Learning Contact Dynamics with LCP Constraint Relaxations

Samuel Pfrommer, Michael Posa

July 24, 2019

DAIR Lab at the University of Pennsylvania

Overview

- Problems with manipulation
- Existing approaches
- How are we learning?
- What are we learning?

Problems with manipulation

- Sudden changes in dynamics when making/breaking contact
- Inconsistencies with Coulomb friction (Painlevé paradox)
- Many simultaneous contacts
- Stick/slip transitions

The Linear Complimentarity Problem (LCP)

- Given matrix M and vector q, find vector λ such that:

$$\begin{aligned} \lambda &\geq 0 \\ M\lambda + q &\geq 0 \quad \text{nonnegativity} \\ \lambda^T(M\lambda + q) &= 0 \quad \text{complimentarity} \end{aligned}$$

Shorthand:

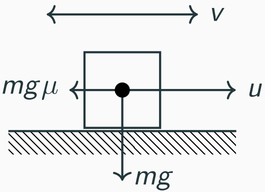
$$0 \leq \lambda \perp M\lambda + q \geq 0$$

- If M positive definite, Lemke's algorithm always finds λ
- Can formulate contact problems using complimentarity
  - Either separation distance is zero, or normal force is
  - Either tangential velocity is zero, or friction force is at boundary of cone (uses slack variables)
  - Solve an LCP for each time step
  - Described by Stewart, Anitescu

A simple example

- 1 DOF block with friction and external force
- Encodes stick/slip, no separation

$$\begin{aligned} v_{k+1} &= v_k + \lambda^+_{k+1} - \lambda^-_{k+1} + u_{k+1} && \text{dynamics equation} \\ 0 \leq \lambda^+_{k+1} &\perp \gamma_{k+1} + v_{k+1} \geq 0 && \text{positive friction if block moving left} \\ 0 \leq \lambda^-_{k+1} &\perp \gamma_{k+1} - v_{k+1} \geq 0 && \text{negative friction if block moving right} \\ 0 \leq \gamma_{k+1} &\perp mg\mu - \lambda^+_{k+1} - \lambda^-_{k+1} \geq 0 && \text{max friction if block is moving} \end{aligned}$$



A simple example (matrix form)

$$0 \le \underbrace{\begin{bmatrix} \lambda^+_{k+1} \\ \lambda^-_{k+1} \\ \gamma_{k+1} \end{bmatrix}}_{\lambda} \perp \underbrace{\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}}_{\mathrm{M}} \begin{bmatrix} \lambda^+_{k+1} \\ \lambda^-_{k+1} \\ \gamma_{k+1} \end{bmatrix} + \underbrace{\begin{bmatrix} v_k + u_{k+1} \\ -v_k - u_{k+1} \\ mg\mu \end{bmatrix}}_{\mathrm{q}} \ge 0$$

- M is not PSD, but with some small adjustments Anitescu shows Lemke always finds solution
- Not necessarily unique
- Generalizes to friction cone approximation

Existing approaches

Learned

- Often in context of policy learning
- Slow and data innefficient
- Doesn't leverage existing understanding of contact dynamics

Hybrid

- Best of both worlds
- Sim-to-real
- Residual physics
- **Differentiation through LCPs**

Analytical

- Only an approximation
- Doesn't fully capture real-world phenomena

LCP differentiation (Belbute-Peres / Amos / Kolter)

- Similar to previous work on QPs
  - KKT conditions for QP are an LCP
- Gives gradients of LCP solutions with respect to M and q
- Forms bilevel optimization problem:

$$
\begin{aligned}
\min_{M,q} \quad & \sum_i (\text{Dynamics}(x_i,\lambda_i) - \bar{x}_i)^2 \\
\text{subject to} \quad & \lambda_i \in LCP(M, q)
\end{aligned}
$$

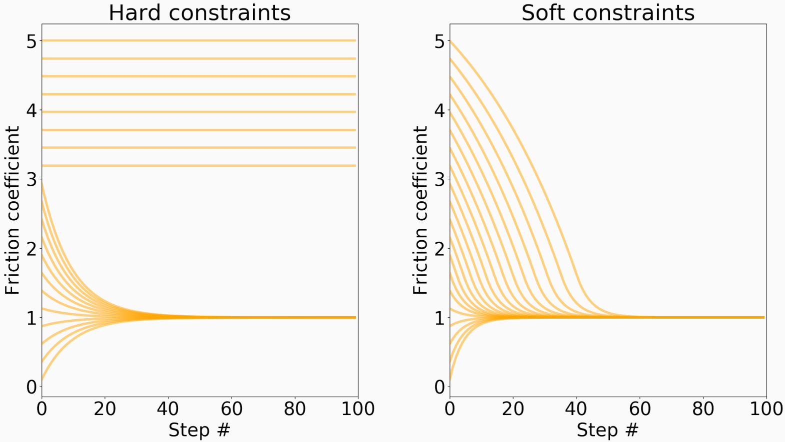
- **Problem:** bad priors can cause zero gradients
- **Cause:** hard constriants in LCP

Reformulating the optimization problem

- Move cost into subproblem
- Soften hard LCP constraints
- Allow unphysical behavior to reduce prediction error

$$
\begin{aligned}
\min_{M,q} \quad & \sum_i c_i \\
\text{subject to} \quad & c_i = \min_{\lambda_i \geq 0} \quad (\text{Dynamics}(x_i, \lambda_i) - \bar{x}_i)^2 \\
& \qquad\qquad + \lambda_i^T (M\lambda_i + q) + \text{hinge}(M\lambda_i + q)
\end{aligned}
$$

Learning $\mu = 1$



The lower level optimization problem

- Can be written in quadratic form:

$$(\mathrm{Dynamics}(x_i,\lambda_i) - \bar{x}_i)^2 = \left(v_{k+1} - (v_k + \lambda_i^+ - \lambda_i^- + u_i)\right)^2$$

$$\text{Write } \beta_i = v_{k+1} - v_k - u_i\text{:}$$

$$= \lambda_i^T \underbrace{\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{D} \lambda_i + \underbrace{\begin{bmatrix} -2\beta_i & 2\beta_i & 0 \end{bmatrix}}_{b} \lambda_i$$

- Get lower-level QP (with hard nonnegativity constraint):

$$\begin{aligned} \min_{\lambda_i \geq 0} \quad & \frac{1}{2}\lambda_i^T (Q + Q^T)\lambda_i + (q^T + b)\lambda_i \\ & M\lambda_i + q \geq 0 \\ & Q = M + D \end{aligned}$$

Convexity guarantees

- Want symmetric part of $M + D$ to have nonnegative eigenvalues
- Eigenvalues of D nonnegative, real
- Eigenvalues of M complex with nonnegative real part
- Extend analysis to full Anitescu formulation
- Scaling D with sufficiently large constant?

Structured learning

- Elementwise learning of M, q undesirable
  - Can prove an ϵ error in single q element breaks sticking
  - Coefficient antisymmetry across rows breaks sticking
- How much structure do we need to impose?
  - Learning physical parameters intuitive but restrictive
  - Eventually progress to noninterpretable parameters

Thank you!

1. Anitescu, M., and F. A. Potra. “Formulating Dynamic Multi-rigid-body Contact Problems with Friction as Solvable Linear Complementarity Problems.” (1996).
2. Stewart, David, and Jeffrey C. Trinkle. “An implicit time-stepping scheme for rigid body dynamics with coulomb friction.” Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065). Vol. 1. IEEE, 2000.
3. de Avila Belbute-Peres, Filipe, et al. “End-to-end differentiable physics for learning and control.” Advances in Neural Information Processing Systems. 2018.
4. Amos, Brandon, and J. Zico Kolter. “Optnet: Differentiable optimization as a layer in neural networks.” Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, 2017.