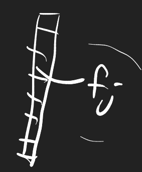
1) Nonlinear Control

$$\min_{s_0 \sim \mu_0} \mathbb{E}\left[\sum_{t=0}^{H-1} \underbrace{c}_{\text{cost}}(s_t, a_t) \;\middle|\; s_{t+1} = \underbrace{f}_{\text{dynamics}}(s_t, a_t) \right]$$

since we consider approximate method based on LQR, we will drop disturbance w_t

Assumption 1 c is 2x differentiable
f is 1x differentiable



$c: \mathbb{R}^{n_s} \times \mathbb{R}^{n_a} \to \mathbb{R}$

$\nabla_s c(s,a) \in \mathbb{R}^{n_s}$

$\nabla_a c(s,a) \in \mathbb{R}^{n_a}$

$\nabla_s^2 c(s,a) \in \mathbb{R}^{n_s \times n_s}$

$\nabla_a^2 c(s,a) \in \mathbb{R}^{n_a \times n_a}$

$\nabla_{as}^2 c(s,a) \in \mathbb{R}^{n_a \times n_s}$

$f: \mathbb{R}^{n_s} \times \mathbb{R}^{n_a} \to \mathbb{R}^{n_s}$

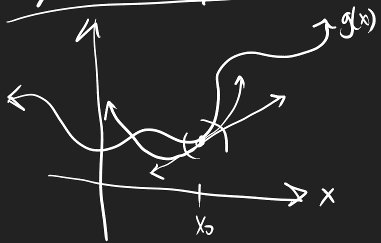
$\nabla_s f(s,a) \in \mathbb{R}^{n_s \times n_s}$

$\nabla_a f(s,a) \in \mathbb{R}^{n_a \times n_s}$

Assumption 2: we have black box access to f & c.
i.e. Query (s,a) and observe $s' = f(s,a)$
$c = c(s,a)$

IDEA: since we know what to do when dynamics are linear & costs quadratic,
1) linearize dynamics
2) 2nd order approx. on costs

2) Linear/Quadric Approximations



Taylor Expansion (1D)

$$g(x) = \underbrace{g(x_0) + g'(x_0)(x-x_0)} + \tfrac{1}{2} g''(x_0)(x-x_0)^2 + \dots$$

when $(x-x_0)$ is small, higher powers are even smaller

$\varepsilon^p \to 0$ as $p \to \infty$ if $|\varepsilon| < 1$

Linear Approx

$$\underline{f(s,a)} \approx f(s_0,a_0) + \underbrace{\nabla_s f(s_0,a_0)}_{\text{"Jacobian"}}{}^T (s - s_0) + \nabla_a f(s_0,a_0)^T (a - a_0)$$

row i col j $\nabla_s f$: $\dfrac{\partial f_j}{\partial s_i}(s_0,a_0)$

row i col j of $\nabla_a f$: $\dfrac{\partial f_j}{\partial a_i}(s_0,a_0)$

$$f(s,a) \approx As + Ba + v$$

Quadratic Approx

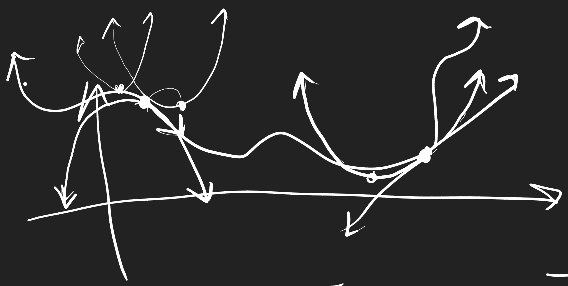
$$c(s,a) \approx c(s_0,a_0) + \nabla_s c(s_0,a_0)^T (s-s_0) + \nabla_a c(s_0,a_0)^T (a-a_0)$$

"Hessian"

$$+ \frac{1}{2}(s-s_0)^T \nabla_s^2 c(s_0,a_0)(s-s_0)$$

$$+ \frac{1}{2}(a-a_0)^T \nabla_a^2 c(s_0,a_0)(a-a_0)$$

$$+ (a-a_0)^T \nabla_{as}^2 c(s_0,a_0)(s-s_0)$$



$$c(s,a) \approx s^T \bar{Q} s + a^T \bar{R} a + a^T M s + q^T s + r^T a + c$$

Practical consideration:

1) put all negative eigenvalues to 0 $\rightarrow$ $-\gamma x^2$ $0 \cdot x$
2) add λI $\lambda > 0$ $\rightarrow \lambda x^2$

$$\bar{Q} = VDV^T = \sum_{i=1}^{ns} \sigma_i v_i v_i^T \qquad Q = \sum_{i=1}^{ns} \max(\sigma_i, 0) v_i v_i^T + \underline{\lambda I}$$

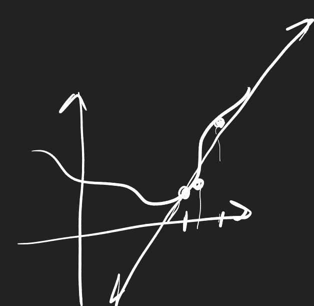
$$D = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_{ns} \end{bmatrix}$$

$$\bar{R} \rightarrow R$$

to summarize,

$$A, B, V, Q, R, q, r, c = \text{APPROX}(f, c, (s_0, a_0))$$

Black Box Access:



$$g'(x) \approx \frac{g(x+\delta) - g(x-\delta)}{2\delta}$$

"finite difference approx"

$$\frac{\partial f_i}{\partial s_j} \approx \frac{f_i(s + \delta e_j, a) - f_i(s - \delta e_j, a)}{2\delta} \qquad e_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow j$$

3) Local LQR control

Goal: stay close to (s^*, a^*)

$$A, B, V, Q, R, M, q, r, c = \text{APPROX}(f, c, (s^*, a^*))$$

$$\min \mathop{\mathbb{E}}_{s_0 \sim \mu_0} \left[\sum_{t=0}^{H-1} s_t^T Q s_t + a_t^T R a_t + a_t^T M s_t + q^T s_t + r^T a_t + c \,\middle|\, s_{t+1} = A s_t + B a_t + V \right]$$

still resuts in quadratic V^* and linear π^*

$$\pi_t^*(s) = K_t^* s + k_t^* \qquad \text{(HW 1)}$$

$$K^*, k^* = LQR(A, B, V, Q, R, M, q, r, c)$$

4) Iterative LQR

problem: when s, a are far from s^*, a^* the approximation can be bad.

Given trajectory $\tau = (s_t, a_t)_{t=0}^{H-1}$,

$$A_t, B_t, V_t, Q_t, R_t, M_t, q_t, r_t, c_t = APPROX(f, c, \tau)$$