A Robust optimization formulation for general nonlinear programming

Yin Zhang, Rice University
(DRAFT)

Large Scale Robust Optimization Workshop
August 31, 2005, Santa Fe

Abstract

For optimal design or control problems involving safety constraints and parameter uncertainty, a central task is to select safety margins. Robust optimization provides a natural mechanism for this task. We propose a general robust formulation for nonlinear programming, and discuss the advantages and limitations of this formulation.

An ODE Control Example

$$\begin{aligned} \min_{y,u\,\in\mathcal{U}} \quad & \left\| y(T) - y^{end} \right\|_2^2 \\ \text{s.t.} \quad & \dot{y} = f(y,u,s), \qquad y(0) = y^{ini}, \\ & y_i(t) \le \gamma, \qquad t \in (0,T) \end{aligned}$$

— $y(t)$ is state, $u(t)$ is control and s is a parameter vector.

— Parameter s contains uncertainty.

— There are important (inequality) “safety constraints”.

– The system and constraints are generally nonlinear.

Task: Optimize while maintaining safety for all “reasonable” s values.

Robust Optimization: Existing Results and Need for Extension

- Robust optimization has been actively studied for linear conic programming: LP, SOCP, SDP (Ben-Tal & Nemirovski, El Ghaoui, ...).

- Data uncertainty are restricted to cases where

  – every uncertain constraint is a linear inequality

  – every uncertain parameter appears linearly

  – e.g. LP: $Ax \leq b$ where (A, b) is uncertain (no $A(s)x \leq b(s)$)

- The applicability of the existing formulations needs to be extended.

- Our extension uses simple ideas: implicit function theorem, linearization

General Nonlinear Programming

$$\begin{aligned} \min_{y,u\,\in\mathcal{U}} \quad & \phi(y,u,s) \\ \text{s.t.} \quad & F(y,u,s) = 0 \\ & G(y,u,s) \le 0. \end{aligned}$$

- $y =$ state variable; $u =$ design/control variable; $s =$ system parameter.
- Assume that state equation $F = 0$ implicitly defines $y = y(u,s)$.
- Assume that a good parameter estimation $\hat{s}$ is available.
- Assume that all inequalities are “safety constraints”.

A Look at A Robust Model

$$
\begin{array}{ll}
\min\limits_{y,u\,\in\mathcal{U}} & \phi(y,u,\hat{s}) \\
\text{s.t.} & F(y,u,\hat{s}) = 0 \\
 & (F_y y_s + F_s)(y,u,\hat{s}) = 0 \\
 & G(y,u,s)e^T \pm \tau(G_y y_s + G_s) \le 0.
\end{array}
$$

A nonlinear program involving only $\hat{s}$, new unknown y_s, $e^T = (1 \cdots 1)$.

$$
\begin{array}{lcl}
\min & \left\| y(T) - y^{end} \right\|_2^2 & \\
\text{s.t.} & \dot{y} = f(y,u,\hat{s}), & y(0) = y^{ini}, \\
 & \dot{y}_s = f_y(y,u,\hat{s})y_s + f_s(y,u,\hat{s}), & y_s(0) = 0, \\
 & y(t)e^T \pm \tau y_s(t) \le 0, & t \in (0,T)
\end{array}
$$

Robust Model Derivation: First Step

We start from the Inequality-only case:

$$\begin{array}{lll} \min & \phi(u) & \\ \text{s.t.} & h_i(u,s) \le 0, \quad \forall s \in \mathcal{S} & \forall i \end{array}$$

$$\Updownarrow$$

$$\max\{h_i(u,s) : s \in \mathcal{S}\} \le 0, \quad \forall i$$

- Can we handle this comfortably?

$$\max_{s \in \mathcal{S}} h_i(u,s)$$

- Maybe too hard for general $\mathcal{S}$ and $H = (h_1 \cdots, h_m)^T$.

Restrict $\mathcal{S}$ and Linearize H

- For $\tau > 0$, define
$$\mathcal{S}_\tau = \{\hat{s} + \tau D\delta : \|\delta\|_p \le 1\}$$
$\hat{s}$ is the nominal parameter value, δ the unit variation (in p-norm) and τ the magnitude of the variation. D is a scale or basis matrix for the variations.

- By Taylor approximation,
$$h_i(u, \hat{s} + \tau D\delta) \approx h_i(u, \hat{s}) + \tau \langle \nabla_s h_i(u, \hat{s}), D\delta \rangle.$$

- Hence,
$$\max_{s \in S_\tau} h_i(u, s) \approx h_i(u, \hat{s}) + \tau \max_{\|\delta\|_p = 1} \langle D^T \nabla_s h_i(u, \hat{s}), \delta \rangle$$

Solution of Linearized Maximization

Recall

$$|\langle c, x\rangle| \le \|x\|_p \|c\|_q \quad \text{for} \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 \le p, q \le +\infty,$$

and the equality can be achieved, i.e.,

$$\max_{\|x\|_p = 1} \langle c, x\rangle = \|c\|_q.$$

- Therefore,

$$\begin{aligned} \max_{s \in S_\tau} h_i(u, s) \quad &\approx \quad h_i(u, \hat{s}) + \tau \max_{\|\delta\|_p = 1} \langle D^T \nabla_s h_i(u, \hat{s}), \delta\rangle \\ &= \quad h_i(u, \hat{s}) + \tau \, \|D^T \nabla_s h_i(u, \hat{s})\|_q. \end{aligned}$$

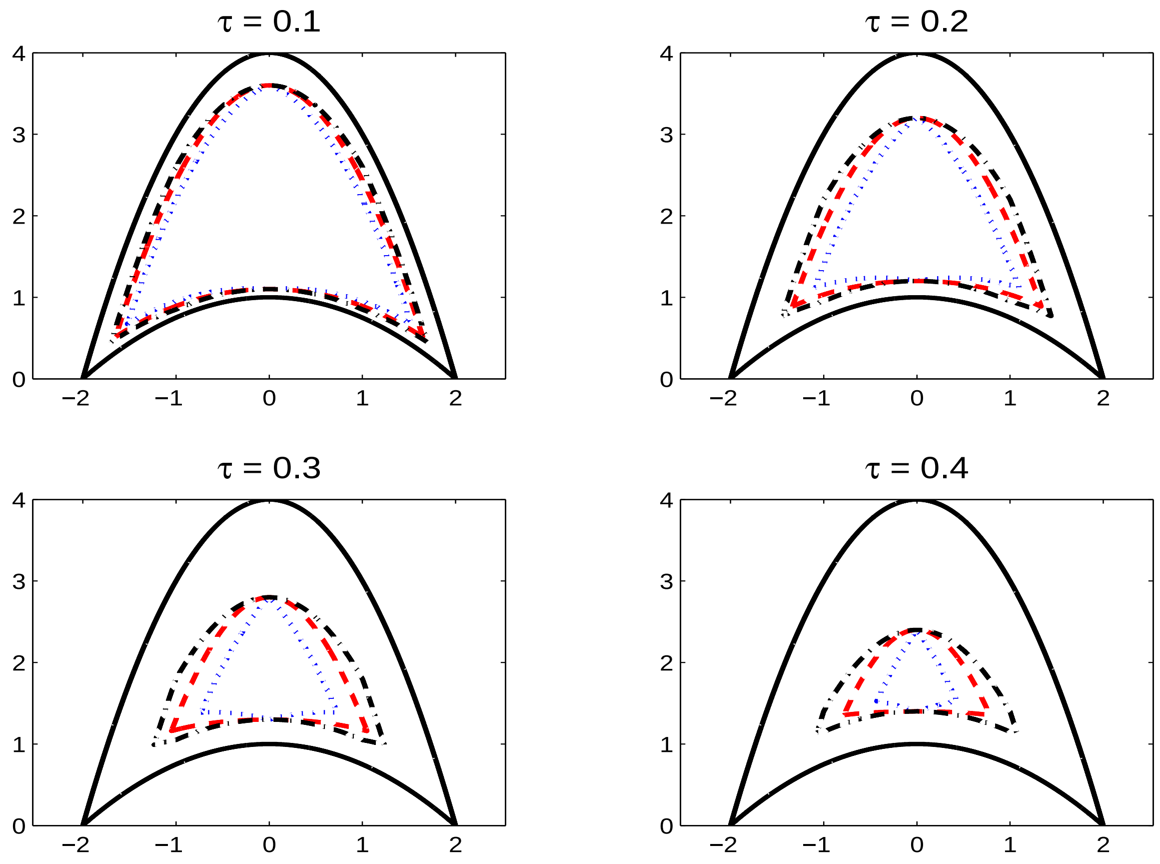
- Now replace $\max h_i(u, s) \le 0$ by $h_i(u, \hat{s}) + \tau \, \|D^T \nabla_s h_i(u, \hat{s})\|_q \le 0$.

Robust Model: Inequality-only

$$\begin{array}{ll} \min & \phi(u) \\ \text{s.t.} & h_i(u,\hat{s}) + \tau \, \|D^T \nabla_s h_i(u,\hat{s})\|_q \le 0, \quad \forall i \end{array}$$

- Only need to optimize once at the nominal value $s = \hat{s}$.
- The added nonnegative terms are "safety margins".
- The margins are proportional to the magnitude τ of the variations.
- Also proportional to the sensitivity of h_i to parameter s at $\hat{s}$.

Robust Feasible Set Illustrations



Feasible set is $\{(x, y) : |x| \leq 2, (ax^2 + b) \leq y \leq 4(ax^2 + b)\}$ with $\hat{a} = 1/4$ and $\hat{b} = 1$. Robust feasible sets are plotted for $q = 1, 2, \infty$ and various τ values.

General Nonlinear Programming

$$\begin{aligned} \min_{y,u\in\mathcal{U}} \quad & \phi(y,u,s) \\ \text{s.t.} \quad & F(y,u,s) = 0 \\ & G(y,u,s) \leq 0 \end{aligned}$$

Approach: **Implicit Function, elimination, and Chain Rule**.

Let $y(u,s)$ satisfy $F(y,u,s) = 0$. Then

$$H(u,s) := G(y(u,s),u,s) \leq 0$$

Even if $G = G(y,u)$, H can still be sensitive to s through the state y.

Be explicit with Implicit Functions

After elimination and robustification:

$$F(y,u,\hat{s}) = 0,\ g_i(y,u,\hat{s}) \le 0 \quad \Longrightarrow \quad h_i(u,\hat{s}) + \tau \| D^T \nabla_s h_i(u,\hat{s}) \|_q \le 0.$$

Recall $H(u,s) = G(y(u,s),u,s)$,

$$\nabla_s h_i = (H_s)^T e_i, \qquad H_s = G_s + G_y y_s.$$

Implicit functions are hard to handle. We put definitions into constraints.

$$\begin{aligned} F(y,u,s) &= 0, \quad \text{for } y(u,s) \\ F_y y_s + F_s &= 0, \quad \text{for } y_s(u,s) \end{aligned}$$

With the above equality imposed,

$$\Longleftrightarrow \quad g_i(y,u,\hat{s}) + \tau \| D^T (G_s + G_y\, y_s)^T e_i \|_q \le 0$$

General Robust Model

$$\begin{array}{ll} \displaystyle\min_{y,u\,\in\mathcal{U}} & \phi(y,u,\hat{s}) \\ \text{s.t.} & F(y,u,\hat{s}) = 0 \\ & \tau(F_y y_s + F_s)(y,u,\hat{s}) = 0 \\ & g_i(y,u,\hat{s}) + \tau\|D^T(G_s + G_y\, y_s)^T e_i\|_q \le 0 \end{array}$$

$$\tau = 0 \implies \text{The original model}$$

- The robust model remains a one-level, nonlinear program involving only the nominal values of uncertain parameters.
- The robust model produces a robust design/control variable u^*. For every parameter value s near $\hat{s}$, we can find a corresponding **safe state** $y(u^*, s)$.

A Robust Model for $p = 1 (q = \infty)$

$$\begin{aligned} \min_{y,u \in \mathcal{U}} \quad & \phi(y, u, \hat{s}) \\ \text{s.t.} \quad & F(y, u, \hat{s}) = 0 \\ & \tau(F_y y_s + F_s)(y, u, \hat{s}) = 0 \\ & G(y, u, \hat{s}) e^T \pm \tau(G_y y_s + G_s) D \leq 0 \end{aligned}$$

If everything is sensitive to $s \in \Re^k$, then:

$n \rightarrow n + n * k$ state variables (y and y_s)

$n \rightarrow n + n * k$ equations (for y and y_s)

$m \rightarrow 2 * m * k$ inequalities

Computational costs remain in the same order if $k = O(1)$
(not so if $k = O(n)$ or $O(m)$).

ODE Example Revisited

$$
\begin{array}{lll}
\min & \left\| y(T) - y^{end} \right\|_2^2 & \\
\text{s.t.} & \dot{y} = f(y, u, \hat{s}), & y(0) = y^{ini}, \\
& \dot{y}_s = f_y(y, u, \hat{s}) y_s + f_s(y, u, \hat{s}), & y_s(0) = 0, \\
& y(t) e^T \pm \tau y_s(t) \leq 0, & t \in (0, T)
\end{array}
$$

Let $y \in \Re^n$ and $s \in \Re^k$.

- A new ODE system of size $n * k$ is added for the sensitivity y_s.
- Safety constraints become a set of $2n * k$ constraints.

1st-Order Robustness

Theorem

Let $(\hat{y}, \hat{u})$ *be strictly feasible to corresponding to* $\hat{s}$ *and* $\tau > 0$*.*

Assume that in the set $\mathcal{S}_\tau$*,*

1. $y(\hat{u}, s)$ *is implicitly defined as a differentiable function of* s *via the equation* $F(y, \hat{u}, s) = 0$*;*
2. *every row of* $[G_y\, y_s + G_s](y(s), \hat{u}, s)$ *is Lipschitz continuous modulo to* L*.*

Then

$$G(y(\hat{u}, s), \hat{u}, s) \leq \frac{L}{2}\tau^2, \ \forall s \in \mathcal{S}_\tau.$$

Parameterized Linear Programs:

$$
\begin{aligned}
\min_{y,u} \quad & \langle c_0(s), y\rangle + \langle d_0(s), u\rangle + \gamma_0(s) \\
\text{s.t.} \quad & A(s)y + B(s)u - h(s) = 0 \\
& \langle c_i(s), y\rangle + \langle d_i(s), u\rangle + \gamma_i(s) \leq 0 \\
& i = 1, 2, \cdots, m
\end{aligned}
$$

Proposition

If all the functions of s are in C^1 (not necessarily linear) and $A(s)$ is invertible, then the robust model of the above parameterized LP is again an LP for $p = 1$ or ∞, and an SOCP for $p = 2$.

— Nonlinear parameter dependency can be handled.

— Equality constraints can be handled (at least partially).

Applicability, Advantages and Limitations

- Provide intelligent safety margins to maintain safety under parameter uncertainty.
- Can handle a wide range of general nonlinear programs.
- Solve a one-level nonlinear program at the nominal value; relatively inexpensive when the number of uncertain parameters is moderate.
- Require good parameter estimates. Parameter variations need to be relatively small around their nominal values.
- Robust only to the 1st-order around the nominal value.
- A technical report is available at: http://www.caam.rice.edu/caam/trs/2004/TR04-13.pdf