

$$\delta\mathbf{x}_{n+1} \approx \mathbf{A}_n\delta\mathbf{x}_n + \mathbf{B}_n\delta\mathbf{u}_n$$

$$\mathbf{A}_n = \frac{\partial \mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}}$$ systems matrix

$$\mathbf{B}_n = \frac{\partial \mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}}$$ control gain matrix

$\mathbf{A}_n$ and $\mathbf{B}_n$ are independent of $\delta\mathbf{x}_n$ and $\delta\mathbf{u}_n$

$\mathbf{A}_n$ and $\mathbf{B}_n$ are time varying

nonlinear → linear, time variant

Quadratization of cost function

$$J = \Phi(\mathbf{x}_N) + \sum_{n=0}^{N-1} L_n(\mathbf{x}_n, \mathbf{u}_n)$$

$$J \approx q_N + \delta\mathbf{x}_N^T \mathbf{q}_N + \frac{1}{2}\delta\mathbf{x}_N^T \mathbf{Q}_N \delta\mathbf{x}_N$$

$$+ \sum_{n=0}^{N-1} \{q_n + \delta\mathbf{x}_n^T \mathbf{q}_n + \delta\mathbf{u}_n^T \mathbf{r}_n + \frac{1}{2}\delta\mathbf{x}_n^T \mathbf{Q}_n \delta\mathbf{x}_n + \frac{1}{2}\delta\mathbf{u}_n^T \mathbf{R}_n \delta\mathbf{u}_n + \delta\mathbf{u}_n^T \mathbf{P}_n \delta\mathbf{x}_n\}$$

Control costs

State costs

'Mixing terms'

$\forall n \in \{0, \cdots, N-1\}:$

$$q_n = L_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n), \qquad \mathbf{q}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}}, \qquad \mathbf{Q}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}^2}$$

$$\mathbf{P}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u} \partial \mathbf{x}}, \qquad \mathbf{r}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}}, \qquad \mathbf{R}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}^2}$$

$n = N:$

$$q_N = \Phi(\bar{\mathbf{x}}_n), \qquad \mathbf{q}_N = \frac{\partial \Phi(\bar{\mathbf{x}}_n)}{\partial \mathbf{x}}, \qquad \mathbf{Q}_N = \frac{\partial^2 \Phi(\bar{\mathbf{x}}_n)}{\partial \mathbf{x}^2}$$

Note that all derivatives w.r.t. $\mathbf{u}$ are zero for the terminal time-step N

Q, R, P are given through definition of cost!

Compute Value function

(2) Quadratic Ansatz for Value function

Ansatz: Quadratic Value function

$$V^*(n+1, \delta \mathbf{x}_{n+1}) = s_{n+1} + \delta \mathbf{x}_{n+1}^T \mathbf{s}_{n+1} + \frac{1}{2} \delta \mathbf{x}_{n+1}^T \mathbf{S}_{n+1} \delta \mathbf{x}_{n+1}$$

$\mathbf{S}_n, \mathbf{s}_n, s_n$ are unknown

and will have to be computed... later!

relabel terms depending on controls

$$\mathbf{g}_n \triangleq \mathbf{r}_n + \mathbf{B}_n^T \mathbf{s}_{n+1}$$
$$\mathbf{G}_n \triangleq \mathbf{P}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{A}_n$$
$$\mathbf{H}_n \triangleq \mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n$$

$$V^*(n, \delta\mathbf{x}_n) = \min_{\mathbf{u}_n} \Big[q_n + s_{n+1} + \delta\mathbf{x}_n^T(\mathbf{q}_n + \mathbf{A}_n^T \mathbf{s}_{n+1}) + \frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1}\mathbf{A}_n)\delta\mathbf{x}_n + \delta\mathbf{u}_n^T(\mathbf{g}_n + \mathbf{G}_n\delta\mathbf{x}_n) + \frac{1}{2}\delta\mathbf{u}_n^T\mathbf{H}_n\delta\mathbf{u}_n \Big] \quad (1.80)$$

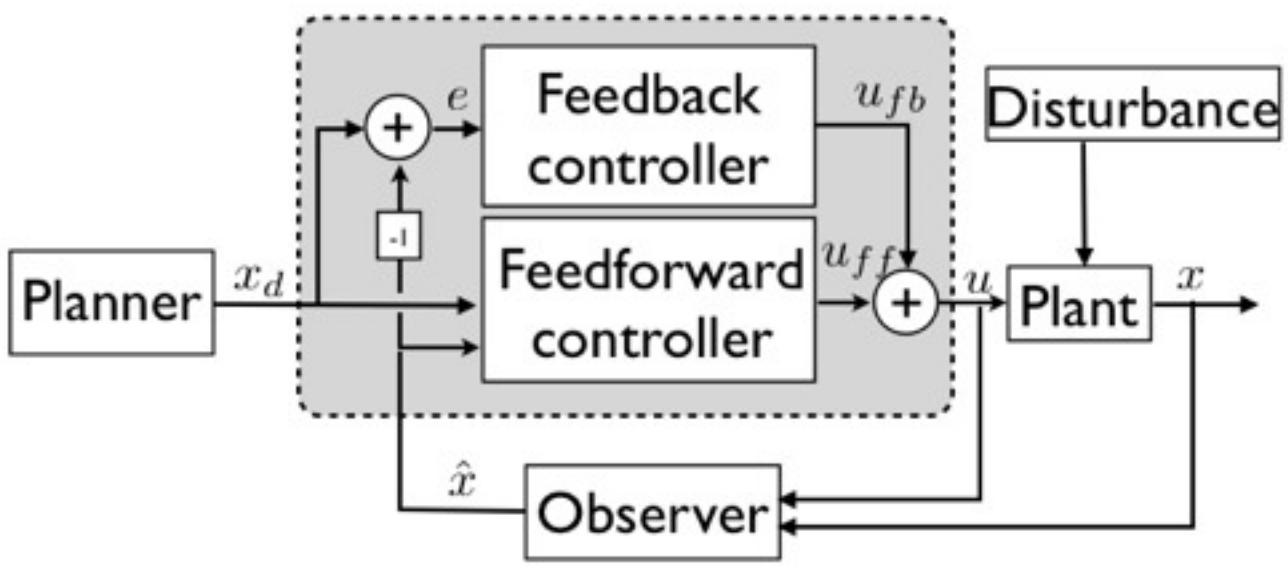
Optimal control: FF/FB

$$\delta \mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

feed-forward term $\delta \mathbf{u}_n^{ff} = -\mathbf{H}_n^{-1}\mathbf{g}_n$

feedback term $\mathbf{K}_n\delta\mathbf{x}_n$ feedback gain matrix $\mathbf{K}_n := -\mathbf{H}_n^{-1}\mathbf{G}_n$

$$\delta \mathbf{u}_n = \delta \mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n$$



A D R L

ETH Zürich

check 'units'

$$\delta \mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

$$\mathbf{H}_n^{-1}\mathbf{g}_n \qquad\qquad \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

$$(\mathbf{R}\ldots)^{-1} \quad (\mathbf{r}\ldots) \qquad (\mathbf{R}\ldots)^{-1} \quad (\mathbf{P}\ldots) \quad \delta\mathbf{x}_n$$

$$\frac{\partial \mathbf{u}^2}{\partial^2 L} \quad \frac{\partial L}{\partial \mathbf{u}} \qquad \frac{\partial \mathbf{u}^2}{\partial^2 L} \quad \frac{\partial^2 L}{\partial \mathbf{u}\partial \mathbf{x}} \quad \delta\mathbf{x}_n$$

$$\partial \mathbf{u} \qquad\qquad \partial \mathbf{u}$$

ADRL

ETH Zürich

check 'units'

$$\delta \mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

$$\mathbf{H}_n^{-1}\mathbf{g}_n \qquad\qquad \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

$$(\mathbf{R}\ldots)^{-1} \quad (\mathbf{r}\ldots) \qquad\qquad (\mathbf{R}\ldots)^{-1} \quad (\mathbf{P}\ldots) \quad \delta\mathbf{x}_n$$

$$\frac{\partial \mathbf{u}^2}{\partial^2 L} \quad \frac{\partial L}{\partial \mathbf{u}} \qquad\qquad \frac{\partial \mathbf{u}^2}{\partial^2 L} \quad \frac{\partial^2 L}{\partial \mathbf{u}\,\partial \mathbf{x}} \quad \delta\mathbf{x}_n$$

$$\partial \mathbf{u} \qquad\qquad \partial \mathbf{u}$$

A D R L

ETH Zürich

$$\delta \mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

feed-forward term $\delta \mathbf{u}_n^{ff} = -\mathbf{H}_n^{-1}\mathbf{g}_n$

feedback gain matrix $\mathbf{K}_n := -\mathbf{H}_n^{-1}\mathbf{G}_n$

functions of unknown $\mathbf{S}_n, \mathbf{s}_n, s_n$

$$\mathbf{g}_n \triangleq \mathbf{r}_n + \mathbf{B}_n^T\mathbf{s}_{n+1}$$
$$\mathbf{G}_n \triangleq \mathbf{P}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{A}_n$$
$$\mathbf{H}_n \triangleq \mathbf{R}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{B}_n$$

Solving for $\mathbf{S}_n, \mathbf{s}_n, s_n$

A D R L

ETH Zürich

EOF Recap

L5

Solving for $\mathbf{S}_n, \mathbf{s}_n, s_n$

$$\delta \mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

feed-forward term $\delta \mathbf{u}_n^{ff} = -\mathbf{H}_n^{-1}\mathbf{g}_n$

feedback term $\mathbf{K}_n\delta\mathbf{x}_n$ feedback gain matrix $\mathbf{K}_n := -\mathbf{H}_n^{-1}\mathbf{G}_n$

replace $\delta \mathbf{u}_n = \delta \mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n$

plug into

$$V^*(n, \delta\mathbf{x}_n) = \min_{\mathbf{u}_n}\Big[q_n + s_{n+1} + \delta\mathbf{x}_n^T(\mathbf{q}_n + \mathbf{A}_n^T\mathbf{s}_{n+1}) + \frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)\delta\mathbf{x}_n + \delta\mathbf{u}_n^T(\mathbf{g}_n + \mathbf{G}_n\delta\mathbf{x}_n) + \frac{1}{2}\delta\mathbf{u}_n^T\mathbf{H}_n\delta\mathbf{u}_n\Big]$$

A D R L

ETH Zürich

$$V^*(n,\delta\mathbf{x}_n) = \min_{\mathbf{u}_n}\Big[q_n + s_{n+1} + \delta\mathbf{x}_n^T(\mathbf{q}_n + \mathbf{A}_n^T\mathbf{s}_{n+1})$$

$$+\frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)\delta\mathbf{x}_n + \delta\mathbf{u}_n^T(\mathbf{g}_n + \mathbf{G}_n\delta\mathbf{x}_n) + \frac{1}{2}\delta\mathbf{u}_n^T\mathbf{H}_n\delta\mathbf{u}_n\Big]$$

$$\delta\mathbf{u}_n = \delta\mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n$$

$$(\delta\mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n)^T(\mathbf{g}_n + \mathbf{G}_n\delta\mathbf{x}_n)$$
$$+\frac{1}{2}(\delta\mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n)^T\mathbf{H}_n(\delta\mathbf{u}_n^{ff} + \mathbf{K}_n\delta\mathbf{x}_n)$$

$$\delta\mathbf{u}_n^{ff^T}\mathbf{g}_n + \delta\mathbf{u}_n^{ff^T}\mathbf{G}_n\delta\mathbf{x}_n + \delta\mathbf{x}_n^T\mathbf{K}_n^T\mathbf{g}_n + \delta\mathbf{x}_n^T\mathbf{K}_n^T\mathbf{G}_n\delta\mathbf{x}_n$$
$$+\frac{1}{2}(\delta\mathbf{u}_n^{ff^T}\mathbf{H}_n\delta\mathbf{u}_n^{ff} + \delta\mathbf{u}_n^{ff^T}\mathbf{H}_n\mathbf{K}_n\delta\mathbf{x}_n + \delta\mathbf{x}_n^T\mathbf{K}_n^T\mathbf{H}_n\delta\mathbf{u}_n^{ff} + \delta\mathbf{x}_n^T\mathbf{K}_n^T\mathbf{H}_n\mathbf{K}_n\delta\mathbf{x}_n)$$

$$\delta\mathbf{u}^{ff^T}\mathbf{g} + \frac{1}{2}\delta\mathbf{u}^{ff^T}\mathbf{H}\delta\mathbf{u}^{ff} + \delta\mathbf{x}^T(\mathbf{G}^T\delta\mathbf{u}^{ff} + \mathbf{K}^T\mathbf{g} + \mathbf{K}^T\mathbf{H}\delta\mathbf{u}^{ff})$$
$$+\frac{1}{2}\delta\mathbf{x}^T(\mathbf{K}^T\mathbf{H}\mathbf{K} + \mathbf{K}^T\mathbf{G} + \mathbf{G}^T\mathbf{K})\delta\mathbf{x}$$

$$V^*(n, \delta\mathbf{x}_n) = \min_{\mathbf{u}_n} \Big[q_n + s_{n+1} + \delta\mathbf{x}_n^T(\mathbf{q}_n + \mathbf{A}_n^T\mathbf{s}_{n+1})$$
$$+ \frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)\delta\mathbf{x}_n + \delta\mathbf{u}_n^T(\mathbf{g}_n + \mathbf{G}_n\delta\mathbf{x}_n) + \frac{1}{2}\delta\mathbf{u}_n^T\mathbf{H}_n\delta\mathbf{u}_n \Big]$$

$$V^*(n, \delta\mathbf{x}_n) = s_n + \delta\mathbf{x}_n^T\mathbf{s}_n + \frac{1}{2}\delta\mathbf{x}_n^T\mathbf{S}_n\delta\mathbf{x}_n$$

Quadratic Ansatz

Optimal control

$$\delta\mathbf{u}^{ff^T}\mathbf{g} + \frac{1}{2}\delta\mathbf{u}^{ff^T}\mathbf{H}\delta\mathbf{u}^{ff} + \delta\mathbf{x}^T(\mathbf{G}^T\delta\mathbf{u}^{ff} + \mathbf{K}^T\mathbf{g} + \mathbf{K}^T\mathbf{H}\delta\mathbf{u}^{ff})$$
$$+ \frac{1}{2}\delta\mathbf{x}^T(\mathbf{K}^T\mathbf{H}\mathbf{K} + \mathbf{K}^T\mathbf{G} + \mathbf{G}^T\mathbf{K})\delta\mathbf{x}$$

$$s_n + \delta\mathbf{x}_n^T\mathbf{s}_n + \frac{1}{2}\delta\mathbf{x}_n^T\mathbf{S}_n\delta\mathbf{x}_n =$$
$$q_n + s_{n+1} + \delta\mathbf{x}_n^T(\mathbf{q}_n + \mathbf{A}_n^T\mathbf{s}_{n+1}) + \frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)\delta\mathbf{x}_n +$$
$$\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}^T}\mathbf{g}_n + \frac{1}{2}\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}^T}\mathbf{H}_n\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}} + \delta\mathbf{x}_n^T(\mathbf{G}_n^T\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}} + \mathbf{K}_n^T\mathbf{g}_n + \mathbf{K}_n^T\mathbf{H}_n\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}})$$
$$+ \frac{1}{2}\delta\mathbf{x}_n^T(\mathbf{K}_n^T\mathbf{H}_n\mathbf{K}_n + \mathbf{K}_n^T\mathbf{G}_n + \mathbf{G}_n^T\mathbf{K}_n)\delta\mathbf{x}_n$$

sort into terms in $\delta\mathbf{x}^a \quad a \in [0, 1, 2]$

$1, \quad \delta\mathbf{x}, \quad \delta\mathbf{x}^T\delta\mathbf{x}$

sort into terms in $1,\ \delta \mathbf{x},\ \delta \mathbf{x}^T \delta \mathbf{x}^T$

$$s_n + \delta \mathbf{x}_n^T \mathbf{s}_n + \frac{1}{2}\delta \mathbf{x}_n^T \mathbf{S}_n \delta \mathbf{x}_n =$$

$$q_n + s_{n+1} + \delta \mathbf{x}_n^T (\mathbf{q}_n + \mathbf{A}_n^T \mathbf{s}_{n+1}) + \frac{1}{2}\delta \mathbf{x}_n^T (\mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{A}_n)\delta \mathbf{x}_n +$$

$$\delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}\,T} \mathbf{g}_n + \frac{1}{2}\delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}\,T} \mathbf{H}_n \delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}} + \delta \mathbf{x}_n^T (\mathbf{G}_n^T \delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}} + \mathbf{K}_n^T \mathbf{g}_n + \mathbf{K}_n^T \mathbf{H}_n \delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}})$$

$$+ \frac{1}{2}\delta \mathbf{x}_n^T (\mathbf{K}_n^T \mathbf{H}_n \mathbf{K}_n + \mathbf{K}_n^T \mathbf{G}_n + \mathbf{G}_n^T \mathbf{K}_n)\delta \mathbf{x}_n$$

$$n \in \{0, \cdots, N-1\}$$

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{A}_n + \mathbf{K}_n^T \mathbf{H}_n \mathbf{K}_n + \mathbf{K}_n^T \mathbf{G}_n + \mathbf{G}_n^T \mathbf{K}_n$$

$$\mathbf{s}_n = \mathbf{q}_n + \mathbf{A}_n^T \mathbf{s}_{n+1} + \mathbf{K}_n^T \mathbf{H}_n \delta \mathbf{u}_n^{ff} + \mathbf{K}_n^T \mathbf{g}_n + \mathbf{G}_n^T \delta \mathbf{u}_n^{ff}$$

$$s_n = q_n + s_{n+1} + \frac{1}{2}\delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}\,T} \mathbf{H}_n \delta \mathbf{u}_n^{ff} + \delta \mathbf{u}_n^{ff\,T} \mathbf{g}_n$$

$$\mathbf{S}_N = \mathbf{Q}_N, \qquad \mathbf{s}_N = \mathbf{q}_N, \qquad s_N = q_N$$

ADRL

ETH Zürich

note symmetry of S (if Q symmetric)!
S positive definite

$$n \in \{0, \cdots, N-1\}$$

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{A}_n + \mathbf{K}_n^T \mathbf{H}_n \mathbf{K}_n + \mathbf{K}_n^T \mathbf{G}_n + \mathbf{G}_n^T \mathbf{K}_n$$

$$\mathbf{s}_n = \mathbf{q}_n + \mathbf{A}_n^T \mathbf{s}_{n+1} + \mathbf{K}_n^T \mathbf{H}_n \delta \mathbf{u}_n^{ff} + \mathbf{K}_n^T \mathbf{g}_n + \mathbf{G}_n^T \delta \mathbf{u}_n^{ff}$$

$$s_n = q_n + s_{n+1} + \frac{1}{2} \delta \mathbf{u}_\mathbf{n}^{\mathbf{ff}^T} \mathbf{H}_n \delta \mathbf{u}_n^{ff} + \delta \mathbf{u}_n^{ff^T} \mathbf{g}_n$$

$$\mathbf{g}_n \triangleq \mathbf{r}_n + \mathbf{B}_n^T \mathbf{s}_{n+1}$$

$$\mathbf{G}_n \triangleq \mathbf{P}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{A}_n$$

$$\mathbf{H}_n \triangleq \mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n$$

$$\mathbf{S}_N = \mathbf{Q}_N, \qquad \mathbf{s}_N = \mathbf{q}_N, \qquad s_N = q_N$$

- ★ S(n) are only a function of known quantities: system matrix, control gain matrix, cost terms
- ★ ...AND future S (backwards)
- ★ Principle of optimality: solve backwards in time

Optimal control

$$(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)$$

$$\delta\mathbf{u}_n = -\mathbf{H}_n^{-1}\mathbf{g}_n - \mathbf{H}_n^{-1}\mathbf{G}_n\delta\mathbf{x}_n$$

$$\delta\mathbf{x}_n \triangleq \mathbf{x}_n - \bar{\mathbf{x}}_n$$

$$\delta\mathbf{u}_n \triangleq \mathbf{u}_n - \bar{\mathbf{u}}_n$$

We have derived the 'incremental' policy, thus total control is

$$\mathbf{u}(n, x) = \bar{\mathbf{u}}_n + \delta\mathbf{u}_n^{ff} + \mathbf{K}_n(\mathbf{x}_n - \bar{\mathbf{x}}_n)$$

Optimal control f(n+1)

feed-forward term $\delta \mathbf{u}_n^{ff} = -\mathbf{H}_n^{-1}\mathbf{g}_n$

feedback gain matrix $\mathbf{K}_n := -\mathbf{H}_n^{-1}\mathbf{G}_n$

using definition can also write these equations as

$$\delta \mathbf{u}^{\mathbf{ff}} = (\mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n)^{-1} (\mathbf{r}_n + \mathbf{B}_n^T \mathbf{s}_{n+1})$$

$$\mathbf{K}_n = (\mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n)^{-1} (\mathbf{P}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{A}_n)$$

ILQGC main iteration

0. *Initialization:* we assume that an initial, feasible policy $\boldsymbol{\mu}$ and initial state $\mathbf{x}_0$ is given. Then, for every iteration (i):

Forward pass

1. *Roll-Out:* perform a forward-integration of the system dynamics (1.70) subject to initial condition $\mathbf{x}_0$ and the current policy $\boldsymbol{\mu}$. Thus, obtain the nominal state- and control input trajectories $\bar{\mathbf{u}}_n^{(i)}, \bar{\mathbf{x}}_n^{(i)}$ for $n = 0, 1, \ldots, N$.

nonlinear system

$$\bar{\mathbf{x}}_{n+1} = \mathbf{f}_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)$$

2. *Linear-Quadratic Approximation:* build a local, linear-quadratic approximation around every state-input pair $(\bar{\mathbf{u}}_n^{(i)}, \bar{\mathbf{x}}_n^{(i)})$ as described in Equations (1.75) to (1.78).

$$\delta\mathbf{x}_{n+1} \approx \mathbf{A}_n\delta\mathbf{x}_n + \mathbf{B}_n\delta\mathbf{u}_n$$

$$\mathbf{A}_n = \frac{\partial\mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{x}}$$

$$\mathbf{B}_n = \frac{\partial\mathbf{f}(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{u}}$$

Backwards pass

3. *Compute the Control Law:* solve equations (1.84) to (1.86) backward in time and design the affine control policy through equation (1.88).

$$\mathbf{u}(n, x) = \bar{\mathbf{u}}_n + \delta\mathbf{u}_n^{ff} + \mathbf{K}_n(\mathbf{x}_n - \bar{\mathbf{x}}_n)$$

4. Go back to 1. and repeat until the sequences $\bar{\mathbf{u}}^{(i+1)}$ and $\bar{\mathbf{u}}^{(i)}$ are sufficiently close.

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n + \mathbf{K}_n^T\mathbf{H}_n\mathbf{K}_n + \mathbf{K}_n^T\mathbf{G}_n + \mathbf{G}_n^T\mathbf{K}_n$$
$$\mathbf{s}_n = \mathbf{q}_n + \mathbf{A}_n^T\mathbf{s}_{n+1} + \mathbf{K}_n^T\mathbf{H}_n\delta\mathbf{u}_n^{ff} + \mathbf{K}_n^T\mathbf{g}_n + \mathbf{G}_n^T\delta\mathbf{u}_n^{ff}$$
$$s_n = q_n + s_{n+1} + \frac{1}{2}\delta\mathbf{u}_\mathbf{n}^{\mathbf{ff}\,T}\mathbf{H}_n\delta\mathbf{u}_n^{ff} + \delta\mathbf{u}_n^{ff\,T}\mathbf{g}_n$$

$$J \approx q_N + \delta\mathbf{x}_N^T\mathbf{q}_N + \frac{1}{2}\delta\mathbf{x}_N^T\mathbf{Q}_N\delta\mathbf{x}_N + \sum_{n=0}^{N-1}\{q_n + \delta\mathbf{x}_n^T\mathbf{q}_n + \delta\mathbf{u}_n^T\mathbf{r}_n + \frac{1}{2}\delta\mathbf{x}_n^T\mathbf{Q}_n\delta\mathbf{x}_n + \frac{1}{2}\delta\mathbf{u}_n^T\mathbf{R}_n\delta\mathbf{u}_n + \delta\mathbf{u}_n^T\mathbf{P}_n\delta\mathbf{x}_n\}$$

$\forall n \in \{0, \cdots, N-1\}$:

$$q_n = L_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n), \quad \mathbf{q}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{x}}, \quad \mathbf{Q}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{x}^2}$$
$$\mathbf{P}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{u}\partial\mathbf{x}}, \quad \mathbf{r}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{u}}, \quad \mathbf{R}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial\mathbf{u}^2}$$

SLQC Recap...

1. Initial guess for policy
2. Solve sub problem: Approximate value function with a linear-quadratic
3. yields new approximative policy
4. repeat

2.1 Forward pass:
integrate to get a state (and controls) trajectory

2.2 Backward pass
Solve simplified optimal control problem around state and control trajectory

3. Adjust guess for optimal control

ILQC
Overview of derivation

Linearize system dynamics
Quadratize cost

Compute value function
Compute optimal control
Solve for Riccati like equation
Solve Riccati like equation

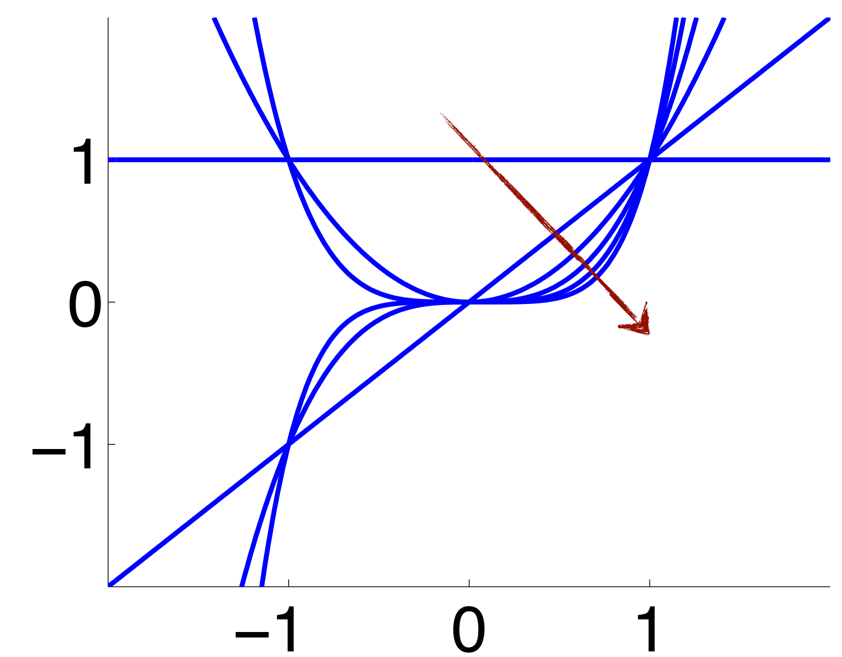
Local approximations...

Taylor series are polynomials

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i$$

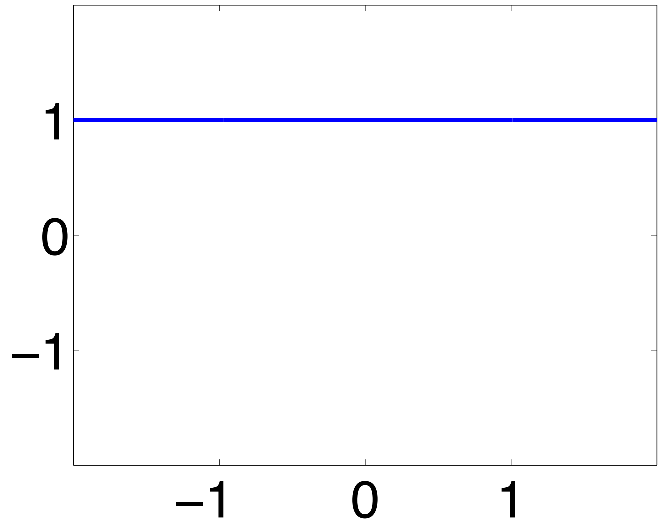
polynomials can (locally) approximate arbitrary function

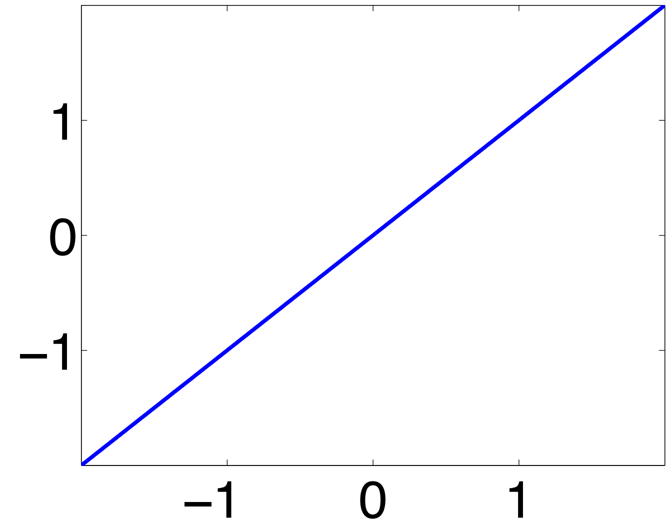
$$f(x) = \sum_{i=0}^{\infty} a_i x^i$$



x^0 x^1 x^2 x^3 x^4 x^5

i = 0



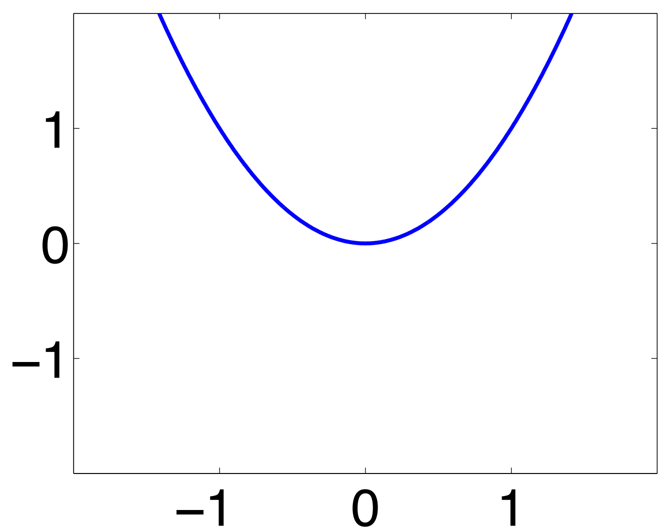


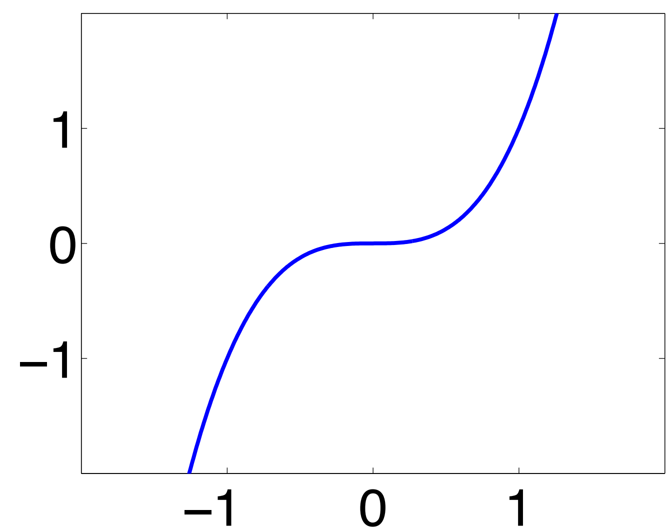
i = 1

symmetric

asymmetric

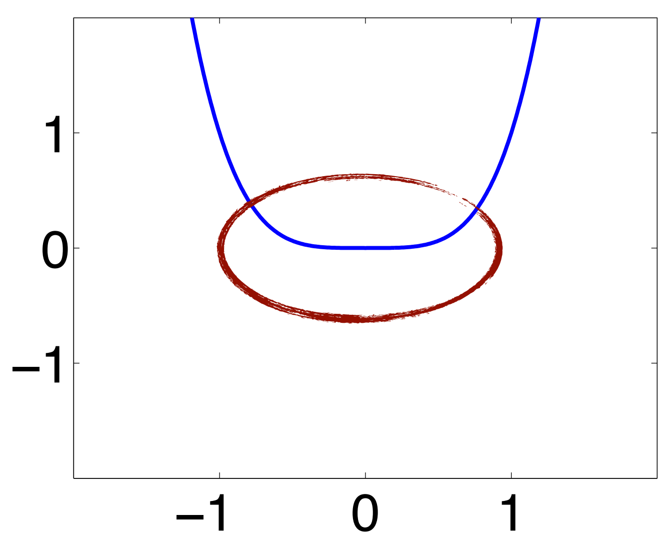
i = 2

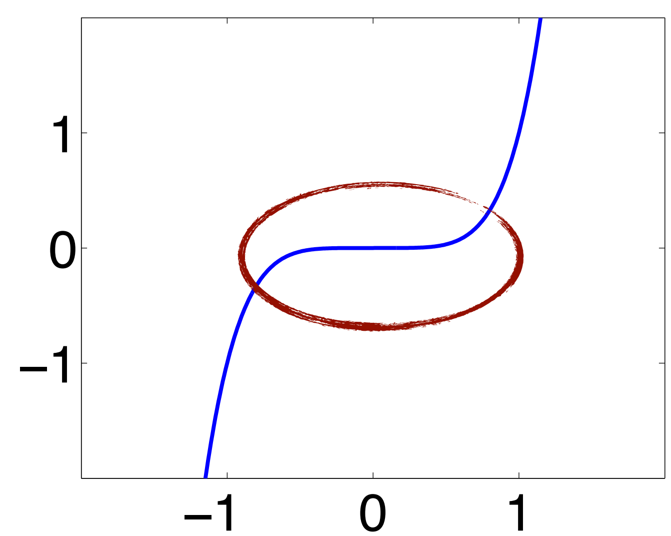




i = 3

i = 4



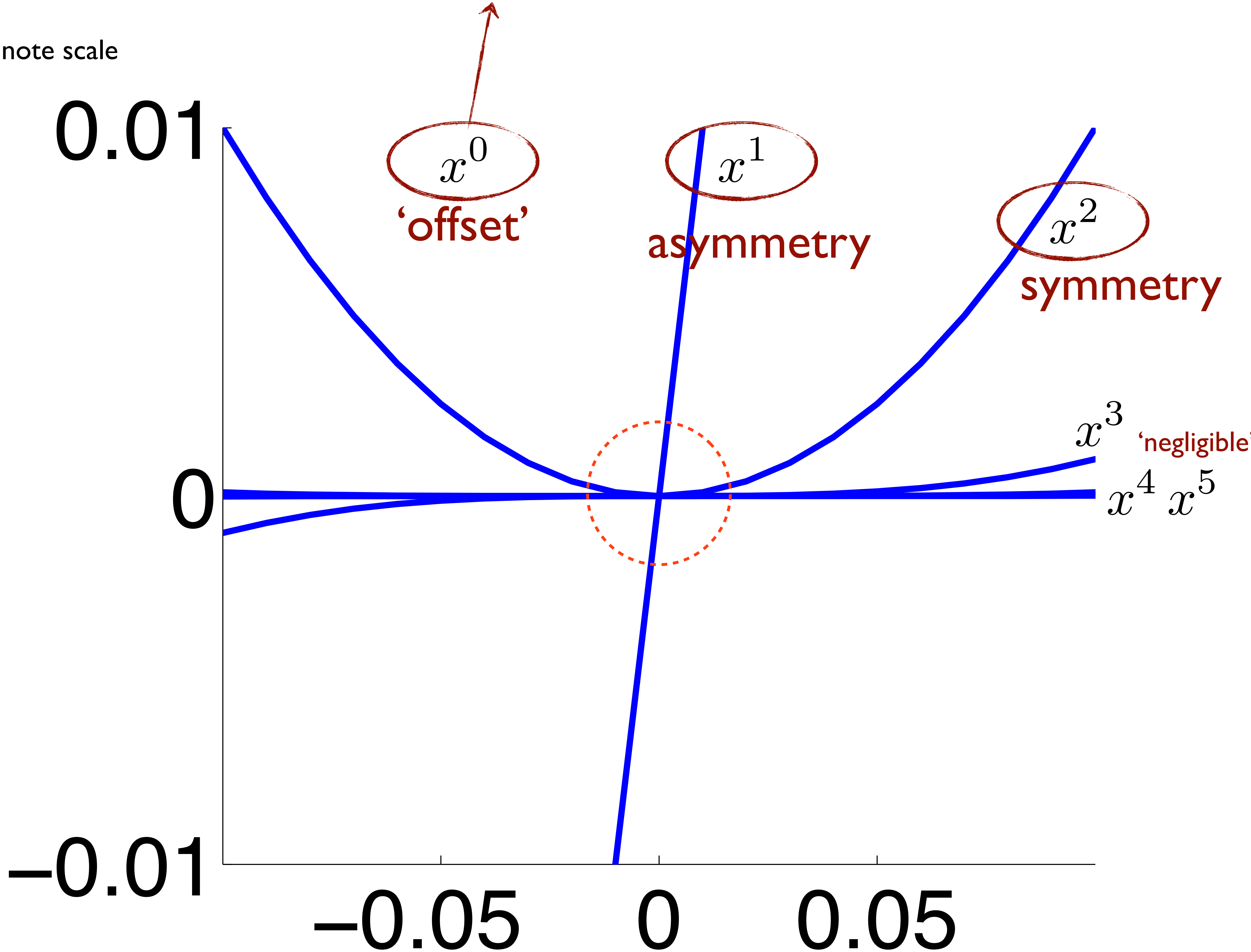


i = 5

A D R L

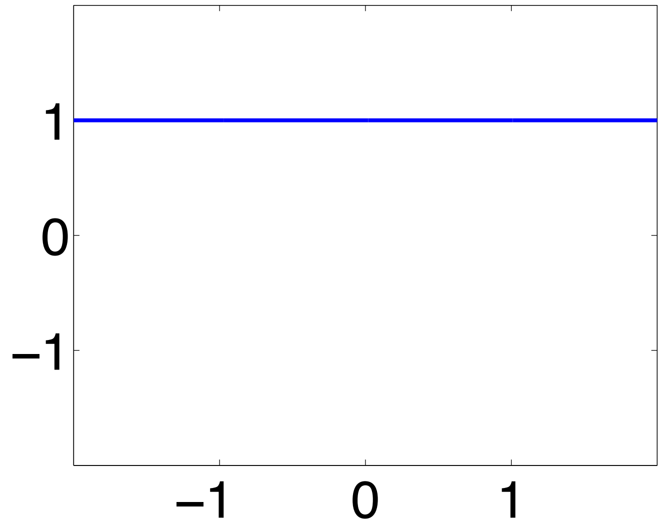
ETH Zürich

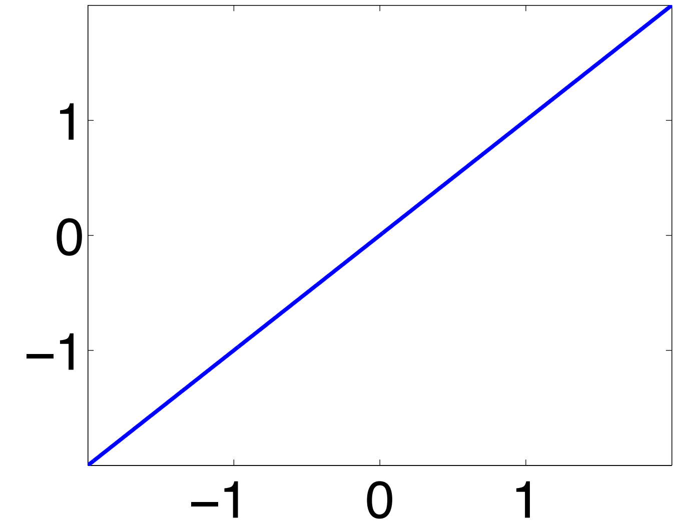
note scale



0.01
0
–0.01
–0.05
0
0.05
x^0
'offset'
x^1
asymmetry
x^2
symmetry
x^3
'negligible'
x^4 x^5

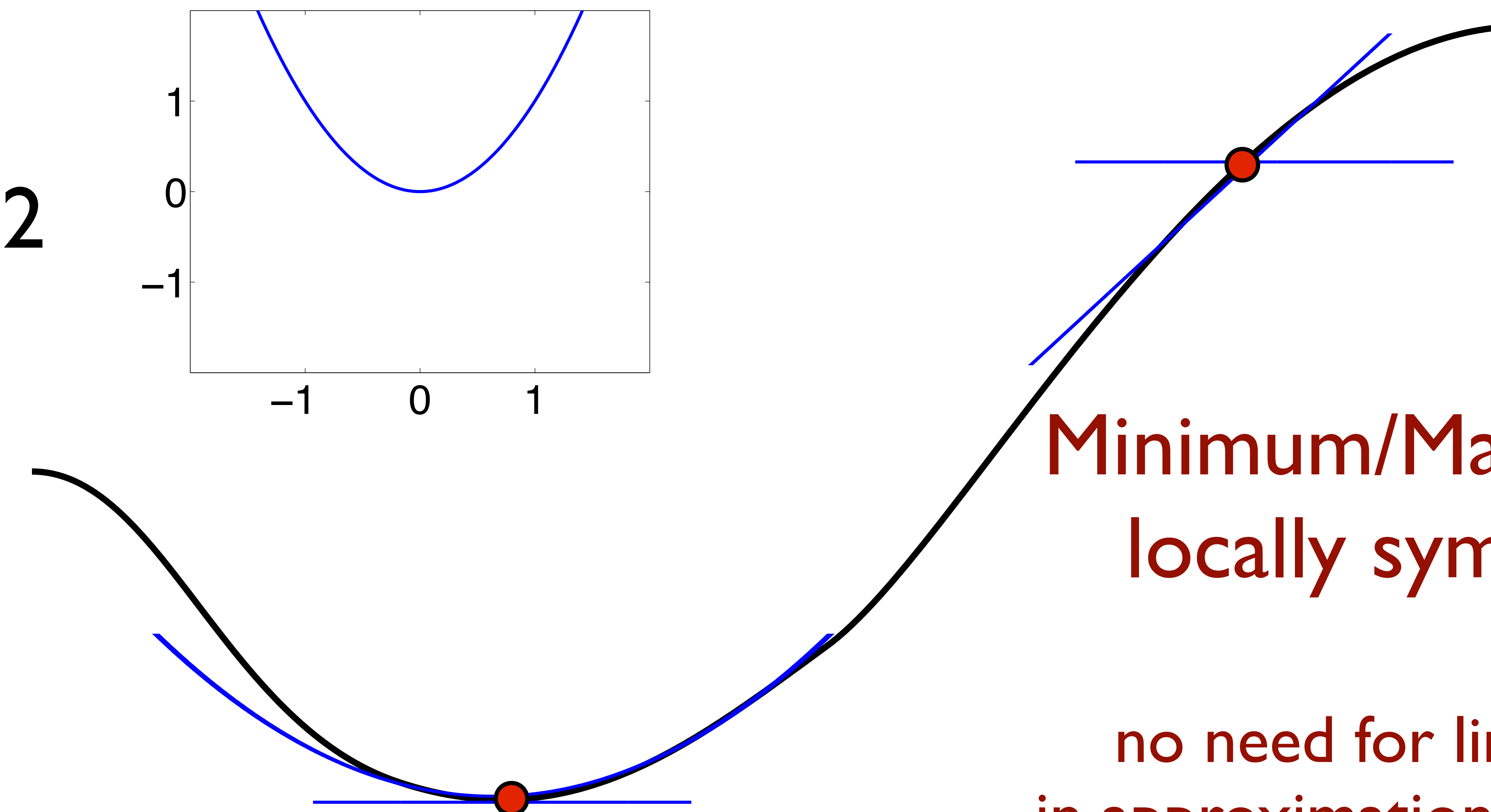
i = 0





i = 1

i = 2



Minimum/Maximum is locally symmetric

no need for linear term
in approximation at minimum

‘Shapeology’ of cost function

- Cost can have arbitrary constant offset
- at minimum:
  - Minimum is locally flat (slope 0) $$\frac{\partial C}{\partial x} = 0$$
  - Cost increases everywhere away from min. $$\frac{\partial^2 C}{\partial x^2} < 0$$
  - Symmetric
- not at minimum:
  - Not locally flat (slope not 0) $$\frac{\partial C}{\partial x} \neq 0$$
  - Cost increases towards ‘one side’ $$\frac{\partial^2 C}{\partial x^2} \in \mathbb{R}$$
  - asymmetric

‘typically’

$$\left|\frac{\partial^2 C}{\partial x^2}\right| \ll \left|\frac{\partial C}{\partial x}\right|$$

LQR - Linear Quadratic Regulator

Linearized System Dynamics

Quadratic cost function

Regulates output to zero

Discrete Time LQR

Linear Regulator

Linear (or linearized) system dynamics:

$$\mathbf{x}_{n+1} = \mathbf{A}_n \mathbf{x}_n + \mathbf{B}_n \mathbf{u}_n$$

Regulator: keep states at 0

$$\delta \mathbf{x}_n = \mathbf{x}_n$$

$$\delta \mathbf{u}_n = \mathbf{u}_n$$

Optimal Regulator

Linear

- Control:

  pure state feedback (no forward)

  linear control enough to stabilize locally

Quadratic

- Cost:

  regulator: optimum at x,u = 0

  increasing for any non-zero x,u

  ⇒ purely quadratic cost

Quadratization of cost function

$$J = \Phi(\mathbf{x}_N) + \sum_{n=0}^{N-1} L_n(\mathbf{x}_n, \mathbf{u}_n)$$

$$J \approx q_N + \delta\mathbf{x}_N^T \mathbf{q}_N + \frac{1}{2}\delta\mathbf{x}_N^T \mathbf{Q}_N \delta\mathbf{x}_N$$
$$+ \sum_{n=0}^{N-1} \{q_n + \delta\mathbf{x}_n^T \mathbf{q}_n + \delta\mathbf{u}_n^T \mathbf{r}_n + \frac{1}{2}\delta\mathbf{x}_n^T \mathbf{Q}_n \delta\mathbf{x}_n + \frac{1}{2}\delta\mathbf{u}_n^T \mathbf{R}_n \delta\mathbf{u}_n + \delta\mathbf{u}_n^T \mathbf{P}_n \delta\mathbf{x}_n\}$$

Control costs

State costs

'Mixing terms'

$\forall n \in \{0, \cdots, N-1\}$:

$$q_n = L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n), \qquad \mathbf{q}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}}, \qquad \mathbf{Q}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}^2}$$

$$\mathbf{P}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u} \partial \mathbf{x}}, \qquad \mathbf{r}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}}, \qquad \mathbf{R}_n = \frac{\partial^2 L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}^2}$$

$n = N$:

$$q_N = \Phi(\bar{\mathbf{x}}_n), \qquad \mathbf{q}_N = \frac{\partial \Phi(\bar{\mathbf{x}}_n)}{\partial \mathbf{x}}, \qquad \mathbf{Q}_N = \frac{\partial^2 \Phi(\bar{\mathbf{x}}_n)}{\partial \mathbf{x}^2}$$

Note that all derivatives w.r.t. $\mathbf{u}$ are zero for the terminal time-step N

Q, R, P are given through definition of cost!

Purely Quadratic cost

$$J = \frac{1}{2}\mathbf{x}_N^T \mathbf{Q}_N \mathbf{x}_N + \sum_{n=0}^{N-1} \frac{1}{2}\mathbf{x}_n^T \mathbf{Q}_n \mathbf{x}_n + \frac{1}{2}\mathbf{u}_n^T \mathbf{R}_n \mathbf{u}_n + \mathbf{u}_n^T \mathbf{P}_n \mathbf{x}_n$$

at optimum no linear term (locally symmetric)

cf with polynomial

$$q_n = L_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n) = 0 \qquad \mathbf{q}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}} = \mathbf{0} \qquad \mathbf{r}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}} = \mathbf{0}$$

$$q_n = L_n(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n) = 0 \qquad \mathbf{q}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{x}} = \mathbf{0} \qquad \mathbf{r}_n = \frac{\partial L(\bar{\mathbf{x}}_n, \bar{\mathbf{u}}_n)}{\partial \mathbf{u}} = \mathbf{0}$$

$$\mathbf{S}_N = \mathbf{Q}_N, \qquad \mathbf{s}_N = 0, \qquad s_N = 0$$

$$\mathbf{g}_n \triangleq \mathbf{r}_n + 0\,\mathbf{s}_{n+1}$$
$$\mathbf{G}_n \triangleq \mathbf{P}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{A}_n$$
$$\mathbf{H}_n \triangleq \mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n$$

$$\delta \mathbf{u}_n^{ff} = -\mathbf{H}\,0\,\mathbf{g}_n$$

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{A}_n + \mathbf{K}_n^T \mathbf{H}_n \mathbf{K}_n + \mathbf{K}_n^T \mathbf{G}_n + \mathbf{G}_n^T \mathbf{K}_n$$
$$\mathbf{s}_n = \mathbf{q}_n + \mathbf{A}_n^T \mathbf{s}_{n+1} + \mathbf{K}_n^T \mathbf{H}_n\,0\,\mathbf{u}_n^{ff} + \mathbf{K}_n^T \mathbf{g}_n + \mathbf{G}_n^T \delta \mathbf{u}_n^{ff}$$
$$s_n = q_n + s_{n+1} + \frac{1}{2} \delta \mathbf{u}_{\mathbf{n}}^{\mathbf{ff}\,T} \mathbf{H}\,0\,\mathbf{u}_n^{ff} + \delta \mathbf{u}_n^{ff\,T} \mathbf{g}_n$$

A D R L

Buchli - OLCAR - 2015

ETH Zürich

Ansatz for Value Function

$$V^*(n+1, \delta\mathbf{x}_{n+1}) = s_{n+1} + \delta\mathbf{x}_{n+1}^T\mathbf{s}_{n+1} + \frac{1}{2}\delta\mathbf{x}_{n+1}^T\mathbf{S}_{n+1}\delta\mathbf{x}_{n+1}$$

$$V^*(n, \mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{S}_n\mathbf{x}$$

Ricatti Equation

$$\mathbf{K}_n := -\mathbf{H}_n^{-1}\mathbf{G}_n$$

$$\begin{aligned}\mathbf{S}_n &= \mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n + \mathbf{K}_n^T\mathbf{H}_n\mathbf{K}_n + \mathbf{K}_n^T\mathbf{G}_n + \mathbf{G}_n^T\mathbf{K}_n \\ &= \mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n - \mathbf{G}_n^T\mathbf{H}_n^{-1}\mathbf{G}_n\end{aligned}$$

$$\mathbf{G}_n \triangleq \mathbf{P}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{A}_n$$
$$\mathbf{H}_n \triangleq \mathbf{R}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{B}_n$$

Discrete time Riccati equation

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n - (\mathbf{P}_n^T + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{B}_n)(\mathbf{R}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{B}_n)^{-1}(\mathbf{P}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)$$

solve backwards $\mathbf{S}_N = \mathbf{Q}_N$

Optimal policy

$$\begin{aligned}\boldsymbol{\mu}^*(n, \mathbf{x}) &= -\mathbf{H}_n^{-1}\mathbf{G}_n\mathbf{x} \\ &= -(\mathbf{R}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{B}_n)^{-1}(\mathbf{P}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)\mathbf{x}\end{aligned}$$

Infinite time LQR

$$J = \sum_{n=0}^{\infty} \frac{1}{2}\mathbf{x}_n^T \mathbf{Q} \mathbf{x}_n + \frac{1}{2}\mathbf{u}_n^T \mathbf{R} \mathbf{u}_n + \mathbf{u}_n^T \mathbf{P} \mathbf{x}_n$$

Value function not a function of time

Algebraic Riccati Equation

$$V^*(n, \mathbf{x}) = \frac{1}{2}\mathbf{x}^T \mathbf{S}_n \mathbf{x}$$

Value function not a function of time

$$\mathbf{S}_n = \mathbf{S}_{n+1} \quad =: \mathbf{S}$$

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{A}_n - (\mathbf{P}_n^T + \mathbf{A}_n^T \mathbf{S}_{n+1} \mathbf{B}_n)(\mathbf{R}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{B}_n)^{-1}(\mathbf{P}_n + \mathbf{B}_n^T \mathbf{S}_{n+1} \mathbf{A}_n)$$

Discrete time algebraic Riccati equation

$$\mathbf{S} = \mathbf{Q} + \mathbf{A}^T \mathbf{S} \mathbf{A} - (\mathbf{P}^T + \mathbf{A}^T \mathbf{S} \mathbf{B})(\mathbf{R} + \mathbf{B}^T \mathbf{S} \mathbf{B})^{-1}(\mathbf{P} + \mathbf{B}^T \mathbf{S} \mathbf{A})$$

$$\boldsymbol{\mu}^*(\mathbf{x}) = -(\mathbf{R} + \mathbf{B}^T \mathbf{S} \mathbf{B})^{-1}(\mathbf{P} + \mathbf{B}^T \mathbf{S} \mathbf{A})\mathbf{x}$$

Solve algebraic Riccati Eq.?

$$\mathbf{S} = \mathbf{Q} + \mathbf{A}^T\mathbf{S}\mathbf{A} - (\mathbf{P}^T + \mathbf{A}^T\mathbf{S}\mathbf{B})(\mathbf{R} + \mathbf{B}^T\mathbf{S}\mathbf{B})^{-1}(\mathbf{P} + \mathbf{B}^T\mathbf{S}\mathbf{A})$$

Can be solved going back to recursive definition:

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{A}_n - (\mathbf{P}_n^T + \mathbf{A}_n^T\mathbf{S}_{n+1}\mathbf{B}_n)(\mathbf{R}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{B}_n)^{-1}(\mathbf{P}_n + \mathbf{B}_n^T\mathbf{S}_{n+1}\mathbf{A}_n)$$

iterate (backwards in time)

initial condition: $\mathbf{S}_\infty = 0$

Computer algebra packages (e.g. Mathematica) can solve such equations

Continuous time LQR

Continuous time LQR

Continuos-time linear time variant system

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$

$$J = \frac{1}{2}\mathbf{x}(T)^T\mathbf{Q}_T\mathbf{x}(T) + \int_0^T \left(\frac{1}{2}\mathbf{x}(t)^T\mathbf{Q}(t)\mathbf{x}(t) + \frac{1}{2}\mathbf{u}(t)^T\mathbf{R}(t)\mathbf{u}(t) + \mathbf{u}(t)^T\mathbf{P}(t)\mathbf{x}(t)\right) dt.$$

Hamilton Jacobi Bellman Equation:

$$-\frac{\partial V^*}{\partial t} = \min_{u \in U}\{L(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T f(x,u)\}$$

$$-\frac{\partial V^*}{\partial t} = \min_{u \in U}\{L(x,u) + \left(\frac{\partial V^*}{\partial x}\right)^T f(x,u)\}$$

$$\min_{u \in U}\{\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \frac{1}{2}\mathbf{u}^T\mathbf{R}\mathbf{u} + \mathbf{u}^T\mathbf{P}\mathbf{x} + \left(\frac{\partial V^*}{\partial x}\right)^T (\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t))\}$$

final value $V(T,\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}_T\mathbf{x}$

Quadratic Ansatz for Value function

$$V^*(t,\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{S}(t)\mathbf{x}$$

$$\frac{\partial V^*(t,\mathbf{x})}{\partial t} = \frac{1}{2}\mathbf{x}^T\dot{\mathbf{S}}(t)\mathbf{x}$$

$$\frac{\partial V^*(t,\mathbf{x})}{\partial \mathbf{x}} = \mathbf{S}(t)\mathbf{x}$$

(cf L3)

substitute into:

$$-\frac{\partial V^*}{\partial t} == \min_{u\in U}\{\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \frac{1}{2}\mathbf{u}^T\mathbf{R}\mathbf{u} + \mathbf{u}^T\mathbf{P}\mathbf{x} + \left(\frac{\partial V^*}{\partial x}\right)^T(\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t))\}$$

$$-\mathbf{x}^T\dot{\mathbf{S}}(t)\mathbf{x} = \min_{u\in U}\{\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} + 2\mathbf{u}^T\mathbf{P}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{B}\mathbf{u}\}$$

A D R L

ETH Zürich

Optimal control

$$-\mathbf{x}^T\dot{\mathbf{S}}(t)\mathbf{x} = \min_{u\in U}\{\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} + 2\mathbf{u}^T\mathbf{P}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{B}\mathbf{u}\}$$

$$\nabla_{\mathbf{u}}\left[\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} + 2\mathbf{u}^T\mathbf{P}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{B}\mathbf{u}\right] = 0$$

$$2\mathbf{R}\mathbf{u} + 2\mathbf{P}\mathbf{x} + 2\mathbf{B}^T\mathbf{S}(t)\mathbf{x} = 0$$

$$\mathbf{u}^*(t,\mathbf{x}) = -\mathbf{R}^{-1}\left(\mathbf{P} + \mathbf{B}^T\mathbf{S}(t)\right)\mathbf{x}$$

A D R L

ETH Zürich

Solve for S

Substitute optimal control

$$\mathbf{u}^*(t,\mathbf{x}) = -\mathbf{R}^{-1}\left(\mathbf{P} + \mathbf{B}^T\mathbf{S}(t)\right)\mathbf{x}$$

in

$$-\mathbf{x}^T\dot{\mathbf{S}}(t)\mathbf{x} = \min_{u\in U}\{\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u} + 2\mathbf{u}^T\mathbf{P}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{S}(t)\mathbf{B}\mathbf{u}\}$$

$$\mathbf{x}^T\Big[\mathbf{S}(t)\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{S}(t) - \left(\mathbf{P}(t) + \mathbf{B}^T(t)\mathbf{S}(t)\right)^T\mathbf{R}^{-1}\left(\mathbf{P}(t) + \mathbf{B}^T(t)\mathbf{S}(t)\right) + \mathbf{Q}(t) + \dot{\mathbf{S}}(t)\Big]\mathbf{x} = 0$$

for all states x

$$\dot{\mathbf{S}} = -\mathbf{S}\mathbf{A} - \mathbf{A}^T\mathbf{S} + \left(\mathbf{P} + \mathbf{B}^T\mathbf{S}\right)^T\mathbf{R}^{-1}\left(\mathbf{P} + \mathbf{B}^T\mathbf{S}\right) - \mathbf{Q}$$

$$\mathbf{S}(T) = \mathbf{Q}_T$$

Stochastic LQR?!