Riccati Eqns & Lyapunov Theory:

Lyapunov Eqn: $\boxed{A^TP + PA + Q = 0}$, $Q = Q^T > 0$

Relationship to stability of A.

$\boxed{A \text{ stable iff } \exists\, P = P^T > 0 \text{ soln}}$

$\frac{d}{dt} e^{A^Tt} Q e^{At}$

$A^T(\cdot) + (\cdot)A$

exponential of e^{At}

e^{At}, A commute

if A stable $\Rightarrow \exists\, P = P^T > 0$ soln

direct computation $\boxed{P = \int_0^\infty e^{A^Tt} Q e^{At}\, dt}$

plugging in $A^T\left(\int_0^\infty e^{A^Tt} Q e^{At} dt\right) + \left(\int_0^\infty e^{A^Tt} Q e^{At} dt\right)A + Q$

$$\int_0^\infty \left(\frac{d}{dt} e^{A^Tt} Q e^{At}\right) dt + Q$$

$$e^{A^Tt} Q e^{At} \Big|_0^\infty + Q = 0$$

$$\underbrace{e^{A^T\infty} Q e^{A\infty}}_{0} - \underbrace{e^{A^T0} Q e^{A0}}_{Q} + Q = 0$$

Vector Field:

$\dot{x} = f(x)$

$x(t) = x(0) + \int - \, dt$

$\dot{x} = A(t)x$

$x(t) = \Phi(t,0)\,x(0)$ → formula complicated

$\dot{\Phi} = A(t)\Phi$

$\dot{x} = e^{At} x(0)$

Cost: $\int_0^\infty x^T Q x \, dt = x(0)^T P x(0)$

$x(t) = e^{At}x(0)$

$$\int_0^\infty \underbrace{x(0)^T e^{A^Tt}}_{x(t)^T} Q e^{At} x(0)\, dt = x(0)^T \int_0^\infty e^{A^Tt} Q e^{At} dt\, x(0)$$

if $\exists\, P = P^T > 0$ soln $\Rightarrow$ A stable

use P to prove stability ← Lyapunov stability theory (the linear case)

$Q = Q^T > 0$, $P = P^T > 0$

λ's > 0 λ's > 0

also for PSD matrices only

eigenvalues = singular values

let $V(x) = x^TPx$:

$\dot{V}(x) = \dot{x}^TPx + x^TP\dot{x}$

$= x^TA^TPx + x^TPAx$

$= x^T(A^TP + PA)x$ want < 0

$= -x^TQx < 0$

if $A^TP + PA = -Q$ $Q = Q^T > 0$

plug in Q

Note: x^TPx : Lyap or energy function

$-x^TQx$: dissipation function

going further...

$$\dot{V}(x) = -x^TQx \le -\lambda_{min}(Q)\, x^Tx \le \frac{-\lambda_{min}(Q)}{\lambda_{max}(P)} x^TPx \Longrightarrow \dot{V}(x) \le \frac{-\lambda_{min}(Q)}{\lambda_{max}(P)} V(x)$$

could be anything in LHP

$$-x^TQx = -x^TRER^Tx = -z^TEz = -\sum_i \lambda_i |z_i|^2 \le -\lambda_{min} z^Tz = -\lambda_{min} x^T\underbrace{RR^T}_{I}x$$

eival of A $\quad Re(\lambda_i) \le \frac{-\lambda_{min}(Q)}{\lambda_{max}(P)}$

Lyapunov Theory:

another way to classify stability ← pretty general

applies to non linear systems as well

Previous

eigenvalues of A

Lyapunov

finding an "energy" that decreasing along system trajectories

Definitions of Stability:

equilibrium at 0

stronger ↓

- Lyapunov stability $\exists \delta$ s.t. $|x(0)| < \delta \rightarrow |x(t)| < \varepsilon \ \forall \varepsilon$
- asymptotic stability $\exists \delta$ s.t. $|x(0)| < \delta \rightarrow |x(t)| \rightarrow 0$
- exponential stability $\exists \delta$ s.t. $|x(0)| < \delta \rightarrow |x(t)| \le c e^{-\lambda t} |x(0)|$

for $c, \lambda > 0$

decay rate bound

$\exists$ a stable linear system that decays slower

Global or Local

whole state space / within a ball

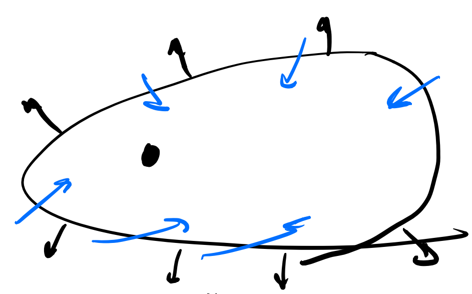
Lyapunov Function: $V(x)$ pos DEF "bowl shaped"

$\dot{x} = f(x)$

- $V(x) \ge 0 \quad \forall x$
- $V(x) = 0$ iff $x = 0$
- all level sets bounded
- $\dot{V}(x) = \underbrace{\frac{\partial V}{\partial x}\dot{x}}_{\text{Lie derivative}} = \frac{\partial V}{\partial x} f(x) \le 0$ ← Lyap stability

< 0 ← asymp stability

Lie derivative "the behavior of V along system trajectories"



Lyapunov analysis: making conc. about trajectories w/out computing them

- $\frac{\partial V}{\partial x} f \le 0 \Rightarrow x(t)$ bounded

  sub level sets are invariant sets

  $V(x(t)) \le V(x(0))$

- $\boxed{\frac{\partial V}{\partial x} f < 0} \Rightarrow x(t) \longrightarrow 0$

asymp stability

Suppose not: $0 < \varepsilon \le V(x(t)) \le V(x(0))$

set $C = \{x \mid \varepsilon \le V(x) \le V(x(0))\} \leftarrow$ compact

closed & bounded

$\sup_{x \in C} \dot{V} = -a < 0 \leftarrow$

lyap. style proof.

$$V(x(T)) = V(x(0)) + \int_0^T \underbrace{\dot{V}\, dt}_{\le -a \times T} \le V(x(0)) - aT$$

if $T > \frac{V(0)}{a} \Rightarrow V(x(T)) < 0$ impossible

- $\boxed{\dot{V} \le \mu V}$ for $\mu < 0 \longrightarrow$

Compare with $\dot{V} = \mu V$

$\dot{V} = \mu V: \; V(t) = e^{\mu t} V(0)$

$\dot{V} \le \mu V: \; V(t) \le e^{\mu t} V(0)$

Exponential stability

Define $u(t) = e^{-\mu t} V(t)$

$$\dot{u} = -\mu e^{-\mu t} V(t) + e^{-\mu t} \dot{V}$$

$$\le \underbrace{-\mu e^{-\mu t} V(t) + e^{-\mu t} \mu V}_{\text{cancels}} = 0$$

$$\Rightarrow u(t) = e^{-\mu t} V(t) \underset{\because \dot{u} \le 0}{\le} u(0) = V(0)$$

$$\Rightarrow \boxed{V(t) \le e^{\mu t} V(0)}$$

Lasalle's Invariance Principle:

allows you to prove asymp stable with just $\dot{V} \leq 0$

extra condition that if $\dot{V} = 0$ then $f(x) \neq 0$

(you have to keep moving through regions where you're not going down hill.)

Sym = RDR^T

skew sym = $\underbrace{R}_{r_1 r_2 r_3 r_4} \begin{bmatrix} 0 & -\lambda_1 & & \\ \lambda_1 & 0 & & \\ & & 0 & -\lambda_2 \\ & & \lambda_2 & 0 \end{bmatrix} \underbrace{R^T}_{\begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \\ r_4^T \end{bmatrix}} \rightarrow$ SVD $\begin{bmatrix} r_1 & r_2 & r_3 & r_4 \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_1 & & \\ & & \lambda_2 & \\ & & & \lambda_2 \end{bmatrix} \begin{bmatrix} - & r_2^T & - \\ - & r_1^T & - \\ - & r_4^T & - \\ - & r_3^T & - \end{bmatrix}$

Interpretations

- $J(x_0) = \int_0^\infty x^T Q x \, dt$ s.t. $\dot{x} = Ax$, $x(0) = x_0$ $\Big\} \rightarrow J(x_0) = x_0^T \int_0^\infty e^{A^T t} Q e^{At} dt \, x_0 \Rightarrow J(x_0) = x_0^T P x_0$ where $A^T P + PA + Q = 0$

- $J(x_0, K) = \int_0^\infty x^T Q x + u^T R u \, dt$
  s.t. $\dot{x} = Ax + Bu$, $x(0) = x_0$
  $u = Kx$

$\rightarrow J(x_0, K) = \underline{x_0^T} \int_0^\infty \underbrace{e^{(A+BK)^T t}}_{\text{closed Loop } A} (Q + \underbrace{K^T R K}_{x^T K^T R K x}) e^{(A+BK)t} dt \, x_0$

$$\left[P = \int_0^\infty e^{(A+BK)^T t} (Q + K^T R K) e^{(A+BK)t} dt \right.$$ assuming $A+BK$ stable

$\Rightarrow \underbrace{(A+BK)^T P + P(A+BK) + Q + K^T R K = 0}_{\text{Lyapunov Eqn}}$

$J(x_0, K) = x_0^T P x_0$

true for any K...

if plug in optimal $\underline{K = -R^{-1} B^T P}$

Lyapunov Eqn where $x(0)^T P x(0)$ gives cost-to-go for K

$$A^TP + PA + K^TB^TP + PBK + Q + K^TRK$$

$$A^TP + PA - PBR^{-1}B^TP - PBR^{-1}B^TP + Q + PBR^{-1}B^TP$$

$$A^TP + PA - \underline{P}BR^{-1}B^T\underline{P} + Q = 0$$

← RICCATI EQN. gives optimal K as $K = -R^{-1}B^TP$

Other examples:

observability gramian: solution to $A^TP + PA + \underbrace{C^TC}_{\text{PSD}} = 0$

controllability gramian: solution to $AP + PA^T + \underbrace{BB^T}_{\text{PSD}} = 0$

PSD ✓ ok is system is obs/cont.

energy in signal $y(t)$ output of

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

for impulse $u = \delta(t)$ with $x(0) = 0$

$$\left(\int_0^\infty y(t)^T y(t)\, dt\right)^{1/2} = \int_0^\infty$$

↳ $(CPC^T)^{1/2}$ where

$$AP + PA^T + BB^T = 0$$

assumed $Q = Q^T > 0$.

ok: $Q = Q^T \geq 0$ if A, Q is observable

$\boxed{(A, Q)} \Leftrightarrow (A, Q^{1/2})$

$$Q^{1/2}Q^{1/2} = Q$$

Riccati Eqns & Hamiltonian Systems:

↳ way to solve using an ODE

Lagrange Multiplier Method for solving LQR

$$\min_u \int_0^T x^TQx + u^TRu \, dt + x(T)^TQ(T)x(T) = J$$

s.t. $\dot{x} = Ax + Bu$, $x(0) = x_0$

Lagrangian: $L = J + \int_0^T \lambda(t)^T(Ax+Bu-\dot{x})\,dt + \lambda(0)^T(x(0)-x_0)$

lagrange multiplier; dynamics included as constraint

1) $\frac{\partial L}{\partial u} = \lambda(t)^TB + u^TR = 0$

$\boxed{u(t) = -R^{-1}B^T\lambda(t)}$ → input stationarity cond

$\lambda(t)$: costate → gradient info of value function

Note $\int_0^T \lambda^T\dot{x}\,dt = \lambda(T)^Tx(T) - \lambda(0)^Tx(0) - \int_0^T \dot{\lambda}^Tx\,dt$ integration by parts

2) $\frac{\partial L}{\partial x} = x(t)^TQ + \lambda(t)^TA + \dot{\lambda}^T = 0$

$\boxed{-\dot{\lambda}(t) = A^T\lambda(t) + Qx(t)}$ ← optimality costate eqn

costate terminal condition

3) $\frac{\partial L}{\partial x(T)} = x(T)^TQ(T) = \lambda(T)^T \Rightarrow \boxed{\lambda(T) = Q(T)x(T)}$

4) $\frac{\partial L}{\partial \lambda} \Rightarrow \boxed{\dot{x} = Ax + Bu}$ ← state eqn (dynamics)

5) $\frac{\partial L}{\partial \lambda(0)} \Rightarrow \boxed{x(0) = x_0}$ ← state initial cond. FEASIBILITY

optimality conds

$$-\dot{\lambda} = A^T\lambda + Qx$$
$$\dot{x} = Ax - BR^{-1}B^T\lambda$$

this doesn't lock in P

but $\lambda_1^T = x_1^TP \ldots$ $\lambda_n^T = x_n^TP$ does lock in P

Comments about $\lambda(t)$:

- costate
- at optimum; it's gradient of the optimal cost-to-go

$\boxed{\lambda^T = x^TP}$ } → suggest that $\cancel{P = x^{-1}\lambda}$ roughly

vectors

$\dot{\lambda} = \dot{x}^TP + x^T\dot{P}$

plug in Riccati

$= -Qx - A^T\lambda$

$x(t)$, $\lambda(t)$ solve Hamiltonian System:

$$\begin{bmatrix}\dot{x}\\ \dot{\lambda}\end{bmatrix} = \begin{bmatrix}A & -BR^{-1}B^T\\ -Q & -A^T\end{bmatrix}\begin{bmatrix}x\\ \lambda\end{bmatrix}$$

ODE, 2 PT BOUNDARY VALUE PROBLEM

physics { $x(0)$ given, $\lambda(T)$ given } hard to solve

state →

← costate

dynamic prog.

↳ other examples

- Markov Decision Processes
- Mean field games
- Dynamic games
- etc.

use this structure to find $P(t)$

Note: ODES can go backward or forward in time..

$\dot{x} = Ax$, $x(0) = x_0$ $\quad x(t) = e^{At}x_0$

$\dot{x} = Ax$, $x(T) = x_T$ $\quad x(t) = e^{-At}x_T$

$(-\dot{x} = -Ax)$

$A = V\underline{E}V^{-1}$ … $x(t) = V\begin{bmatrix}\frac{1}{e^{\lambda_1 t}} & & \\ & \ddots & \\ & & \frac{1}{e^{\lambda_n t}}\end{bmatrix}V^{-1}x_T$

$\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$

$P(T) = Q_T$

MATRIX ODE:

$$\begin{bmatrix}\dot X\\ \dot Y\end{bmatrix} = \begin{bmatrix}A & -BR^{-1}B^T\\ -Q & -A^T\end{bmatrix}\begin{bmatrix}X\\ Y\end{bmatrix}$$

solving n times with $X(T) = I$ ← arbitrary

$Y(T) = Q_T$

n is # of states

at time T

$\lambda_i(T) = P(T)x_i(T)$

$\Rightarrow Y = Q_T \Rightarrow \lambda_i(T) = Q_T[:,i]$

solving for n

$$X(t) = n\begin{bmatrix}x_1(t) \cdots x_n(t)\end{bmatrix} \;\&\; Y = \begin{bmatrix}\lambda_1(t) \cdots \lambda_n(t)\end{bmatrix}$$

all at once

You could $\begin{bmatrix}X\\ Y\end{bmatrix} = \begin{bmatrix}I\\ Q_T\end{bmatrix}M = \begin{bmatrix}M\\ Q_T M\end{bmatrix}$

$$\begin{bmatrix}X\\ Y\end{bmatrix} = \begin{bmatrix}x_1 & x_2 & \cdots\\ \lambda_1 & \lambda_2 & \cdots\end{bmatrix}$$

$\Downarrow$

$\lambda_1 = Px_1,\ \lambda_2 = Px_2, \cdots$

$Y = PX$

so by solving for multiple x's and λ's together... we're getting a full characterization of P in every direction

$$\boxed{P(t) = Y(t)X^{-1}(t)}$$

$Y M M^{-1} X^{-1}$

→ $\boxed{P \text{ solves the Riccati Eqn}}$

→ can check

$\dot P = \dot Y X^{-1} + Y \dot{X^{-1}}$

$-X^{-1}\dot X X^{-1}$

plug in dynamics

$X^{-1}X = I$

$\dot{X^{-1}}X + X^{-1}\dot X = 0$

$\dot{X^{-1}} = -X^{-1}\dot X X^{-1}$

Interpretation

$x_i(t)$: is the state that will evolve to e_i (stand. basis vec) a time T

$\lambda_i(t)$: is the gradient of CTG at x_i

Solves Riccati Diff EQ

ALGEBRAIC RICCATI EQN:

NOTE:

$$\underbrace{\begin{bmatrix}I & 0\\ -P & I\end{bmatrix}}_{Z^{-1}}\underbrace{\begin{bmatrix}A & -BR^{-1}B^T\\ -Q & -A^T\end{bmatrix}}_{H}\underbrace{\begin{bmatrix}I & 0\\ P & I\end{bmatrix}}_{Z} = \begin{bmatrix}A+BK & -BR^{-1}B^T\\ \cancel{\neq}\, 0 & -(A+BK)^T\end{bmatrix}$$

ALG. RIC EQN

H: $\underbrace{\lambda_1 \cdots \lambda_n}_{\text{eigenvalues of } A+BK}$, $\underbrace{-\lambda_1 \cdots -\lambda_n}_{\text{eigenvalues } -A-BK}$

stable

diagonalize $A+BK = TET^{-1}$

$$\Rightarrow \underbrace{\begin{bmatrix}T^{-1} & 0\\ 0 & T^T\end{bmatrix}\begin{bmatrix}I & 0\\ -P & I\end{bmatrix}}_{\begin{bmatrix}T^{-1} & 0\\ T^TP & T^T\end{bmatrix}} H \underbrace{\begin{bmatrix}I & 0\\ P & I\end{bmatrix}\begin{bmatrix}T & 0\\ 0 & T^{-T}\end{bmatrix}}_{\begin{bmatrix}T & 0\\ PT & T^{-T}\end{bmatrix}}$$

$$= \begin{bmatrix}E & -T^{-1}BR^{-1}B^TT^{-T}\\ 0 & -E\end{bmatrix}$$

$$H \begin{bmatrix} T & 0 \\ PT & T^{-T} \end{bmatrix} = \begin{bmatrix} T & 0 \\ PT & T^{-T} \end{bmatrix} \begin{bmatrix} E & -T^{-1}BR^{-1}B^T T^{-1} \\ 0 & -E \end{bmatrix}$$

$$H \begin{bmatrix} T \\ PT \end{bmatrix} = \begin{bmatrix} T \\ PT \end{bmatrix} E$$

cols are eigenvectors of H

→ with eigen values $\lambda_1 \dots \lambda_n$

So ...
to compute P solves Alg. Ric. Eqn.
compute n evecs of H corresponding to stable evals ...

$$\begin{bmatrix} X(T) \\ Y(T) \end{bmatrix} = \begin{bmatrix} T \\ P(T)T \end{bmatrix} = \begin{bmatrix} T \\ Q_T T \end{bmatrix}$$

put them in $\begin{bmatrix} F \\ G \end{bmatrix}$

$$\boxed{P = GF^{-1}}$$

More on costates (optimal control)

before: using costates to characterize optimum $\frac{\partial J}{\partial u} = 0$

now: use costates to compute $\frac{\partial J}{\partial u}$ → to do gradient descent.

Optimization:

Descent methods → gradient descent.

want to minimize $\frac{\partial J}{\partial u}$.

- compute $\frac{\partial J}{\partial u}$
- update $u^+ = u - \alpha \frac{\partial J}{\partial u}^+$

α: step size (line search)

How to compute $\frac{\partial J}{\partial u}$?

$$\min_{u(t)} \int_0^T \ell(x,u,t)\,dt + \ell(x,T)$$

$$\text{s.t.} \quad \dot{x} = f(x,u,t), \quad x(0) = x_0$$

$$\min_{u[t]} \sum_{t=0}^{T-1} \ell[x,u,t] + \ell[x,T]$$

$$\text{s.t.} \quad x[t+1] = f(x,u,t), \quad x[0] = x_0$$

Trick: augment dynamics to get rid of running cost.

new state: z s.t. $z[t+1] = z[t] + \ell(x,u,t)$

↳ $$\min_{u} \; \ell(x,T) + z[T] = \bar{J}$$

s.t.
$$\begin{pmatrix} x[t+1] \\ z[t+1] \end{pmatrix} = \begin{pmatrix} f(x,u,t) \\ \ell(x,u,t) + z[t] \end{pmatrix} \qquad \begin{aligned} x[0] &= x_0 \\ z[0] &= 0 \end{aligned}$$

$$\bar{x}[t+1] = \bar{f}(x,u,t) \qquad \bar{x}[0] = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$$

what happens if perturb $u + \Delta u$ ← signal

traj changes...

$$\cancel{\bar{x}[t+1]} + \Delta \bar{x}[t+1] \approx \cancel{\bar{f}} + \underbrace{\frac{\partial \bar{f}}{\partial \bar{x}}}_{:= A_t} \Delta \bar{x}[t] + \underbrace{\frac{\partial \bar{f}}{\partial u}}_{:= B_t} \Delta u[t]$$ linearize...

$$A_t = \begin{array}{c|c} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} n \\ 1 \end{array} & \left(\begin{array}{c|c} \left.\frac{\partial f}{\partial x}\right|_t & 0 \\ \hline \left.\frac{\partial \ell}{\partial x}\right|_t & 1 \end{array}\right) \end{array} \begin{array}{c} \Delta x \\ \Delta z \end{array} \qquad B_t = \begin{array}{c|c} & m \\ \begin{array}{c} n \\ 1 \end{array} & \begin{pmatrix} \left.\frac{\partial f}{\partial u}\right|_t \\ \left.\frac{\partial \ell}{\partial u}\right|_t \end{pmatrix} \end{array}$$

$$\Delta \bar{x}_T = \left[A_{T-1} \cdots A_0 \;\middle|\; A_{T-1} \cdots A_1 B_0 \; \cdots \;\middle|\; A_{T-1} B_{T-2} \;\middle|\; B_{T-1} \right] \begin{bmatrix} \Delta \bar{x}_0 \\ \hline \Delta u_0 \\ \vdots \\ \Delta u_{T-1} \end{bmatrix}$$

drift term

like LTV "controllability matrix"

$$\Delta \bar{J} = \frac{\partial \bar{J}}{\partial \bar{x}_T} \Delta \bar{x}_T \qquad \Delta \bar{J} = \frac{\partial \bar{J}}{\partial x_0, u} \begin{pmatrix} \Delta x_0 \\ \Delta u \end{pmatrix}$$

$$= \left[\left.\frac{\partial \ell}{\partial x}\right|_T \; 1 \right] \left[A_{T-1} \cdots A_0 \;\middle|\; A_{T-1} \cdots A_1 B_0 \; \cdots \;\middle|\; A_{T-1} B_{T-2} \;\middle|\; B_{T-1} \right] \begin{bmatrix} \Delta \bar{x}_0 \\ \hline \Delta u_0 \\ \vdots \\ \Delta u_{T-1} \end{bmatrix}$$

LTV DT state transition

call

$$\begin{bmatrix} \frac{\partial l}{\partial x_T} & 1 \end{bmatrix} \begin{bmatrix} \left.\frac{\partial f}{\partial x}\right|_{T-1} & 0 \\ \left.\frac{\partial l}{\partial x}\right|_{T-1} & 1 \end{bmatrix} \cdots \begin{bmatrix} \left.\frac{\partial f}{\partial x}\right|_{t} & 0 \\ \left.\frac{\partial l}{\partial x}\right|_{t} & 1 \end{bmatrix} := \begin{bmatrix} \lambda[t]^T & 1 \end{bmatrix}$$

$$\lambda[t-1] = \lambda[t]^T \left.\frac{\partial f}{\partial x}\right|_{t-1} + \frac{\partial l}{\partial x_{t+1}} \qquad \lambda[T] = \left.\frac{\partial l}{\partial x^T}\right|_T$$

backwards recursion

given λ $\quad \frac{\partial J}{\partial u_t} = \lambda[t+1]^T \left.\frac{\partial f}{\partial u}\right|_t + \left.\frac{\partial l}{\partial u}\right|_t \qquad \frac{\partial J}{\partial x_0} = \lambda[0]^T$

In continuous time version

$$\min_u \int_0^T l(x,u,t)\,dt + l(x,T)$$

$$\text{s.t. } \dot{x} = f(x,u,t), \quad x(0) = x_0$$

similar sys. augmentation ...

$$-\dot{\lambda}^T = \lambda^T \left.\frac{\partial f}{\partial x}\right|_t + \left.\frac{\partial l}{\partial x}\right|_t, \quad \lambda(T) = \left.\frac{\partial l}{\partial x}\right|_T \qquad \frac{\partial J}{\partial u} = \lambda(t)^T \left.\frac{\partial f}{\partial u}\right|_t + \left.\frac{\partial l}{\partial u}\right|_t$$

Compare w/ LQR

$$-\dot{\lambda}^T = \lambda^T A + x^T Q, \quad \lambda^T(T) = x^T Q_T \qquad \frac{\partial J}{\partial u} = \lambda^T B + u^T R \quad \cancel{=0}$$

$J(u)$, $\frac{\partial J}{\partial u}$ $\rightarrow$ do gradient descent on your cost. $\quad$ optimal control lol

signal or long vector