Performance Bounds for Constrained Linear Stochastic Control

Yang Wang and **Stephen Boyd**

Electrical Engineering Department
Stanford University

Workshop on Optimization and Control with Applications, Harbin, June 9 2009

Outline

- constrained linear stochastic control problem
- some heuristic control schemes
  - projected linear control
  - control-Lyapunov
  - certainty-equivalent model predictive control (MPC)
- the linear quadratic case
- performance bound
- performance bound parameter choices for control schemes
- numerical examples

Linear stochastic system

- linear dynamical system with process noise:

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad t = 0, 1, \ldots,$$

  - $x_t \in \mathbf{R}^n$ is the state
  - $u_t \in \mathcal{U}$ is the control input
  - $\mathcal{U} \subset \mathbf{R}^m$ is the input constraint set, with $0 \in \mathcal{U}$
  - $w_t \in \mathbf{R}^n$ is zero mean IID process noise, $\mathbf{E}\, w_t w_t^T = W$

- state feedback control policy:

$$u_t = \phi(x_t), \quad t = 0, 1, \ldots,$$

$\phi : \mathbf{R}^n \to \mathcal{U}$ is the state feedback function

Objective

- objective is average stage cost:

$$J = \limsup_{T\to\infty} \frac{1}{T} \mathbf{E} \sum_{t=0}^{T-1} (\ell_x(x_t) + \ell_u(u_t))$$

  - $\ell_x : \mathbf{R}^n \to \mathbf{R}$ is state stage cost function
  - $\ell_u : \mathcal{U} \to \mathbf{R}$ is the input state cost function

- ℓ_x, ℓ_u, $\mathcal{U}$ need not be convex

Stochastic control problem

- stochastic control problem: *choose feedback function* ϕ *to minimize* J
- infinite dimensional nonconvex optimization problem
- problem data:
  - dynamics and input matrices A, B
  - distribution of process noise w_t
  - state and input cost functions ℓ_x, ℓ_u
  - input constraint set $\mathcal{U}$
- $\phi^\star$ denotes an optimal feedback function
- $J^\star$ denotes optimal objective value

'Solution' via dynamic programming

- find $V^\star : \mathbf{R}^n \to \mathbf{R}$ and α with

$$V^\star(z) + \alpha = \min_{v \in \mathcal{U}} \left(\ell_u(v) + \mathbf{E}\, V^\star(Az + Bv + w_t) \right)$$

  (expectation is over w_t)

- optimal feedback function is then

$$\phi^\star(z) = \operatorname*{argmin}_{v \in \mathcal{U}} \left(\ell_u(v) + \mathbf{E}\, V^\star(Az + Bv + w_t) \right)$$

- optimal value of stochastic control problem is $J^\star = \alpha$

Stochastic control problem

- generally very hard to solve
  (even more: how would we represent a general function ϕ?)

- can be effectively solved
  - when the problem dimensions are very small, *e.g.*, $n = m = 1$
  - when $\mathcal{U} = \mathbf{R}^m$ and ℓ_x, ℓ_u are convex quadratic;
    in this case optimal policy is linear: $\phi^\star(z) = Kz$

- many suboptimal methods have been proposed
  - can evaluate J for a given ϕ via Monte Carlo simulation
  - but how suboptimal is it?

- this talk: *an effective method for finding a (good) lower bound on* $J^\star$

Projected linear state feedback

- a simple suboptimal policy:

$$\phi_{\rm pl}(z) = \mathcal{P}(K_{\rm pl}z)$$

  - $K_{\rm pl} \in \mathbf{R}^{m\times n}$ is a gain matrix (to be chosen)
  - $\mathcal{P}$ is projection onto $\mathcal{U}$

- when $\mathcal{U}$ is a box, $i.e.$, $\mathcal{U} = \{u \mid \|u\|_\infty \leq U^{\rm max}\}$, reduces to *saturated linear state feedback*

$$\phi_{\rm pl}(z) = U^{\rm max}\,\mathbf{sat}((1/U^{\rm max})K_{\rm pl}z)$$

  $\mathbf{sat}$ is (entrywise) unit saturation

Control-Lyapunov policy

- control-Lyapunov policy is

$$\phi_{\mathrm{clf}}(z) = \underset{v \in \mathcal{U}}{\operatorname{argmin}} \left(\ell_u(v) + \mathbf{E}\, V_{\mathrm{clf}}(Az + Bv + w_t)\right)$$

  – $V_{\mathrm{clf}} : \mathbf{R}^n \to \mathbf{R}$ (which is to be chosen) is the *control-Lyapunov function*
  – when $V_{\mathrm{clf}} = V^\star$, this is optimal policy

- when V_{clf} is quadratic, the control-Lyapunov policy simplifies to

$$\phi_{\mathrm{clf}}(z) = \underset{v \in \mathcal{U}}{\operatorname{argmin}} \left(\ell_u(v) + V_{\mathrm{clf}}(Az + Bv)\right)$$

since $\mathbf{E}\, w_t = 0$, and term involving $\mathbf{E}\, w_t w_t^T = W$ is constant

Certainty-equivalent model predictive control (MPC)

- $\phi_{\mathrm{mpc}}(z)$ is found by solving (possibly approximately)

$$\begin{array}{ll} \text{minimize} & \sum_{\tau=0}^{T-1} (\ell_x(\tilde{x}_\tau) + \ell_u(v_\tau)) + V_{\mathrm{mpc}}(\tilde{x}_T) \\ \text{subject to} & \tilde{x}_{\tau+1} = A\tilde{x}_\tau + Bv_\tau, \quad \tau = 0, \ldots, T-1 \\ & v_\tau \in \mathcal{U}, \quad \tau = 0, \ldots, T-1 \\ & \tilde{x}_0 = z \end{array}$$

  - variables are $v_0, \ldots, v_{T-1}$, $\tilde{x}_0, \ldots, \tilde{x}_T$
  - $V_{\mathrm{mpc}} : \mathbf{R}^n \to \mathbf{R}$ is the terminal cost (to be chosen)
  - T is the planning horizon (also to be chosen)

- let solution be $v_0^\star, \ldots, v_{T-1}^\star$, $\tilde{x}_0^\star, \ldots, \tilde{x}_T^\star$

- MPC policy is $\phi_{\mathrm{mpc}}(z) = v_0^\star$

Parameters in heuristic control policies

- performance of suboptimal policies depends on choice of parameters ($K_{\rm pl}$, $V_{\rm clf}$, $V_{\rm mpc}$ and T)
- one choice for $V_{\rm clf}$, $V_{\rm mpc}$: (quadratic) value function for some unconstrained linear quadratic problem
- one choice for $K_{\rm pl}$: optimal gain matrix for some unconstrained linear quadratic problem
- we will suggest some parameters later . . .

The performance bound

our method:

- computes a lower bound $J^{\mathrm{lb}} \leq J^\star$ using convex optimization (hence is tractable)
- bound is computed for each specific problem instance
- (at this time) cannot guarantee tightness of bound

Judging a heuristic policy

- suppose we have a heuristic policy ϕ with objective J (evaluated by Monte Carlo, say)
- since $J^{\rm lb} \leq J^\star \leq J$, if $J - J_{\rm lb}$ is small, then
  - policy ϕ is nearly optimal
  - bound $J^{\rm lb}$ is nearly tight
- if $J - J^{\rm lb}$ is big, then for this problem instance, either
  - policy is poor, or,
  - bound is poor (or both)
- examples suggest that $J - J^{\rm lb}$ is often small

Unconstrained linear quadratic control

- can effectively solve stochastic control problem when
  - $\mathcal{U} = \mathbf{R}^m$ (no constraints)
  - $\ell_x(z) = z^T Q z$, $\ell_u(v) = v^T R v$, $\quad Q \succeq 0$, $R \succeq 0$
- optimal cost is $J^\star_{\rm lq} = \mathbf{Tr}(P^\star_{\rm lq} W)$
- optimal state feedback function is $\phi^\star(z) = K^\star_{\rm lq} z$, where

$$K^\star_{\rm lq} = -(R + B^T P^\star_{\rm lq} B)^{-1} B^T P^\star_{\rm lq} A$$

- $P^\star_{\rm lq}$ is positive semidefinite solution of ARE

$$P^\star_{\rm lq} = Q + A^T P^\star_{\rm lq} A - A^T P^\star_{\rm lq} B (R + B^T P^\star_{\rm lq} B)^{-1} B^T P^\star_{\rm lq} A$$

Linear quadratic control via LMI/SDP

- can characterize $J^\star_{\rm lq}$ and $P^\star_{\rm lq}$ via the semidefinite program (SDP)

$$\begin{array}{ll} \mbox{maximize} & \mathbf{Tr}(PW) \\ \mbox{subject to} & P \succeq 0 \\ & \left[\begin{array}{cc} R + B^TPB & B^TPA \\ A^TPB & Q + A^TPA - P \end{array} \right] \succeq 0 \end{array}$$

  - variable is P
  - optimal point is $P = P^\star_{\rm lq}$; optimal value is $J^\star_{\rm lq}$

- solution does not depend on W, as long as $W \succ 0$

- constraints are convex in (P, Q, R), so $J^\star_{\rm lq}(Q,R)$ is a *concave* function of (Q,R)

Basic bound

- suppose $Q \succeq 0$, $R \succeq 0$, s satisfy

$$z^T Q z + v^T R v + s \leq \ell_x(z) + \ell_u(v) \quad \text{for all } z \in \mathbf{R}^n,\ v \in \mathcal{U}$$

  i.e., quadratic stage costs are everywhere smaller than $\ell_x + \ell_v$

- then $J_{\mathrm{lq}}^\star(Q, R) + s$ is a lower bound on $J^\star$
- follows from monotonicity of stochastic control cost w.r.t. stage costs
- lefthand side is optimal value of unconstrained quadratic problem

Optimizing the bound

- can optimize the lower bound over Q, R, s by solving

$$\begin{array}{ll} \text{maximize} & J_{\mathrm{lq}}^{\star}(Q,R)+s \\ \text{subject to} & Q \succeq 0, \quad R \succeq 0, \\ & z^TQz + v^TRv + s \leq \ell_x(z) + \ell_u(v) \quad \text{for all } z \in \mathbf{R}^n,\ v \in \mathcal{U} \end{array}$$

- a convex optimization problem
  - objective is concave
  - constraints are convex
  - last constraint is convex in Q, R, s for each z and v
- last constraint is semi-infinite, parameterized by the (infinite) set $z \in \mathbf{R}^n$, $u \in \mathcal{U}$

Optimizing the bound

- semi-infinite constraint makes problem difficult in general
- can solve exactly in a few cases
- in other cases, can replace semi-infinite constraint with conservative approximation, which still gives a lower bound

Quadratic stage cost and finite input set

- can solve optimization problem exactly when
  - $\ell_x(z) = z^T Q_0 z$, $\ell_u(v) = v^T R_0 v$, $\quad Q \succeq 0$, $R \succeq 0$
  - $\mathcal{U} = \{u_1, \ldots, u_K\}$ (finite input constraint set)

- constraint

$$z^T Q z + v^T R v + s \leq \ell_x(z) + \ell_u(v) \quad \text{for all } z \in \mathbf{R}^n,\ v \in \mathcal{U}$$

becomes

$$Q \preceq Q_0, \qquad u_i^T R u_i + s \leq u_i^T R_0 u_i, \quad i = 1, \ldots, K$$

- to optimize the bound we solve SDP (with variables P, Q, R, s)

$$
\begin{array}{ll}
\text{maximize} & \mathbf{Tr}(PW) + s \\
\text{subject to} & P \succeq 0, \quad Q \succeq 0, \quad R \succeq 0, \quad Q \preceq Q_0 \\
& \left[\begin{array}{cc} R + B^T PB & B^T PA \\ A^T PB & Q + A^T PA - P \end{array}\right] \succeq 0 \\
& u_i^T R u_i + s \leq u_i^T R_0 u_i, \quad i = 1, \ldots, K
\end{array}
$$

- monotone in Q, so we can set $Q = Q_0$ w.l.o.g.

S-procedure relaxation

- suppose stage costs are quadratic
- suppose we can find $R_1, \ldots, R_M$ and $s_1, \ldots, s_M$ for which

$$\mathcal{U} \subseteq \tilde{\mathcal{U}} = \{v \mid v^T R_i v + s_i \leq 0,\ i = 1, \ldots, M\}$$

- a sufficient condition for

$$z^T Q z + v^T R v + s \leq \ell_x(z) + \ell_u(v) \text{ for all } z \in \mathbf{R}^n,\ v \in \mathcal{U}$$

is

$$z^T Q z + v^T R v + s \leq z^T Q_0 z + v^T R_0 v \text{ for all } z \in \mathbf{R}^n,\ v \in \tilde{\mathcal{U}}$$

- equivalent to $Q \preceq Q_0$ and

$$v^T R_i v + s_i \le 0,\ i = 1, \ldots, M \quad \Longrightarrow \quad v^T R v + s \le v^T R_0 v$$

- which is implied by $Q \preceq Q_0$ and the existence of $\lambda_1, \ldots, \lambda_M \ge 0$ with

$$R - R_0 \preceq \sum_{i=1}^{M} \lambda_i R_i, \qquad s \le \sum_{i=1}^{M} \lambda_i s_i$$

  (by the S-procedure)

- so $J_{\rm lq}^\star(Q, R) + s$ is a still a lower bound on $J^\star$

- to optimize the bound we solve the SDP

$$
\begin{array}{ll}
\text{maximize} & \mathbf{Tr}(PW) + s_0 \\
\text{subject to} & P \succeq 0, \quad Q \succeq 0, \quad R \succeq 0, \quad Q \preceq Q_0 \\
& \left[\begin{array}{cc} R + B^T P B & B^T P A \\ A^T P B & Q + A^T P A - P \end{array} \right] \succeq 0 \\
& R - R_0 \preceq \sum_{i=1}^M \lambda_i R_i, \quad s_0 \leq \sum_{i=1}^M \lambda_i s_i \\
& \lambda_i \geq 0, \quad i = 1, \ldots, M
\end{array}
$$

with variables P, Q, R, $\lambda_1, \ldots, \lambda_M$, $s_0, \ldots, s_M$

- can set $Q = Q_0$ w.l.o.g.

Suboptimal control policies

- optimizing the lower bound gives $P_{\rm lb}$
- can interpret $\mathbf{Tr}(P_{\rm lb}W)$ as optimal cost of an unconstrained quadratic problem that approximates (and underestimates) our problem
- suggests that

  $$V_{\rm lb}(z) = z^T P_{\rm lb} z,$$

  and

  $$K_{\rm lb} = -(R_{\rm lb} + B^T P_{\rm lb} B)^{-1} B^T P_{\rm lb} A$$

  are good choices of parameters for suboptimal control policies
- examples show this is the case

Numerical examples

- illustrate bounds for 3 examples
  - small problem with trilevel inputs
  - large problem with box constraints
  - discretized mechanical control system
- compare lower bound with various heuristic policies
  - projected linear state feedback
  - model predictive control
  - control-Lyapunov policy

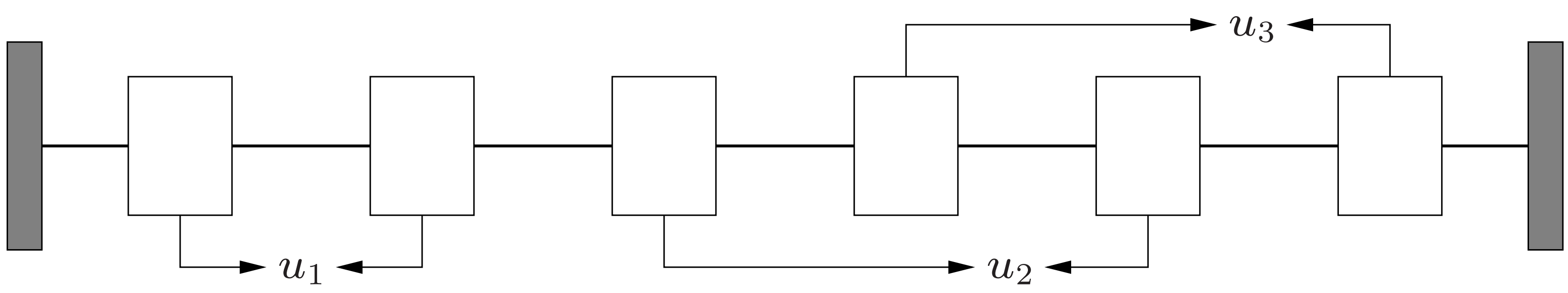
Small problem with trilevel inputs

- $n = 8$, $m = 2$
- A, B matrices randomly generated; A scaled so $\max_i |\lambda_i(A)| = 1$
- quadratic stage costs with $R_0 = I$, $Q_0 = I$
- $w_t \sim \mathcal{N}(0, 0.25I)$
- finite input set: $\mathcal{U} = \{-0.2, 0, 0.2\}^2$

Large problem with box constraints

- $n = 30$, $m = 10$
- A, B matrices randomly generated; A scaled so $\max_i |\lambda_i(A)| = 1$
- quadratic stage costs with $R_0 = I$, $Q_0 = I$
- $w_t \sim \mathcal{N}(0, 0.25I)$
- box input constraints: $\mathcal{U} = \{v \in \mathbf{R}^m \mid \|v\|_\infty \leq 0.1\}$

Discretized mechanical control system



- 6 masses connected by springs; 3 input tensions between masses
- quadratic stage costs with $R_0 = I$, $Q_0 = I$
- w_t uniform on $[-0.5, 0.5]$
- box input constraints: $\mathcal{U} = \{v \in \mathbf{R}^m \mid \|v\|_\infty \leq 0.1\}$

Heuristic policies

- projected linear state feedback with $K_{\rm pl} = K_{\rm lq}^\star$
- control-Lyapunov policy with $V_{\rm clf}(z) = z^T P_{\rm lb} z$
- model predictive control (MPC) with $T = 30$, $V_{\rm mpc}(z) = z^T P_{\rm lb} z$

  (for trilevel example we solve convex relaxation with $u(t) \in [-0.2, 0.2]$, then round value to $\{-0.2, 0, 0.2\}$)

Results

	small trilevel	large random	masses
PLSF	12.9	31.3	269.8
CLF	10.8	25.6	61.1
MPC	10.9	25.7	58.9
J^{lb}	9.1	23.8	43.2

- control-Lyapunov with P_{lb} and MPC achieve similar performance
- control-Lyapunov policy can be computed *very* fast (in tens of microseconds); MPC policy can be computed in milliseconds
- bound J_{lb} is reasonably close to J for these examples

Conclusions

- we've shown how to find lower bounds on optimal performance for constrained linear stochastic control problems
- requires solution of convex optimization problem, hence is tractable
- provides only provable lower bound on optimal performance that we are aware of
- as a by-product, provides excellent choice for quadratic control-Lyapunov function
- in many cases, gives everything you want:
  - a provable lower bound on performance
  - a relatively simple heuristic policy that comes close

References

- Wang & Boyd, *Performance Bounds for Linear Stochastic Control*, Systems Control Letters 58(3):178–182, 2008
- Wang & Boyd, *Performance Bounds and Suboptimal Policies for Linear Stochastic Control via LMIs*, (manuscript)
- Cogill & Lall, *Suboptimality Bounds in Stochastic Control: A Queueing Example*, Proceedings ACC p1642–1647, 2006

  (similar results in MDP setting)
- Lincoln & Rantzer, *Relaxing Dynamic Programming*, IEEE T-AC 51(8):1249-2006, 2006