## · Causal state-feedback control

linear dynamical system, over finite horizon

$$x_{t+1} = Ax_t + Bu_t + w_t \qquad t = 0, \cdots T-1$$

$x_t \in \mathbb{R}^n$ is state, $u_t \in \mathbb{R}^m$ is the input at time $t$

$w_t$ is the process noise (or exogeneous input) at time $t$

$X_t = (x_0 \cdots x_t)$ is state history

Causal state-feedback control

$$u_t = \phi_t(X_t) = \psi_t(x_0\ w_0 \ \cdots\ x_{t-1}, w_{t-1}) \qquad t=0 \cdots T-1$$

$\phi_t : \mathbb{R}^{(t+1)\times n} \to \mathbb{R}^m$ called the control policy at time $t$

## · Stochastic finite horizon control

$(x_0, w_0 \cdots x_{t-1}, w_{t-1})$ is a random variable ( a stochastic process )

objective: $J = E\left( \sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T) \right)$

convex stage cost $\ell_t : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$

convex terminal cost $\ell_T : \mathbb{R}^n \mapsto \mathbb{R}$

$J$ depends on control policies $\phi_0, \cdots, \phi_{T-1}$

constraints: $u_t \in \mathcal{U}_t \qquad t=0, \cdots, T-1$

convex input constraint sets $\mathcal{U}_0, \cdots \mathcal{U}_{T-1}$

Stochastic control problem: choose control policies $\phi_0, \cdots, \phi_{T-1}$ to minimize $J$

the variable is $\phi_t$ upon which $u$ is chosen

key idea, we have to recourse ( feeback, closed-loop control )

we can change $u_t$ based on the observed state history $x_0, \cdots x_t$

## · Solution via dynamic programming

let $V_t(X_t)$ be optimal value of objective, from $t$ on, starting from initial state history $X_t$

$\left( V_T(X_T) = \ell_T(x_T) \qquad J^* = E\, V_0(X_0) \right)$

2 extremes:

1. prescient controller: tell you ahead of time what the disturbance sequence is
2. full ignorant policy: choose $u_0, u_1 \cdots u_T$ before you know any of $w$s

$V_t$ can be found by backward recursion:

this is random because $w_t$ is unknown

$$V_t(X_t) = \inf_{v \in \mathcal{U}} \left\{ \ell_t(x_t, v) + E\left[V_{t+1}(\underbrace{Ax_t + Bv + w_t}) \mid X_t\right] \right\}$$

optimal policy is causal state feedback

$$\phi_t^*(X_t) = \underset{v \in \mathcal{U}}{\arg\min} \left\{ \ell_t(x_t, v) + E\left[V_{t+1}(Ax_t + Bv + w_t) \mid x_t\right] \right\}$$

· Independent process noise

assume $x_0\ w_0 \cdots x_{T-1}\ w_{T-1}$ are independent

$V_t$ depends only on current state $x_t$ (not the history $X_t$)

Bellman equations: $V_T(x_T) = \ell_T(x_T)$

$$V_t(x_t) = \inf_{v \in \mathcal{U}} \left\{ \ell_t(x_t, v) + E\left[ V_{t+1}(Ax_t + Bv + w_t) \right] \right\}$$

· Linear quadratic stochastic control

special case of linear stochastic control

$\mathcal{U}_t = \mathbb{R}^m$

$x_0\ w_0 \cdots x_{T-1}, w_{T-1}$ are independent

$$E[x_0] = 0 \qquad E[w_i] = 0 \qquad E[x_0 x_0^T] = \Sigma \qquad E[w_t w_t^T] = W_t$$

$$\ell_t(x_t, u_t) = x_t^T Q_t x_t + u_t^T R_t u_t \qquad \text{with } Q_t \geq 0 \quad R_t > 0$$

$$\ell_T(x_T) = x_T^T Q_T x_T \qquad \text{with } Q_T \geq 0$$

can show value function are quadratic

$$V_t(x_t) = x_t^T P_t x_t + q_t \qquad t = 0, \dots, T$$

$$\begin{bmatrix} A & B^T \\ B & C \end{bmatrix}$$

row 2 − $BA^{-1}$ row 1

$C - BA^{-1}B^T$

Bellman recursion: $P_T = Q_T \quad q_T = 0$

$$V_t(z) = \inf_v \Big\{ z^T Q_t z + v^T R_t v + E_{w_t}\left[ (Az + Bv + w_t)^T P_{t+1} (Az + Bv + w_t) + q_{t+1} \right] \Big\}$$

$E[w_t^T P_{t+1}(Az + Bv)] = 0$ since $w_t$ is zero mean

$E[w_t^T P_{t+1} w_t] = E[\mathrm{tr}(P_{t+1} w_t w_t^T)] = \mathrm{tr}(P_{t+1} W_t)$

works out to

$$P_t = A^T P_{t+1} A - A^T P_{t+1} B (B^T P_{t+1} B + R_t)^{-1} B^T P_{t+1} A + Q_t$$

$$= A^T \underbrace{\left( P_{t+1} - P_{t+1} B (B^T P_{t+1} B + R_t)^{-1} B^T P_{t+1} \right)}_{\text{Schur complement}} A + Q_t$$

$$q_t = q_{t+1} + \mathrm{Tr}(W_t P_{t+1})$$

Running sum

optimal policy is linear state feedback

$$\phi_t^*(x_t) = K_t x_t$$

$$K_t = -(B^T P_{t+1} B + R_t)^{-1} B^T P_{t+1} A$$ does not depend on $\Sigma$, $W_0 \cdots W_{T-1}$

optimal cost

$$J^* = E\, V_0(x_0)$$
$$= \mathrm{Tr}(\Sigma P_0) + q_0$$
$$= \mathrm{Tr}(\Sigma P_0) + \sum_{t=0}^{T-1} \mathrm{Tr}(W_t P_{t+1})$$

## · Branch and bound algorithms

methods for global optimization for nonconvex problems

nonheuristic

maintain provable lower and upper bound on global objective value

terminate with certificate proving $\epsilon$-suboptimality

often slow, exponential worst case performance

with luck, can sometimes work well

## · Basic idea

rely on two subroutines that (efficiently) compute a lower and an upper bound of the optimal value over a given region

upper bound can be found by choosing any point in the region, or by local minimization

lower bound can be found by convex relaxation, duality, Lipschitz or other bounds

partition feasible set into convex sets, and find lower / upper bound for each

form global lower and upper bound → if close, quit

↘ else, refine partition and repeat

e.g. $\min\ c^T x$

s.t. $Ax \le b$ $\quad x_i \in \{0,1\}$ $\quad i = 1, 2, \cdots 10$

this gives $2^{10} = 1024$ LPs

when solving one of those LPs with interior point method,

each iteration gives a dual value

if take a pattern (e.g. 1, 1, ⋯), the instant dual value

goes over $f(x_{best})$, quit solving for this pattern