Stochastic Model Predictive Control

- stochastic finite horizon control
- stochastic dynamic programming
- certainty equivalent model predictive control

Causal state-feedback control

- linear dynamical system, over finite time horizon:

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

  - $x_t \in \mathbf{R}^n$ is state, $u_t \in \mathbf{R}^m$ is the input at time t
  - w_t is the process noise (or exogeneous input) at time t

- $X_t = (x_0, \ldots, x_t)$ is the state history up to time t

- causal state-feedback control:

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

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

Stochastic finite horizon control

- $(x_0, w_0, \ldots, w_{T-1})$ is a random variable
- objective: $J = \mathbf{E}\left(\sum_{t=0}^{T-1} \ell_t(x_t, u_t) + \ell_T(x_T)\right)$
  - convex stage cost functions $\ell_t : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, $t = 0, \ldots, T-1$
  - convex terminal cost function $\ell_T : \mathbf{R}^n \to \mathbf{R}$
- J depends on control policies $\phi_0, \ldots, \phi_{T-1}$
- constraints: $u_t \in \mathcal{U}_t$, $t = 0, \ldots, T-1$
  - convex input constraint sets $\mathcal{U}_0, \ldots, \mathcal{U}_{T-1}$
- **stochastic control problem:** choose control policies $\phi_0, \ldots, \phi_{T-1}$ to minimize J, subject to constraints

Stochastic finite horizon control

- an infinite dimensional problem: variables are *functions* $\phi_0, \ldots, \phi_{T-1}$
  - can restrict policies to finite dimensional subspace, $e.g.$, ϕ_t all affine
- key idea: we have **recourse** (a.k.a. feedback, closed-loop control)
  - we can change u_t based on the observed state history $x_0, \ldots, x_t$
  - cf standard ('open loop') optimal control problem, where we commit to $u_0, \ldots, u_{T-1}$ ahead of time
- in general case, need to evaluate J (for given control policies) via Monte Carlo simulation

'Solution' via dynamic programming

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

- $V_T(X_T) = \ell_T(x_T)$; $J^\star = \mathbf{E}\, V_0(x_0)$

- V_t can be found by backward recursion: for $t = T-1, \ldots, 0$

$$V_t(X_t) = \inf_{v \in \mathcal{U}} \{\ell_t(x_t, v) + \mathbf{E}(V_{t+1}((X_t,\ Ax_t + Bv + w_t))|X_t)\}$$

- V_t, $t = 0, \ldots, T$ are convex functions

- optimal policy is causal state feedback

$$\phi_t^\star(X_t) = \operatorname*{argmin}_{v \in \mathcal{U}} \{\ell_t(x_t, v) + \mathbf{E}(V_{t+1}((X_t,\ Ax_t + Bv + w_t))|X_t)\}$$

Independent process noise

- assume $x_0, w_0, \ldots, w_{T-1}$ are independent
- V_t depends only on the current state x_t (and not the state history X_t)
- Bellman equations: $V_T(x_T) = \ell_T(x_T)$; for $t = T-1, \ldots, 0$,

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

- optimal policy is a function of current state x_t

$$\phi^\star(x_t) = \operatorname*{argmin}_{v \in \mathcal{U}} \{ \ell_t(x_t, v) + \mathbf{E}\, V_{t+1}(Ax_t + Bv + w_t) \}$$

Linear quadratic stochastic control

- special case of linear stochastic control
- $\mathcal{U}_t = \mathbf{R}^m$
- $x_0, w_0, \ldots, w_{T-1}$ are independent, with

$$\mathbf{E}\, x_0 = 0, \quad \mathbf{E}\, w_t = 0, \quad \mathbf{E}\, x_0 x_0^T = \Sigma, \quad \mathbf{E}\, w_t w_t^T = W_t$$

- $\ell_t(x_t, u_t) = x_t^T Q_t x_t + u_t^T R_t u_t$, with $Q_t \succeq 0$, $R_t \succ 0$
- $\ell_T(x_T) = x_T^T Q_T x_T$, with $Q_T \succeq 0$

- can show value functions are quadratic, *i.e.*,

$$V_t(x_t) = x_t^T P_t x_t + q_t, \quad t = 0, \ldots, T$$

- Bellman recursion: $P_T = Q_T$, $q_T = 0$; for $t = T-1, \ldots, 0$,

$$\begin{array}{rcl} V_t(z) & = & \inf_v \{ z^T Q_t z + v^T R_t v \\ & & + \mathbf{E}((Az + Bv + w_t)^T P_{t+1}(Az + Bv + w_t) + q_{t+1}) \} \end{array}$$

- works out to

$$\begin{array}{rcl} P_t & = & A^T P_{t+1} A - A^T P_{t+1} B (B^T P_{t+1} B + R_t)^{-1} B^T P_{t+1} A + Q_t \\ q_t & = & q_{t+1} + \mathbf{Tr}(W_t P_{t+1}) \end{array}$$

- optimal policy is linear state feedback: $\phi_t^\star(x_t) = K_t x_t$,

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

(which, strangely, does not depend on Σ, $W_0, \ldots, W_{T-1}$)

- optimal cost

$$\begin{aligned} J^\star &= \mathbf{E}\, V_0(x_0) \\ &= \mathbf{Tr}(\Sigma P_0) + q_0 \\ &= \mathbf{Tr}(\Sigma P_0) + \sum_{t=0}^{T-1} \mathbf{Tr}(W_t P_{t+1}) \end{aligned}$$

Certainty equivalent model predictive control

- at every time t we solve the certainty equivalent problem

$$
\begin{array}{ll}
\text{minimize} & \sum_{\tau=t}^{T-1} \ell_t(x_\tau, u_\tau) + \ell_T(x_T) \\
\text{subject to} & u_\tau \in \mathcal{U}_\tau, \quad \tau = t, \ldots, T-1 \\
& x_{\tau+1} = Ax_\tau + Bu_\tau + \hat{w}_{\tau|t}, \quad \tau = t, \ldots, T-1
\end{array}
$$

  with variables $x_{t+1}, \ldots, x_T$, $u_t, \ldots, u_{T-1}$ and data x_t, $\hat{w}_{t|t}, \ldots, \hat{w}_{T-1|t}$

- $\hat{w}_{t|t}, \ldots, \hat{w}_{T-1|t}$ are predicted values of $w_t, \ldots, w_{T-1}$ based on X_t ($e.g.$, conditional expectations)
- call solution $\tilde{x}_{t+1}, \ldots, \tilde{x}_T$, $\tilde{u}_t, \ldots, \tilde{u}_{T-1}$
- we take $\phi^{\mathrm{mpc}}(X_t) = \tilde{u}_t$
  - ϕ^{mpc} is a function of X_t since $\hat{w}_{t|t}, \ldots, \hat{w}_{T-1|t}$ are functions of X_t

Certainty equivalent model predictive control

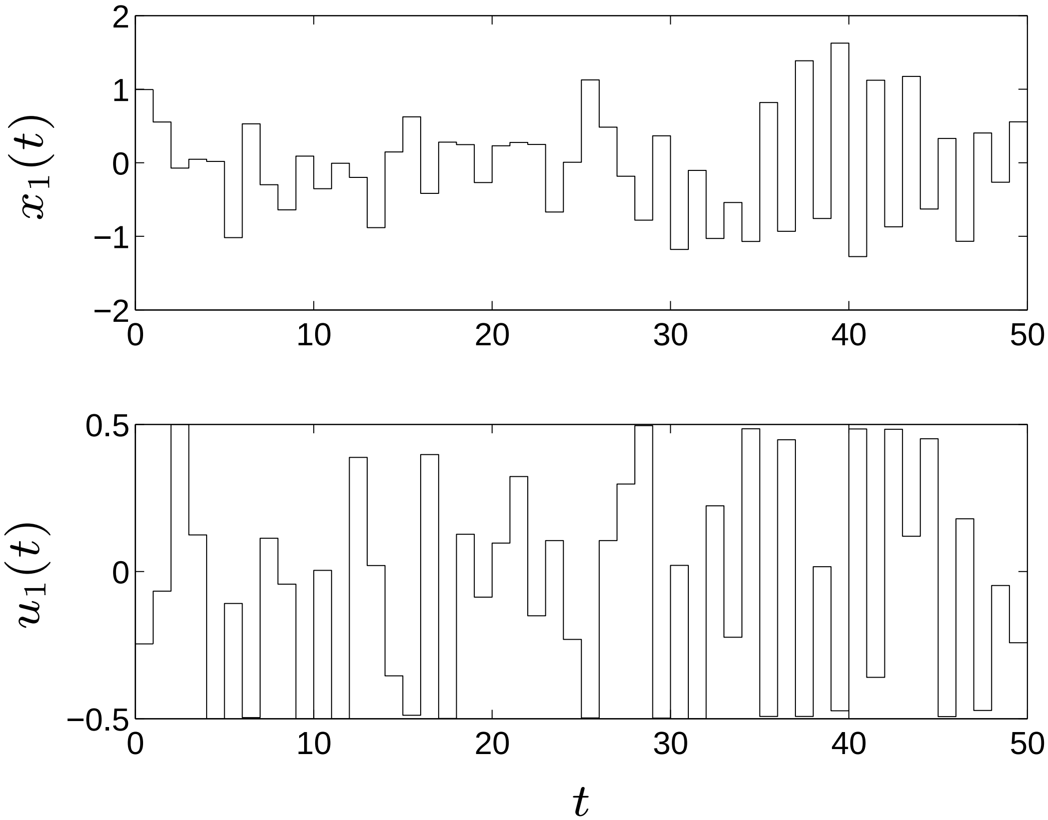
- widely used, *e.g.*, in 'revenue management'
- based on (bad) approximations:
  - future values of disturbance are exactly as predicted; there is no future uncertainty
  - in future, no recourse is available
- yet, often works very well

Example

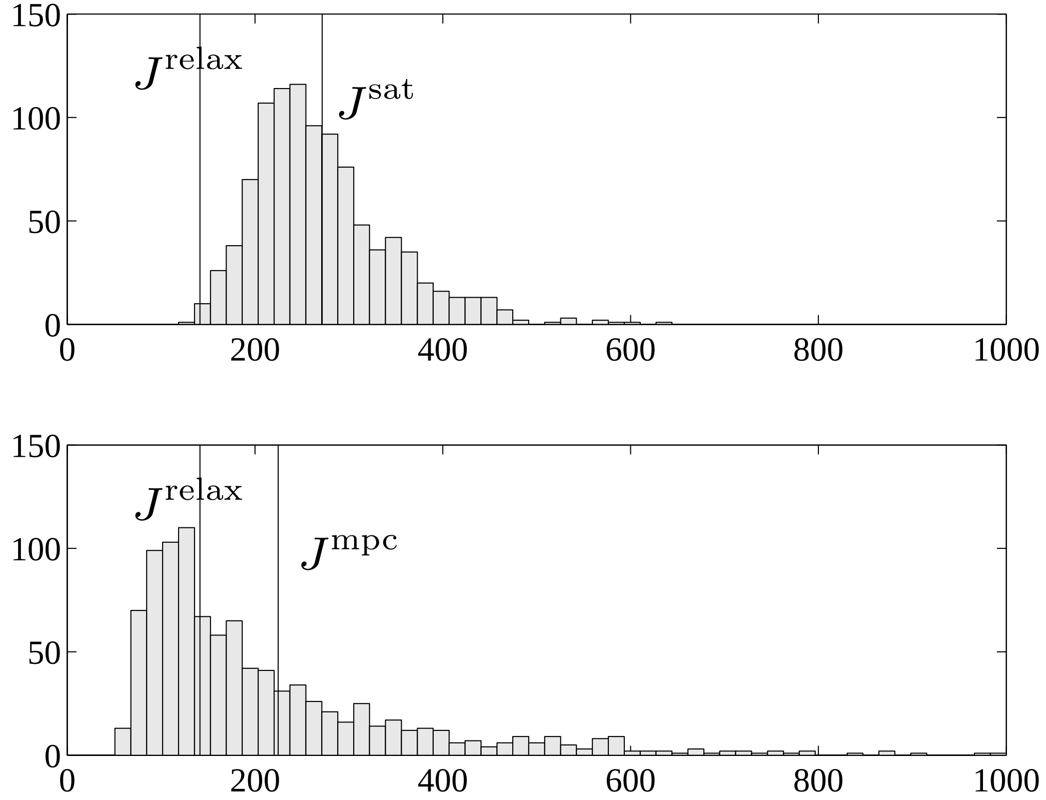
- system with $n = 3$ states, $m = 2$ inputs; horizon $T = 50$
- A, B chosen randomly
- quadratic stage cost: $\ell_t(x, u) = \|x\|_2^2 + \|u\|_2^2$
- quadratic final cost: $\ell_T(x) = \|x\|_2^2$
- constraint set: $\mathcal{U} = \{u \mid \|u\|_\infty \leq 0.5\}$
- $x_0, w_0, \ldots, w_{T-1}$ iid $\mathcal{N}(0, 0.25I)$

Stochastic MPC: Sample trajectory

sample trace of x_1 and u_1



Cost histogram



Simple lower bound for quadratic stochastic control

- $x_0, w_0, \ldots, w_{T-1}$ independent
- quadratic stage and final cost
- relaxation:
  - ignore $\mathcal{U}_t$; yields linear quadratic stochastic control problem
  - solve relaxed problem exactly; optimal cost is $J^{\rm relax}$
- $J^\star \geq J^{\rm relax}$
- for our numerical example,
  - $J^{\rm mpc} = 224.7$ (via Monte Carlo)
  - $J^{\rm sat} = 271.5$ (linear quadratic stochastic control with saturation)
  - $J^{\rm relax} = 141.3$