dynamic programming

goal: apply nonlinear programming techniques to solve multi-stage optimization problems, i.e. optimal control problems

refs: Stengel 1994 § 3.4
Lewis, Vrabie, Syrmos 2012 Ch 6

- suppose now we are given DT DE $x^+ = f(x,u)$, $x \in \mathbb{R}^d$, $u \in \mathbb{R}^m$ and we wish to choose inputs over time $u : [0,t] \to \mathbb{R}^m$ to minimize

$$c(x,u) = \underbrace{\ell(t, x(t))}_{\text{"final" cost}} + \underbrace{\sum_{\tau=0}^{t-1} \mathcal{L}(\tau, x(\tau), u(\tau))}_{\text{"running" cost}}$$

- Richard Bellman published key insight in 1957:

idea: the optimal control $u(\tau)$ to apply at time τ depends only on $x(\tau)$ — not on previous states/inputs

$\to$ leads naturally to working backward from final time:

- letting $v_\tau^*(x(\tau))$ denote lowest (i.e. optimal) cost achievable from state $x(\tau) \in \mathbb{R}^d$ at time τ,

$$\underbrace{v_\tau^*(x(\tau))}_{\text{"value" of } x(\tau)} = \min_{u(\tau) \in \mathbb{R}^m} \left[\mathcal{L}(\tau, x(\tau), u(\tau)) + v_{\tau+1}^*\big(\underbrace{x(\tau+1)}_{= f(x(\tau), u(\tau))}\big) \right]$$

* this is referred to as a Bellman equation

$\to$ this tells us (in principle) how to determine optimal sequence of control inputs by solving a sequence of NLP backward in time!

→ consider $x^+ = ax + bu,\ x, u \in \mathbb{R}$, / q, r, a, b are given

$$c_\tau(x,u) = \sum_{s=\tau}^{t} q(s)x(s)^2 + r(s)u(s)^2$$

→ determine optimal input & value at $\tau = t$

- $c_t(x,u) = q(t)x(t)^2 + r(t)u(t)^2$

- $v_t^*(x(t)) = q(t)x(t)^2$ is optimal value w/ optimal input $u^*(t) = 0$

→ determine optimal input & value at $\tau = t-1$ using Bellman's equation

- $$v_{t-1}^*(x(t-1)) = \min_{u(t-1)} \Big[q(t-1)x(t-1)^2 + r(t-1)u(t-1)^2 + q(t)\underbrace{x(t)}_{\substack{\| \\ ax(t-1)+bu(t-1)}}^2 \Big]$$

$$= \min_{u_{t-1}} \left[q_{t-1}x_{t-1}^2 + r_{t-1}u_{t-1}^2 + q_t(a x_{t-1} + b u_{t-1})^2 \right]$$

- differentiating wrt u_{t-1} yields $2r_{t-1}u_{t-1} + 2q_t(a x_{t-1} + b u_{t-1}) \cdot b = 0$

$$\Leftrightarrow \quad u_{t-1} = \frac{-abq_t}{b^2 + r_t} x_{t-1} \quad \leftarrow \text{linear in } x_{t-1}!$$

∘ in principle, this process could be repeated to determine optimal input & value for all $\tau \in \{1, \dots, t\}$

→ we'll use Bellman's equation both analytically (verify optimality) & computationally (synthesize optimal inputs)