Second-Order Cone Programming

Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, Hervé Lebret

July 15 1997

Outline

- second-order cone programming (SOCP)
- examples
  - sum and maximum of norms
  - hyperbolic constraints (log-Chebychev, matrix-fractional)
  - robust LP
  - robust least-squares
  - robust QP
- applications
  - antenna array weight design
  - robust (finite horizon) optimal control

Second-order cone (SOC) $\diamond$

- Standard second-order cone of dimension k

$$\mathcal{C}_k = \left\{ \left[\begin{array}{c} u \\ t \end{array} \right] \;\middle|\; u \in \mathbf{R}^{k-1},\ t \in \mathbf{R},\ \|u\| \leq t \right\}$$

(also quadratic, ice-cream, or Lorentz cone)

- Second-order cone constraint of dimension k

$$\|Ax + b\| \leq c^T x + d \quad \Longleftrightarrow \quad \left[\begin{array}{c} A \\ c^T \end{array} \right] x + \left[\begin{array}{c} b \\ d \end{array} \right] \ \in\ \mathcal{C}_k,$$

variable: $x \in \mathbf{R}^n$

parameters: $A \in \mathbf{R}^{(k-1)\times n}$, $b \in \mathbf{R}^{k-1}$, $c \in \mathbf{R}^n$, $d \in \mathbf{R}$

Second-order cone program

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\| \leq c_i^T x + d_i, \quad i = 1, \ldots, L \end{array}$$

$A_i \in \mathbf{R}^{(n_i-1)\times n}$, optimization variable $x \in \mathbf{R}^n$

- convex, nondifferentiable problem
- includes LP, QP as special cases
- many IP methods, most LP and SDP methods are easily adapted
  (Nesterov & Nemirovsky, Lobo & al., Alizadeh & al., Andersen & Andersen, Xue & Ye, etc.)
- implementations available (SOCP, SDPPACK)

Relation to semidefinite programming (SDP)

- SOCP can be solved via SDP

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \geq 0, \quad i = 1, \ldots, L \end{array}$$

- solving SOCP is more efficient than SDP, $O(\sqrt{L})$ vs. $O(\sqrt{\sum_{i=1}^{L} n_i})$

The dual SOCP

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^{L} \left(b_i^T z_i + d_i w_i \right) \\ \text{subject to} & \|z_i\| \le w_i, \qquad i = 1, \ldots, L \\ & \sum_{i=1}^{L} \left(A_i^T z_i + c_i w_i \right) = f \end{array}$$

optimization variables $z_i \in \mathbf{R}^{n_i - 1}$, $w \in \mathbf{R}^L$

- also an SOCP
- primal and dual optimal values are equal if, $e.g.$, both problems are strictly feasible

Primal-dual formulation ⋄

- the difference between the primal and dual objectives is called the *duality gap* associated with x, z, w

$$\eta(x,z,w) \;=\; f^T x + \sum_{i=1}^{L} \left(b_i^T z_i + d_i w_i\right) \;=\; \sum_{i=1}^{L} \left(z_i^T u_i + w_i t_i\right)$$

  $\eta(x,z,w) \geq 0$ *(weak duality)*

  if the primal and dual problems are strictly feasible, then there exist primal and dual feasible points such that $\eta(x,z,w) = 0$

- primal-dual problem

$$\begin{array}{ll} \text{minimize} & \eta(x,z,w) \\ \text{subject to} & \textit{(primal and dual constraints)} \end{array}$$

Sum of norms

$$\text{minimize} \quad \sum_{i=1}^{p} \|A_i x + b_i\| \qquad\qquad (A_i \in \mathbf{R}^{n_i \times n})$$

can express as an SOCP with auxiliary variables $t_1, \ldots, t_p$

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{p} t_i \\ \text{subject to} & \|A_i x + b_i\| \leq t_i, \quad i = 1, \ldots, p \end{array}$$

- we can easily incorporate other second-order cone constraints in the problem, *e.g.*, linear inequalities on x

Example: complex ℓ_1-norm approximation

$$\text{minimize} \quad \|Ax - b\|_1 \qquad (x \in \mathbf{C}^q,\ A \in \mathbf{C}^{p\times q},\ b \in \mathbf{C}^p)$$

(ℓ_1-norm on $\mathbf{C}^p$ is $\|v\|_1 = \sum_{i=1}^p |v_i|$)

can write as an SOCP with p constraints of dimension three

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^p t_i \\ \text{subject to} & \\ & \left\| \begin{bmatrix} \mathrm{Re}\, a_i^T & -\mathrm{Im}\, a_i^T \\ \mathrm{Im}\, a_i^T & \mathrm{Re}\, a_i^T \end{bmatrix} z - \begin{bmatrix} \mathrm{Re}\, b_i \\ \mathrm{Im}\, b_i \end{bmatrix} \right\| \leq t_i, \quad i = 1, \ldots, p \end{array}$$

in the variables $z = [\mathrm{Re}\, x^T \ \mathrm{Im}\, x^T]^T \in \mathbf{R}^{2q}$, and $t \in \mathbf{R}^p$

Maximum of norms

$$\text{minimize} \quad \max_{i=1,\ldots,p} \|A_i x + b_i\| \qquad\qquad (A_i \in \mathbf{R}^{n_i \times n})$$

can express as an SOCP with auxiliary variable t

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \|A_i x + b_i\| \leq t, \quad i = 1, \ldots, p \end{array}$$

- we can easily incorporate other second-order cone constraints in the problem, $e.g.$, linear inequalities on x

Example: complex ℓ_∞-norm approximation

$$\text{minimize} \quad \|Ax - b\|_\infty \qquad\qquad (x \in \mathbf{C}^q,\ A \in \mathbf{C}^{p\times q},\ b \in \mathbf{C}^p)$$

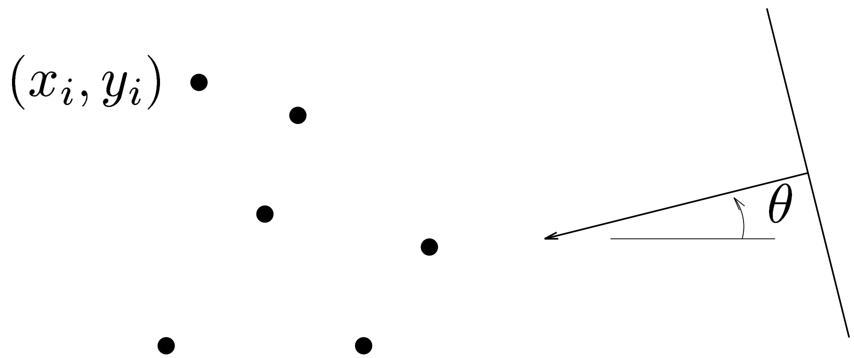
(ℓ_∞-norm on $\mathbf{C}^p$ is $\|v\|_\infty = \max_{i=1,\ldots,p} |v_i|$)

can write as an SOCP with p constraints of dimension three

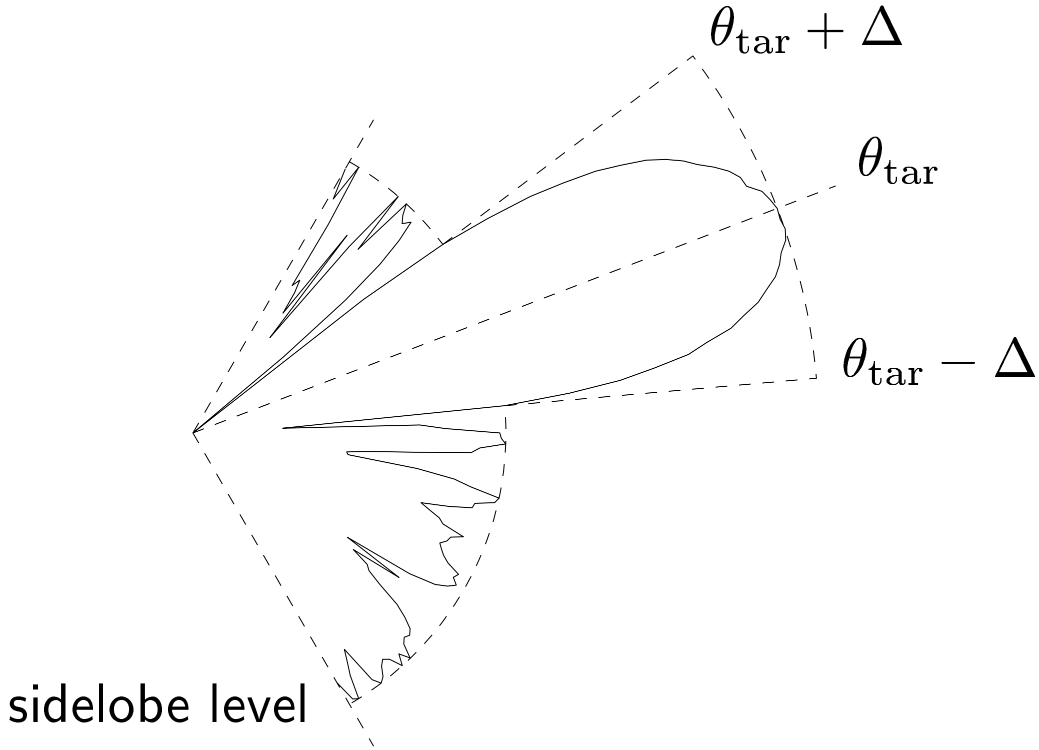
$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \\ & \left\| \begin{bmatrix} \text{Re}\, a_i^T & -\text{Im}\, a_i^T \\ \text{Im}\, a_i^T & \text{Re}\, a_i^T \end{bmatrix} z - \begin{bmatrix} \text{Re}\, b_i \\ \text{Im}\, b_i \end{bmatrix} \right\| \leq t, \quad i = 1, \ldots, p \end{array}$$

in the variables $z = [\text{Re}\, x^T\ \text{Im}\, x^T]^T \in \mathbf{R}^{2q}$, and $t \in \mathbf{R}$

Application: antenna array weight design



- antenna weights $w_1, \ldots, w_n \in \mathbf{C}$
- combined output $f(\theta) \in \mathbf{C}$: $f(\theta) = \sum_{i=1}^{n} w_i \, e^{j(x_i \cos\theta + y_i \sin\theta)}$



$$\begin{array}{ll} \text{minimize} & \max_{|\theta - \theta_{\rm tar}| > \Delta} |f(\theta)| \\ \text{subject to} & f(\theta_{\rm tar}) = 1 \end{array}$$

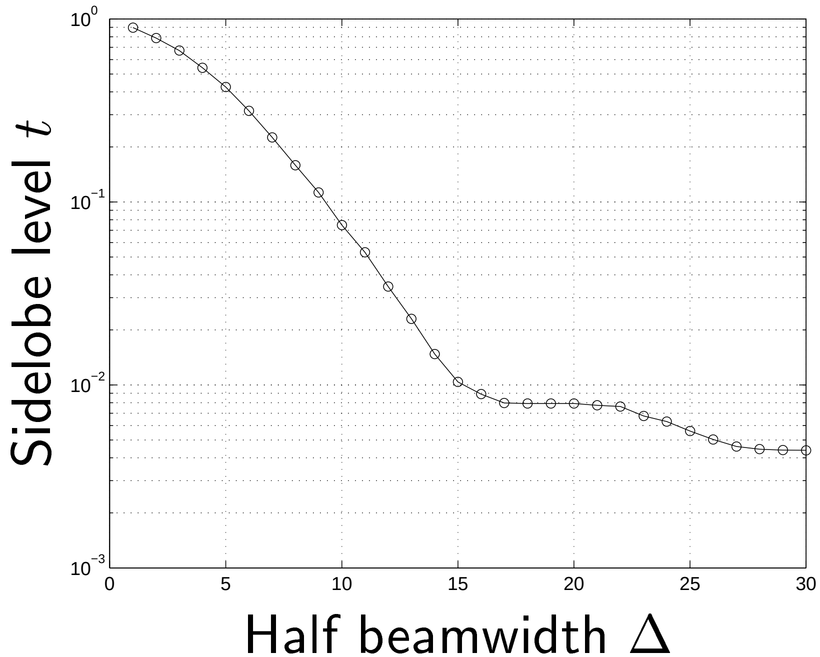
discretize θ, approximate problem (with $\theta_{\rm tar} = \theta_0$)

$$\begin{array}{ll} \mbox{minimize} & t \\ \mbox{subject to} & |f(\theta_i)| \leq t, \quad \mbox{for } |\theta_i - \theta_0| > \Delta \\ & f(\theta_0) = 1 \end{array}$$

- eliminate equality, complex ℓ_∞-norm approximation problem
- using real and imaginary parts of the variables and data: an SOCP

numerical example

- data from field measurements, 8 antenna elements, 121 angles ($\pm 60^\circ$)
- compute tradeoff curve



Hyperbolic constraints

half hyperboloid

$$\{\ (w, x, y) \in (\mathbf{R}^n, \mathbf{R}, \mathbf{R})\ \mid\ w^T w \leq xy,\ \ x, y \geq 0\ \}$$

- convex
- can represent as second-order cone constraint

$$w^T w \leq xy,\ \ x, y \geq 0 \quad \Longleftrightarrow \quad \left\| \begin{bmatrix} 2w \\ x - y \end{bmatrix} \right\| \leq x + y$$

Logarithmic Chebychev approximation

fact

$$|\log(a_i^T x) - \log(b_i)| = \log\max(\, a_i^T x / b_i \,,\, b_i / a_i^T x \,)$$

problem

$$\text{minimize} \quad \max_i |\log(a_i^T x) - \log(b_i)|$$

can be written as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & 1/t \leq a_i^T x / b_i \leq t, \quad i = 1, \ldots, p \end{array}$$

- hyperbolic and linear constraints $\Longrightarrow$ SOCP
- solve $Ax \approx b$ measuring the error by the maximum logarithmic deviation

Matrix-fractional programming

$$\begin{array}{ll} \text{minimize} & (Fx+g)^T (P_0 + x_1 P_1 + \cdots + x_p P_p)^{-1} (Fx+g) \\ \text{subject to} & P_0 + x_1 P_1 + \cdots + x_p P_p > 0, \quad x \geq 0 \end{array}$$

$P_i = P_i^T \in \mathbf{R}^{n \times n}$, variable $x \in \mathbf{R}^p$

- can solve with SDP, for any P_i
- for $P_i \geq 0$, can write as

$$\begin{array}{ll} \text{minimize} & t_0 + t_1 + \cdots + t_p \\ \text{subject to} & P_0^{1/2} y_0 + P_1^{1/2} y_1 + \cdots + P_p^{1/2} y_p = Fx + g \\ & y_0^T y_0 \leq t_0 \\ & y_i^T y_i \leq t_i x_i, \quad t_i,\ x_i \geq 0, \quad i = 1, \ldots, p \end{array}$$

Robust linear program

- c and b_i fixed, a_i uncertain

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \ \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array}$$

- the uncertainty in a_i is described by an ellipsoid

$$\mathcal{E}_i = \{\overline{a}_i + P_i u \mid \|u\| \leq 1\}$$

with $P_i = P_i^T \geq 0$

Robust LP as an SOCP

- note that
$$\max\{ a_i^T x \mid a_i \in \mathcal{E}_i \} = \overline{a}_i^T x + \|P_i x\|$$
and therefore
$$a_i^T x \le b_i \text{ for all } a_i \in \mathcal{E}_i \quad \Longleftrightarrow \quad \overline{a}_i^T x + \|P_i x\| \le b_i$$

- the robust LP can be expressed as the SOCP
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|P_i x\| \le -\overline{a}_i^T x + b_i, \quad i = 1, \ldots, m \end{array}$$

- the additional norm terms act as 'regularization terms', discouraging large x in directions with considerable uncertainty in the parameters a_i

Probabilistic interpretation of robust LP

- a_i is an R.V. with Gaussian distribution, mean $\overline{a}_i$, variance Σ_i
- we want to solve (for $\pi_i \geq 0.5$)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \text{Prob}\{a_i^T x \leq b_i\} \geq \pi_i, \quad i = 1, \ldots, m \end{array}$$

- rewrite the constraint (Φ is the c.d.f. of standard Gausian)

$$\text{Prob}\{a_i^T x \leq b_i\} \geq \pi_i \quad \Longleftrightarrow \quad b_i - \overline{a}_i^T x \geq \Phi^{-1}(\pi_i) \|\Sigma_i^{1/2} x\|$$

r.h.s. is a second-order cone constraint

- equivalent to deterministic case with $P_i = \Phi^{-1}(\pi_i)\, \Sigma_i^{1/2}$

Robust least-squares (ball uncertainty) $\diamond$

$$\text{minimize} \quad \max_{\|\delta A\| \le \rho,\ \|\delta b\| \le \xi} \|(A + \delta A)x - (b + \delta b)\|$$

overdetermined set of equations with uncertainty in coefficients, minimize largest possible residual

- equivalent to minimizing a sum of norms

$$\max_{\|\delta A\| \le \rho,\ \|\delta b\| \le \xi} \|(A + \delta A)x - (b + \delta b)\| \;=\; \|Ax - b\| + \rho\|x\| + \xi$$

- can be more efficiently solved by other methods (SVD), SOCP becomes interesting when we add other constraints, $e.g.$, nonnegativity of x

Robust least-squares

- assume that a_i, the rows of A, have independent errors ($e.g.$, they correspond to different time samples)

- rows of A lie within an ellipsoid: $a_i \in \mathcal{E}_i$, where

$$\mathcal{E}_i = \{\overline{a}_i + P_i u \mid \|u\| \leq 1\} \qquad (P_i = P_i^T > 0)$$

- problem: minimize the worst-case residual

$$\text{minimize} \quad \max_{a_i \in \mathcal{E}_i} \left(\sum_{i=1}^{n} \left(a_i^T x - b_i\right)^2 \right)^{1/2}$$

- work out the objective function in a closed form, and the problem can be formulated as

$$\text{minimize} \quad \left(\sum_{i=1}^{n} \left(\left| \overline{a}_i^T x - b_i \right| + \|P_i x\| \right)^2 \right)^{1/2}$$

which can be cast as the SOCP

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & \|t\| \leq s \\ & u_i + \|P_i x\| \leq t_i, \quad i = 1, \ldots, n \\ & \left| \overline{a}_i^T x - b_i \right| \leq u_i, \quad i = 1, \ldots, n \end{array}$$

- easily extended for uncertainty in b, can add other constraints

Robust quadratic programming

- P lies within the ellipsoid $\mathcal{E} = \left\{ P_0 + \sum_{i=1}^{m} P_i u_i \;\middle|\; \|u\| \leq 1 \right\}$

- robust QP: minimize $\max_{P \in \mathcal{E}} \; x^T P \, x + 2q^T x + r$

  (same ideas apply for quadratic constraints, *i.e.*, QCQP)

- work out objective in a closed form by noting that

$$\max_{\|u\| \leq 1} \sum_{i=1}^{m} x^T P_i \, x \, u_i \quad = \quad \left(\sum_{i=1}^{m} \left(x^T P_i \, x \right)^2 \right)^{1/2}$$

- the problem can be cast as

$$\begin{array}{ll} \text{minimize} & t + v + 2q^T x + r \\ \text{subject to} & \|u\| \leq t \\ & x^T P_i x \leq u_i, \qquad i = 1, \ldots, m \\ & x^T P_0 x \leq v \end{array}$$

which is equivalent to the SOCP

$$\begin{array}{ll} \text{minimize} & t + v + 2q^T x + r \\ \text{subject to} & \|u\| \leq t \\ & \left\| \begin{bmatrix} 2P_i^{1/2} x \\ u_i - 1 \end{bmatrix} \right\| \leq u_i + 1, \quad u_i \geq 0, \quad i = 1, \ldots, m \\ & \left\| \begin{bmatrix} 2P_0^{1/2} x \\ v - 1 \end{bmatrix} \right\| \leq v + 1, \quad v \geq 0 \end{array}$$

Application: (finite horizon) optimal control and LP $\diamond$

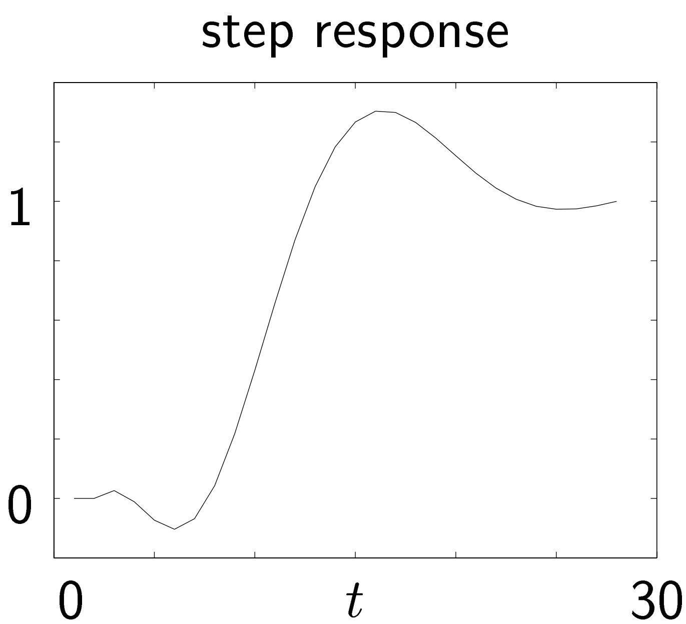
with: C. Crusius, A. Hansson

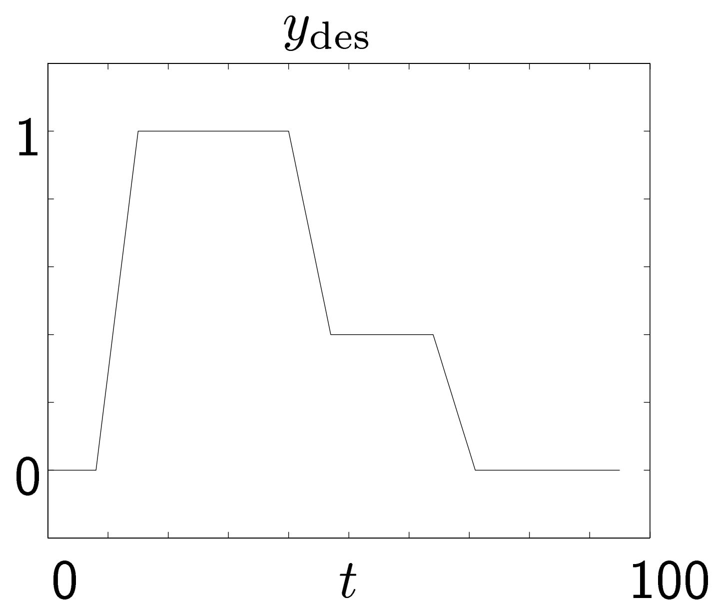
- discrete-time linear system, input $u(t) \in \mathbf{R}$, output $y(t) \in \mathbf{R}$
- optimization variable: input sequence $u = (u(1), \ldots, u(N))$
- desired output sequence: $y_{\rm des}(1), \ldots, y_{\rm des}(M)$

minimize	$E = \max_{t=1}^{M} \lvert y(t) - y_{\rm des}(t) \rvert$	(peak tracking error)
subject to	$\lvert u(t) \rvert \leq U$	(limit on input amplitude)
	$\lvert u(t+1) - u(t) \rvert \leq S$	(limit on input slew rate)

- can be cast as LP, hence efficiently solved

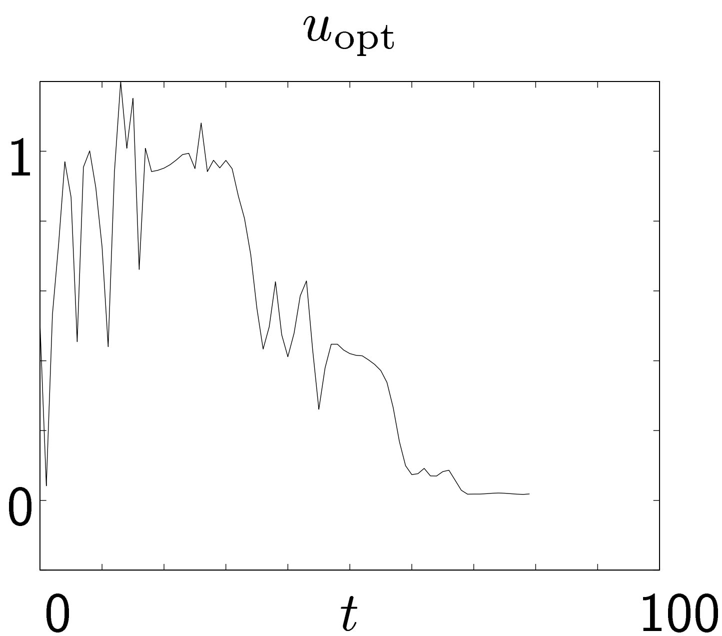
Example ◇

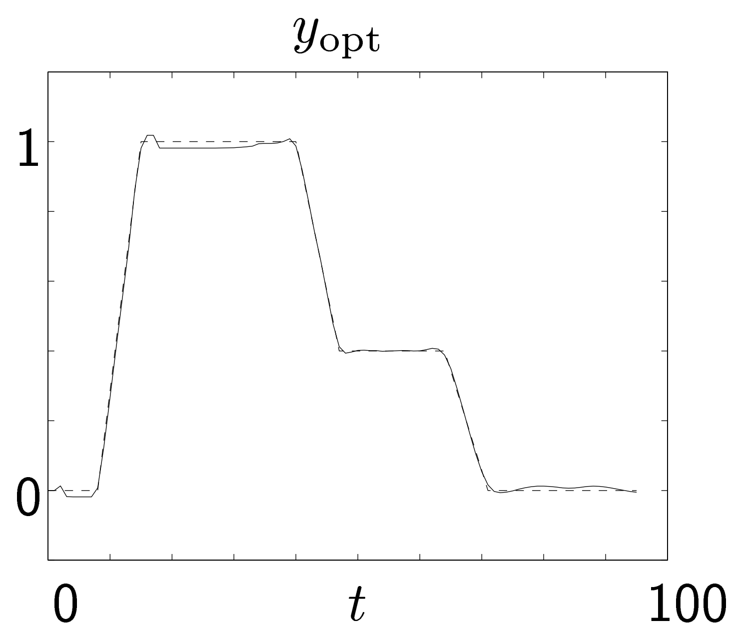




- input amplitude limit: $U = 1.2$
- input slew rate limit: $S = 0.5$

Optimal input and resulting output ◇





- optimal tracking error $.019$
- amplitude & slew rate limit are active

optimal control via LP (or QP)

- can handle multiple inputs/outputs, many other constraints
- used in predictive (receding horizon) control
- can exploit problem structure (sparsity, state equations) to improve efficiency

Using robust LP: robust optimal control $\diamond$

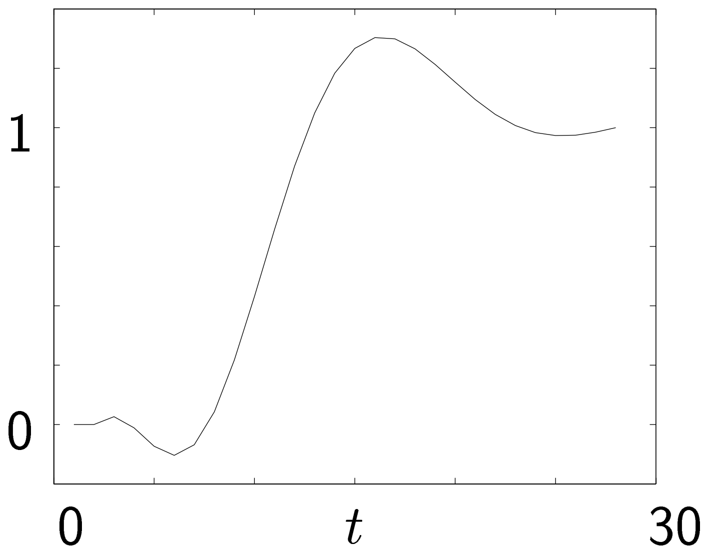
- plant uncertainty described by ellipsoid of step responses: $s \in \mathcal{S}$
- $\mathcal{S}$ might come from measurements, sys ID confidence bounds . . .

$$\begin{array}{ll} \text{minimize} & E_{\rm wc} = \max_{s \in \mathcal{S}} \max_{t=1}^{M} |y(t) - y_{\rm des}(t)| \\ \text{subject to} & |u(t)| \leq U \\ & |u(t+1) - u(t)| \leq S \end{array}$$

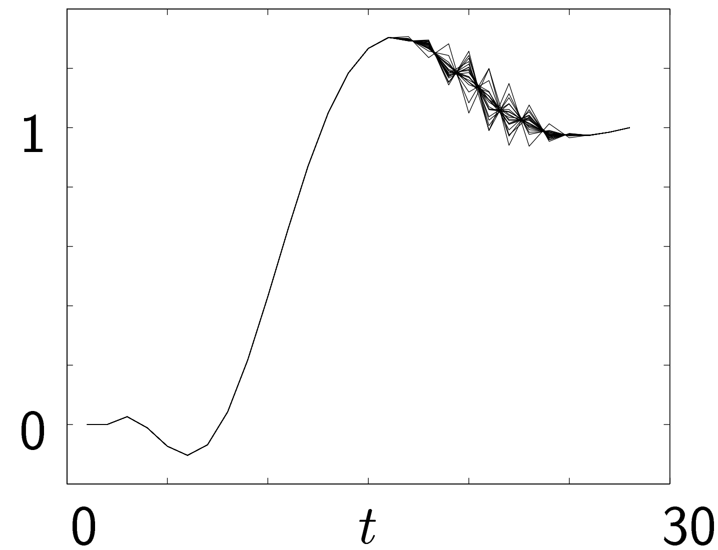
- objective is **worst-case, peak tracking error**
- can be cast as robust LP, hence efficiently solved as SOCP

Robust optimal control: example ◇

nominal step response

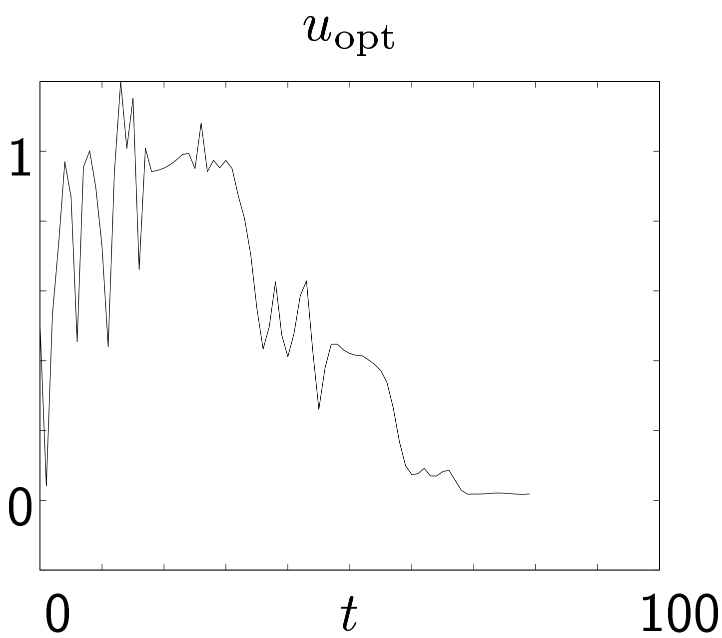


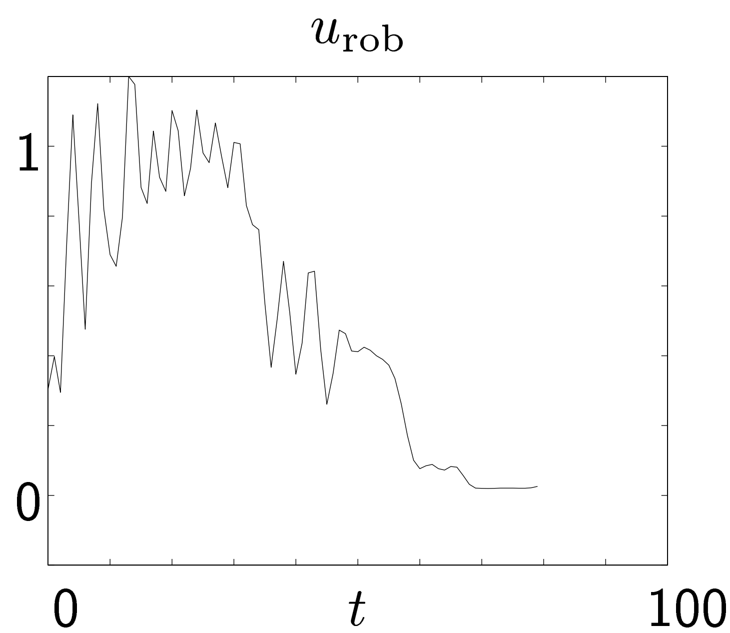
typical $s \in \mathcal{S}$



- same amplitude and slew rate limits on u

Nominal and robust optimal inputs ⋄





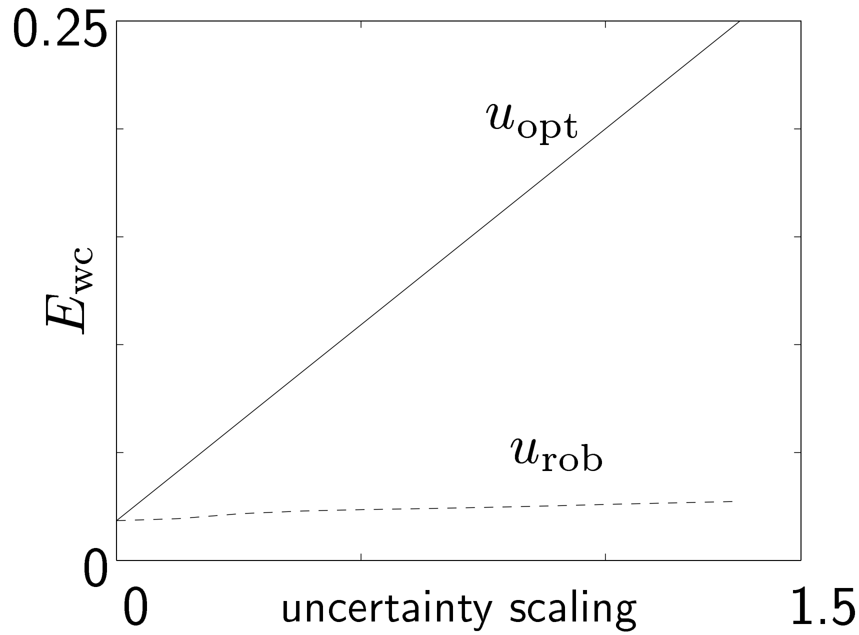
- nominal & robust optimal inputs don't look too different . . .

Resulting tracking errors ◇

	u_{opt}	u_{rob}
E (nominal plant)	.019	.026
E_{wc} (worst-case plant)	.200	.026

- u_{rob} has **worse performance** with nominal plant
- but, **far better robustness** to plant variations

Performance / robustness trade-off ◇



- scale plant uncertainty by scaling uncertainty ellipsoid
- compare nominal optimal $u_{\rm opt}$ and robust optimal $u_{\rm rob}$ for each uncertainty level

Conclusions

Second-order cone programming:

- convex, efficiently solved, lies between LP and SDP
- interior-point methods are easily generalized from LP (or specialized from SDP) to SOCP
- implementations available
- flexible formulation, many problems
- quadratic & hyperbolic constraints, many engineering applications
- robustness problems (LP, QP, least-squares)

References

- *http://www-isl.stanford.edu/people/boyd/socp-talk.html*
- M. S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, Second-order cone programming, *submitted to Linear Algebra and its Applications*, 1997
- Yu. Nesterov and A. Nemirovsky, Interior-point polynomial methods in convex programming, *SIAM, Philadelphia, PA*, 1994
- SOCP, M. S. Lobo, L. Vandenberghe, S. Boyd, *http://www-isl.stanford.edu/people/boyd/SOCP.html*
- SDPPACK, F. Alizadeh, J.-P. A. Haeberly, M. V. Nayakkankuppam, M. L. Overton, S. Schmieta, *http://cs.nyu.edu/cs/faculty/overton/sdppack/sdppack.html*