Last Time:

- Convex MPC Examples
- Nonlinear Traj Opt
- DDP/iLQR

Today:
- DDP details + extensions
- Constraints

---

* DDP Recap:

- Solve the unconstrained Traj Opt problem:

$$\min_{\substack{x_{1:N} \\ u_{1:N-1}}} J = \sum_{n=1}^{N-1} l(x_n, u_n) + l_N(x_N)$$

$$\text{s.t. } x_{n+1} = f(x_n, u_n)$$

- Backward Pass:

$$V_n(x+\Delta x) \approx V(x) + p_n^T \Delta x + \frac{1}{2} \Delta x^T P_n \Delta x$$

$$P_N = \nabla^2 l_N(x) \quad , \quad p_N = \nabla l_N(x)$$

$$V_{N-1}(x+\Delta x) = \min_{\Delta u} S(x+\Delta x, u+\Delta u)$$

$$\Rightarrow \boxed{\begin{aligned} \Delta u_{N-1} &= -d_{N-1} - K_{N-1} \Delta x_{N-1} \\ P_{N-1} &= G_{xx} + K^T G_{uu} K - G_{xu} K - K^T G_{ux} \\ p_{N-1} &= g_x - K^T g_u + K^T G_{uu} d - G_{xu} d \end{aligned}}$$

- Forward Rollout

$$\Delta J = 0$$
$$x_1' = x_1$$

for $k = 1 : N-1$

$$u_n' = u - \alpha d_n - K_n(x_n' - x_n)$$

$$x_{n+1}' = f(x_n', u_n')$$

$$\Delta J \leftarrow \Delta J + \alpha g_{u,n} d_n$$

end

- Line Search:

$$\alpha = 1$$

do:

$$x', u', \Delta J = \text{rollout}(x, u, d, K, \alpha)$$

$$\alpha \leftarrow c\alpha$$

while $J(x', u') < J(x, u) - b\Delta J$

$$x, u \leftarrow x', u'$$

- repeat until $|d|_\infty < tol$

Armijo parameters : $c \sim \frac{1}{2}$, $b \sim 1e^{-4} - 0.01$

* Examples:

  - Cartpole + acrobot swing up

  - DDP can converge in fewer iterations but iLQR often wins in wall-clock time

  - Problems are nonconvex $\Rightarrow$ can land in different local optima depending on initial guess

* Regularization

  - Just like like standard Newton, $V(x)$ and/or $S(x,u)$ Hessians can become indefinite in backward pass

  - Regularization is definitely necessary for DDP, often a good idea with iLQR as well.

  - Many options for regularizing:

  * Add a multiple of identity to $\nabla^2 S(x,u)$

  * Regularize P_n or G_n as needed in the backward pass

  * Regularize just $G_{uu} = \nabla^2_{uu} S(x,u)$ (this is the only matrix you have to invert):

$$d = G_{uu}^{-1} g_u \quad , \quad K = G_{uu}^{-1} G_{ux}$$

  - This last one is good for iLQR but not DDP

  - Regularization should not be required for iLQR but can be necessary due to floating-point error.

* DDP Notes:

+ Can be very fast (itterations + wall-clock)

+ One of the most efficient methods due to exploitation of DP structure

+ Always dynamically feasible due to forward rollout
  ⇒ can always execute on robot

+ Comes with TVLQR tracking controller for free

  ⇒ Can be very effective for online use

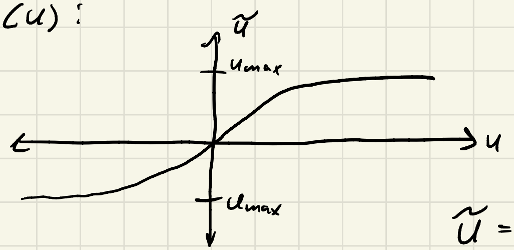
- Does not natively handle constraints

- Does not support infeasible initial guess for state trajectory due to forward rollout. Bad for "maze" or "bug-trap" problems.

- Can suffer from numerical ill-conditioning on long trajectories.

* Handling Constraints in DDP

- Many options depending on type of constraint
- Torque limits can be handled with a "squashing function" e.g. $\tanh(u)$:



$$\tilde{u} = u_{max} \tanh(u/u_{max})$$

- Effective, but adds nonlinearity and may need more itterations
- Better option: solve box-constrained QP in the backward pass:

$$\Delta u = \underset{\Delta u}{\operatorname{argmin}} \; S(x+\Delta x, u+\Delta u)$$

$$\text{s.t.} \quad u_{min} \leq u + \Delta u \leq u_{max}$$

- State constraints are harder. Often penalties are added to cost function. Can cause ill-conditioning.
- Better option: Wrap entire DDP algorithm in an Augmented Lagrangian method.
- AL method adds linear (multiplier) and quadratic (penalty) terms to the cost $\Rightarrow$ fits into DDP nicely.