NLP Formulations of Chance Constraints With Application to OPF

Andreas Wächter
joint with Alejandra Peña-Ordieres, Jim Luedtke, Dan Molzahn, and Line Roald

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL, USA
andreas.waechter@northwestern.edu

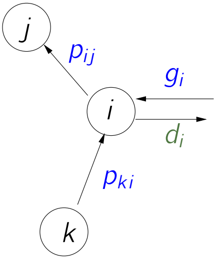
2019 Grid Science Conference
Santa Fe, NM

January 10, 2019

Motivation: DC Optimal Power Flow

- d_i: demand at bus i
- g_i: power generation at bus i

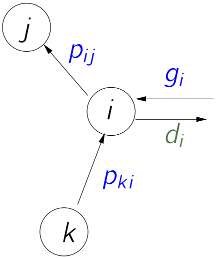
$$\min_{g} \; f(g)$$
$$\text{s.t.} \; \sum_i g_i = \sum_i d_i$$
$$g_i^L \le g_i \le g_i^U \qquad \forall_i$$
$$\qquad \forall_{i,j}$$



Motivation: DC Optimal Power Flow

- d_i: demand at bus i
- g_i: power generation at bus i
- $p_{ij} = m_{ij}(g - d)$: power flow in line from bus i to bus j
- m_{ij}: Power transfer distribution factors (PTDFs)

$$
\begin{aligned}
\min_{g} \quad & f(g) \\
\text{s.t.} \quad & \sum_i g_i = \sum_i d_i \\
& g_i^L \le g_i \le g_i^U && \forall_i \\
& p_{ij}^L \le m_{ij}(g - d) \le p_{ij}^U && \forall_{i,j}
\end{aligned}
$$



DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^{\xi} = d_i + \xi_i$
- Random variable ξ_i

DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^\xi = d_i + \xi_i$
- Random variable ξ_i
- Total demand change: $\Delta = \sum_i \xi_i$
- Generator response: $g_i^\xi = g_i + \alpha_i \Delta$
- Participation factors: α_i ($\alpha \geq 0$ and $\sum_i \alpha_i = 1$)

DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^{\xi} = d_i + \xi_i$
- Random variable ξ_i
- Total demand change: $\Delta = \sum_i \xi_i$
- Generator response: $g_i^{\xi} = g_i + \alpha_i \Delta$
- Participation factors: α_i $\quad$ ($\alpha \geq 0$ and $\sum_i \alpha_i = 1$)

$$
\begin{aligned}
&\min_{g} \; f(g) \\
&\text{s.t.} \; \left\{ \begin{array}{lllll} g_i^L \leq g_i & & \leq g_i^U & & \forall_i \\ p_{ij}^L \leq m_{ij}(g_i & & - d_i &) \leq p_{ij}^U & \forall_{i,j} \end{array} \right\} \\
&\qquad \sum_i g_i = \sum_i d_i,
\end{aligned}
$$

DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^\xi = d_i + \xi_i$
- Random variable ξ_i
- Total demand change: $\Delta = \sum_i \xi_i$
- Generator response: $g_i^\xi = g_i + \alpha_i \Delta$
- Participation factors: α_i $\quad (\alpha \geq 0$ and $\sum_i \alpha_i = 1)$

$$
\begin{aligned}
&\min_{g,\alpha} \; f(g) \\
&\text{s.t.} \; \left\{ \begin{array}{ll} g_i^L \leq g_i + \alpha_i \Delta \leq g_i^U & \forall_i \\ p_{ij}^L \leq m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \leq p_{ij}^U & \forall_{i,j} \end{array} \right\} \\
&\qquad \sum_i g_i = \sum_i d_i, \quad \alpha \geq 0, \quad \sum_i \alpha_i = 1
\end{aligned}
$$

DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^{\xi} = d_i + \xi_i$
- Random variable ξ_i
- Total demand change: $\Delta = \sum_i \xi_i$
- Generator response: $g_i^{\xi} = g_i + \alpha_i \Delta$
- Participation factors: α_i $(\alpha \geq 0$ and $\sum_i \alpha_i = 1)$

Robust formulation: Feasible for any ξ.

$$\min_{g,\alpha} \; f(g)$$

$$\text{s.t.} \; \left\{ \begin{array}{ll} g_i^L \leq g_i + \alpha_i \Delta \leq g_i^U & \forall_i \\ p_{ij}^L \leq m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \leq p_{ij}^U & \forall_{i,j} \end{array} \right\} \text{ for all } \xi$$

$$\sum_i g_i = \sum_i d_i, \quad \alpha \geq 0, \quad \sum_i \alpha_i = 1$$

DC Optimal Power Flow with Uncertainty

- Uncertain demand: $d_i^\xi = d_i + \xi_i$
- Random variable ξ_i
- Total demand change: $\Delta = \sum_i \xi_i$
- Generator response: $g_i^\xi = g_i + \alpha_i \Delta$
- Participation factors: α_i $\qquad (\alpha \geq 0$ and $\sum_i \alpha_i = 1)$

Chance-constrained formulation: Feasible with a specified probability

$$
\begin{aligned}
&\min_{g,\alpha} \; f(g) \\
&\text{s.t. } \mathbb{P}_\xi \left\{ \begin{array}{ll} g_i^L \leq g_i + \alpha_i \Delta \leq g_i^U & \forall_i \\ p_{ij}^L \leq m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \leq p_{ij}^U & \forall_{i,j} \end{array} \right\} \geq 0.95 \\
&\qquad \sum_i g_i = \sum_i d_i, \quad \alpha \geq 0, \quad \sum_i \alpha_i = 1
\end{aligned}
$$

Single vs. Joint Chance Constraints

Joint chance constraints:

$$\mathbb{P}_{\xi}\left\{\begin{array}{ll} g_i^L \leq g_i + \alpha_i \Delta \leq g_i^U & \forall_i \\ p_{ij}^L \leq m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \leq p_{ij}^U & \forall_{i,j} \end{array}\right\} \geq 0.95$$

Single chance constraints:

$$\mathbb{P}_{\xi}\left\{g_i^L \leq g_i + \alpha_i \Delta \leq g_i^U\right\} \geq 0.95 \quad \forall_i$$

$$\mathbb{P}_{\xi}\left\{p_{ij}^L \leq m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \leq p_{ij}^U\right\} \geq 0.95 \quad \forall_{i,j}$$

- [Bienstock et al. 14], [Roald et al. 17]
- Easier to solve
- Do not quite capture the real goal

General Problem Statement

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_\xi \{c(x,\xi) \leq 0\} \geq 1 - \alpha$$

$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$c(x,\xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}^m$$

- Random variable ξ
- Assume $f(x)$ and $c(x,\xi)$ are sufficiently smooth for all ξ
- $X \subseteq \mathbb{R}^n$: Captures additional constraints

General Problem Statement

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_\xi \{c(x,\xi) \le 0\} \ge 1 - \alpha$$

$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$c(x,\xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}^m$$

- Random variable ξ
- Assume $f(x)$ and $c(x,\xi)$ are sufficiently smooth for all ξ
- $X \subseteq \mathbb{R}^n$: Captures additional constraints
- Even if $c(\cdot, \xi)$ is linear, feasible region can be nonconvex

General Problem Statement

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_{\xi}\{c(x,\xi) \leq 0\} \geq 1-\alpha$$

$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$c(x,\xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}^m$$

- Random variable ξ
- Assume $f(x)$ and $c(x,\xi)$ are sufficiently smooth for all ξ
- $X \subseteq \mathbb{R}^n$: Captures additional constraints
- Even if $c(\cdot,\xi)$ is linear, feasible region can be nonconvex
- We are looking for local minimizers

General Problem Statement

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_\xi \{c(x,\xi) \leq 0\} \geq 1 - \alpha$$

$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$c(x,\xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}^m$$

- Random variable ξ
- Assume $f(x)$ and $c(x,\xi)$ are sufficiently smooth for all ξ
- $X \subseteq \mathbb{R}^n$: Captures additional constraints
- Even if $c(\cdot,\xi)$ is linear, feasible region can be nonconvex
- We are looking for local minimizers
- Distinction:
  - Single chance constraint: $m = 1$ (first part of talk)
  - Joint chance constraints: $m > 1$

Many Solution Approaches

- Equivalent convex formulations, e.g.
  - [Calafiore, El Ghaoui 06], [Bienstock et al. 14], [Roald et al. 17], . . .
- Feasible convex approximations, e.g.
  - [Pintér 89], [Rockafellar, Uryasev 00], . . .
- Special optimization methods, e.g.
  - [Dentcheva, Prékopa, Ruszczyński 00], . . .
- Sample Average Approximation, e.g.
  - [Calafiore, Campi 05, 06], [Nemirovski, Shapiro 06], [Luedtke 10, 14], [Kücükyavuz 12], [Liu et al. 16], . . .

Disadvantages

- Strong assumptions on random variable (e.g., normal), or
- Restricted to linear or convex problems, or
- Restricted to single chance constraints, or
- Results in conservative solutions, or
- Requires solution of difficult mixed-integer linear programs

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem
- does not require knowledge of distribution function

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem
- does not require knowledge of distribution function
- is not restricted to convex functions

Wish List

Develop approximation technique for chance constraints that

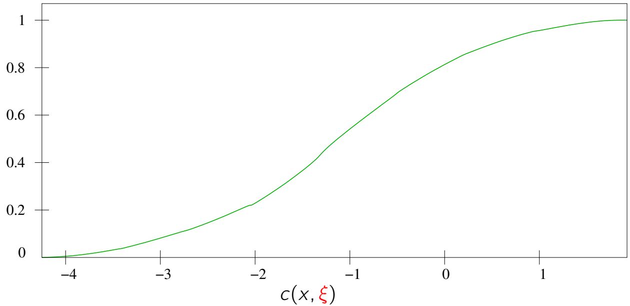
- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem
- does not require knowledge of distribution function
- is not restricted to convex functions
- avoids combinatorial complexity

Wish List

Develop approximation technique for chance constraints that

- results in a single differentiable constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem
- does not require knowledge of distribution function
- is not restricted to convex functions
- avoids combinatorial complexity
- can be extended to handle joint chance constraints

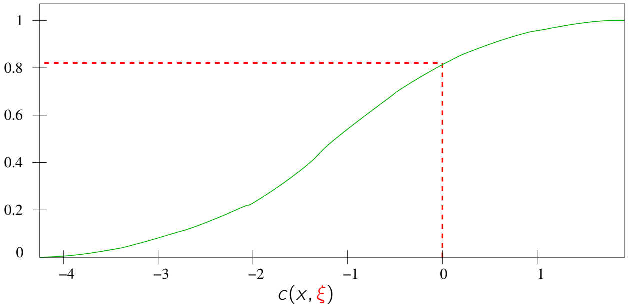
Cumulative Distribution Function



- Cumulative distribution function of constraint values for fixed x

$$\boxed{\Phi(t; x) = \mathbb{P}\{c(x, \xi) \le t\}}$$

Cumulative Distribution Function



- Cumulative distribution function of constraint values <u>for fixed x</u>

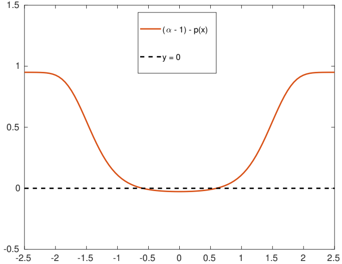
$$\Phi(t; x) = \mathbb{P}\{c(x, \xi) \leq t\}$$

- Probabilistic constraint:

$$p(x) = \Phi(0; x) \geq 1 - \alpha$$

Shape of Probabilistic Constraint Function

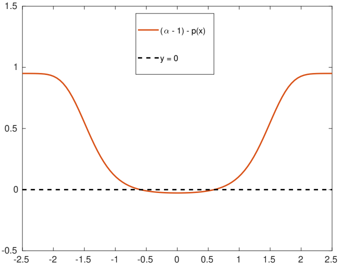
$$p(x) \geq 1 - \alpha \quad \Longleftrightarrow \quad (\alpha - 1) - p(x) \leq 0$$



$$c(x, \xi) = x^2 - 2 + \xi$$
$$\xi \sim \mathcal{N}(0, 1)$$

Shape of Probabilistic Constraint Function

$$\boxed{p(x) \geq 1 - \alpha} \quad \Longleftrightarrow \quad \boxed{(\alpha - 1) - p(x) \leq 0}$$



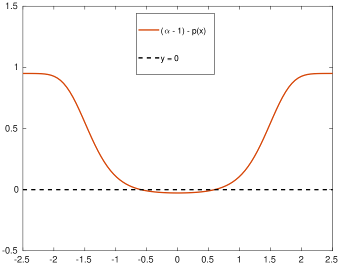
$$\boxed{\begin{array}{c} c(x,\xi) = x^2 - 2 + \xi \\ \xi \sim \mathcal{N}(0,1) \end{array}}$$

Issues

- Always nonconvex

Shape of Probabilistic Constraint Function

$$\boxed{p(x) \geq 1 - \alpha} \quad \Longleftrightarrow \quad \boxed{(\alpha - 1) - p(x) \leq 0}$$



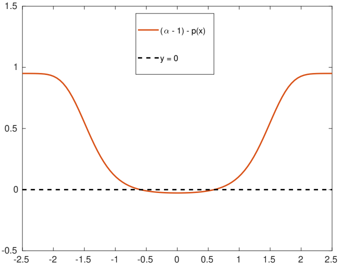
$$\boxed{\begin{array}{c} c(x,\xi) = x^2 - 2 + \xi \\ \xi \sim \mathcal{N}(0,1) \end{array}}$$

Issues

- Always nonconvex
- Linearization is poor approximation (bad for NLP solvers)

Shape of Probabilistic Constraint Function

$$\boxed{p(x) \geq 1 - \alpha} \quad \Longleftrightarrow \quad \boxed{(\alpha - 1) - p(x) \leq 0}$$

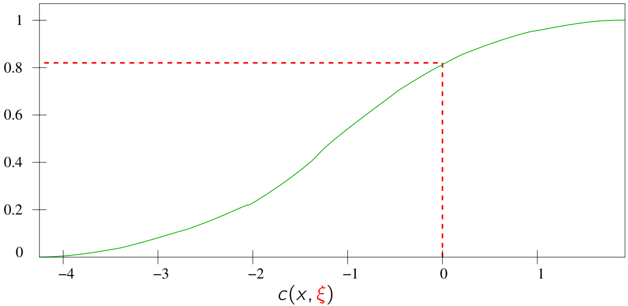


$$\boxed{\begin{array}{c} c(x, \xi) = x^2 - 2 + \xi \\ \xi \sim \mathcal{N}(0, 1) \end{array}}$$

Issues

- Always nonconvex
- Linearization is poor approximation (bad for NLP solvers)
- Used in some nonlinear programming formulations
  - [Hu et al. 13], [Shan et al. 14], [Bremer et al. 15], [Tovar-Facio et al. 18]

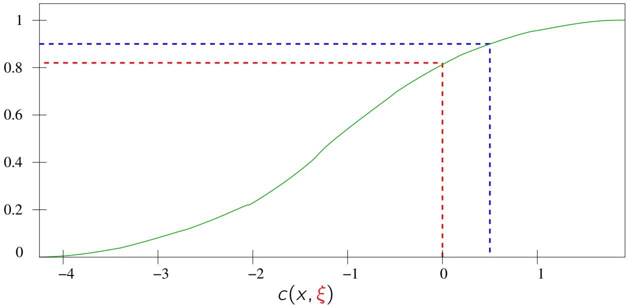
Quantile Formulation



- Cumulative distribution function

$$\boxed{\Phi(t; x) = \mathbb{P}\{c(x, \xi) \leq t\}}$$

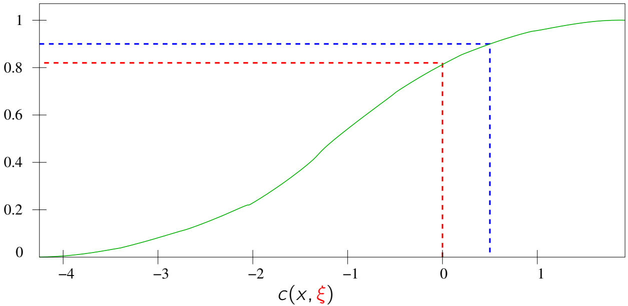
Quantile Formulation



- Cumulative distribution function

$$\boxed{\Phi(t; x) = \mathbb{P}\{c(x, \xi) \leq t\}}$$

Quantile Formulation



- Cumulative distribution function

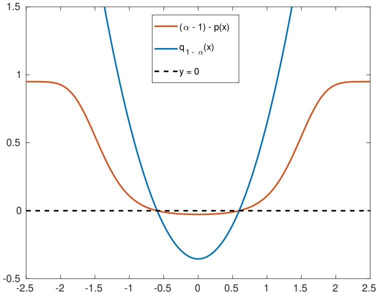
$$\boxed{\Phi(t; x) = \mathbb{P}\{c(x, \xi) \leq t\}}$$

- Quantile formulation:

$$\boxed{q(x) = \Phi^{-1}(1 - \alpha; x) \leq 0}$$

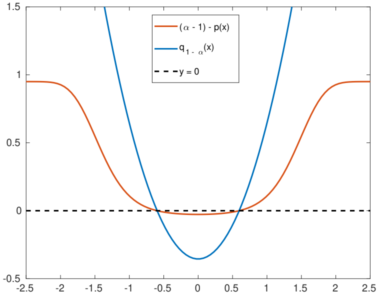
Choice of Formulation

$$\boxed{(1-\alpha) - p(x) \leq 0} \quad \Longleftrightarrow \quad \boxed{q(x) \leq 0}$$



Choice of Formulation

$$\boxed{(1-\alpha) - p(x) \leq 0} \quad \Longleftrightarrow \quad \boxed{q(x) \leq 0}$$

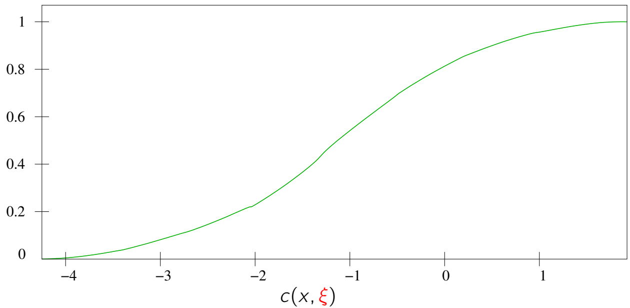


$$\begin{aligned} \min_{x} \ & x \\ \text{s.t.} \ & p(x) \geq 1-\alpha \\ & (\text{or } q(x) \leq 0) \\ & x_L \leq x \leq x_U \end{aligned}$$

Knitro iterations (starting point $x_0 = 3$):

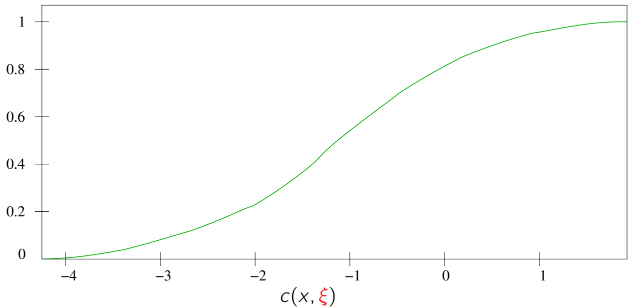
$[x_L, x_U]$	$[-1, 1]$	$[-10, 10]$	$[-100, 100]$	$(-\infty, \infty)$
$p(x)$	6	36	14	5(fail)
$q(x)$	6	6	6	6

Approximation of Cumulative Distribution Function



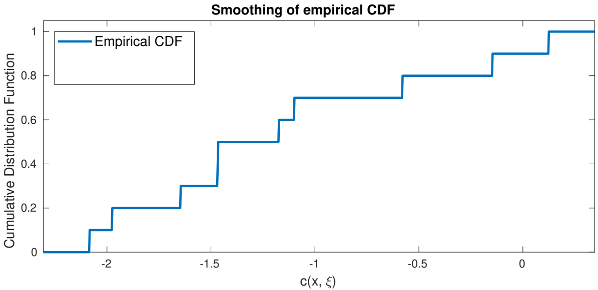
- Recall: $\Phi(t; x) = \mathbb{P}\{Y \leq t\}$ with random variable $Y = c(x, \xi)$

Approximation of Cumulative Distribution Function



- Recall: $\Phi(t; x) = \mathbb{P}\{Y \leq t\}$ with random variable $Y = c(x, \xi)$
- Draw samples $\{\xi_1, \ldots, \xi_N\}$, set $y_i = c(x, \xi_i)$

Approximation of Cumulative Distribution Function



- Recall: $\Phi(t;x) = \mathbb{P}\{Y \le t\}$ with random variable $Y = c(x,\xi)$
- Draw samples $\{\xi_1, \ldots, \xi_N\}$, set $y_i = c(x,\xi_i)$
- "Empirical CDF", using sample average approximation:

$$\widehat{\Phi}^N(t;x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(y_i - t) \qquad \mathbb{1}(s) = \begin{cases} 1 & \text{if } s \le 0 \\ 0 & \text{if } s > 0 \end{cases}$$

Smoothing Indicator Function

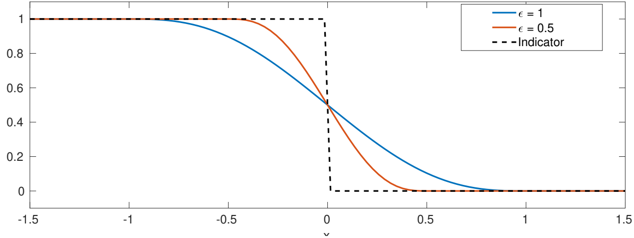
- Empirical cdf: $$\widehat{\Phi}^N(t;x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(y_i - t)$$

  is not smooth because of the indicator function

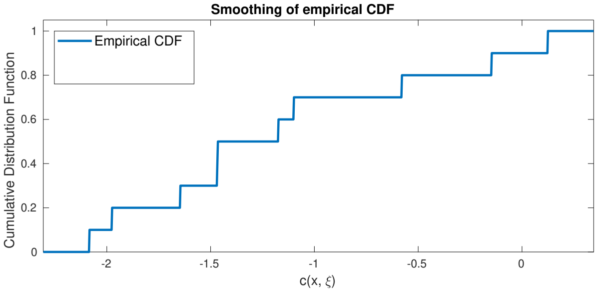
Smoothing Indicator Function

- Empirical cdf: $$\widehat{\Phi}^N(t;x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(y_i - t)$$
  is not smooth because of the indicator function
- Smoothed indicator function (twice continuously differentiable)

$$\bar{\mathbb{1}}^{\epsilon}(s) = \begin{cases} 1 & \text{if } s \le -\epsilon \\ \in (0,1) & \text{if } -\epsilon < s < \epsilon \\ 0 & \text{if } s \ge \epsilon \end{cases}$$



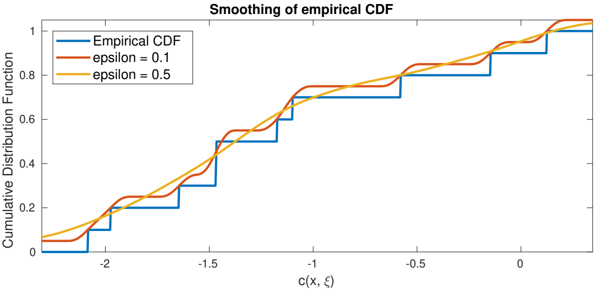
Smoothed Empirical CDF



- Empirical cdf:

$$\widehat{\Phi}^N(t;x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(y_i - t)$$

Smoothed Empirical CDF



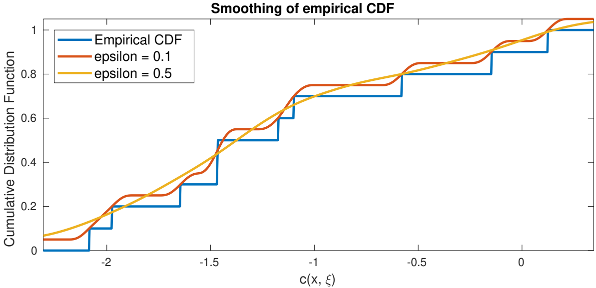
- Empirical cdf:

$$\widehat{\Phi}^N(t;x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(y_i - t)$$

- Smoothed cdf [Azzalini 81]:

$$\bar{\Phi}^{N,\epsilon}(t;x) = \frac{1}{N}\sum_{i=1}^{N} \bar{\mathbb{1}}^{\epsilon}(y_i - t)$$

Smoothed Empirical CDF



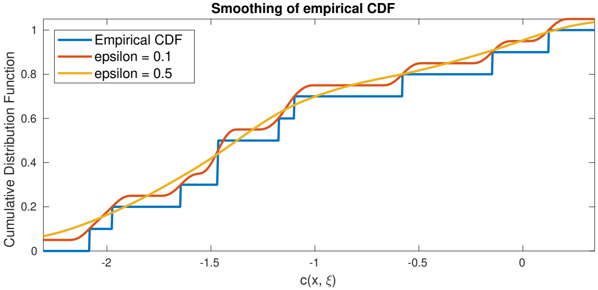
- Empirical cdf:

$$\widehat{\Phi}^N(t; x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}(y_i - t)$$

- Smoothed cdf [Azzalini 81]:

$$\bar{\Phi}^{N,\epsilon}(t; x) = \frac{1}{N} \sum_{i=1}^{N} \bar{\mathbb{1}}^{\epsilon}(y_i - t) = \frac{1}{N} \sum_{i=1}^{N} \bar{\mathbb{1}}^{\epsilon}(c(x, \xi_i) - t)$$

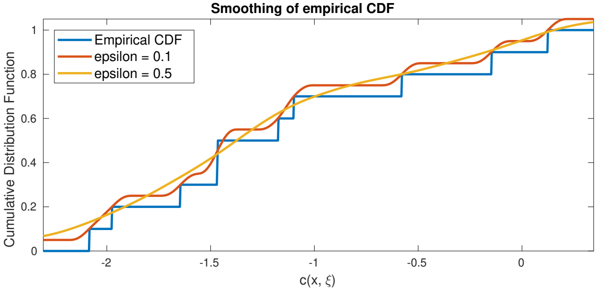
Smoothed Quantile Estimates



- True quantile

$$\boxed{q(x) = \Phi^{-1}(1 - \alpha; x)}$$

Smoothed Quantile Estimates



- True quantile

$$\boxed{q(x) = \Phi^{-1}(1-\alpha; x)}$$

- Smoothed quantile estimate

$$\boxed{\bar{q}^{N,\epsilon}(x) = (\bar{\Phi}^{N,\epsilon})^{-1}(1-\alpha; x)}$$

Differentiable Quantile Constraint

$$\begin{aligned} \min_{x \in X} \; & f(x) \\ \text{s.t. } & \bar{q}^{N,\epsilon}(x) \leq 0 \end{aligned}$$

- Smoothed quantile estimate

$$\bar{q}^{N,\epsilon}(x) = (\bar{\Phi}^{N,\epsilon})^{-1}(1 - \alpha; x)$$

Differentiable Quantile Constraint

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(x) \leq 0$$

- Smoothed quantile estimate

$$\boxed{\bar{q}^{N,\epsilon}(x) = (\bar{\Phi}^{N,\epsilon})^{-1}(1-\alpha; x)}$$

- Compute $\bar{q}^{N,\epsilon}(x)$ as solution t^* of

$$1 - \alpha = \bar{\Phi}^{N,\epsilon}(t; x)$$

Differentiable Quantile Constraint

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(x) \leq 0$$

- Smoothed quantile estimate

$$\bar{q}^{N,\epsilon}(x) = (\bar{\Phi}^{N,\epsilon})^{-1}(1-\alpha; x)$$

- Compute $\bar{q}^{N,\epsilon}(x)$ as solution t^* of

$$1-\alpha = \bar{\Phi}^{N,\epsilon}(t; x) = \frac{1}{N}\sum_{i=1}^{N} \bar{\mathbb{1}}^{\epsilon}(c(x,\xi_i) - t)$$

Differentiable Quantile Constraint

$$\min_{x \in X} \; f(x) \quad \text{s.t.} \; \bar{q}^{N,\epsilon}(x) \leq 0$$

- Smoothed quantile estimate

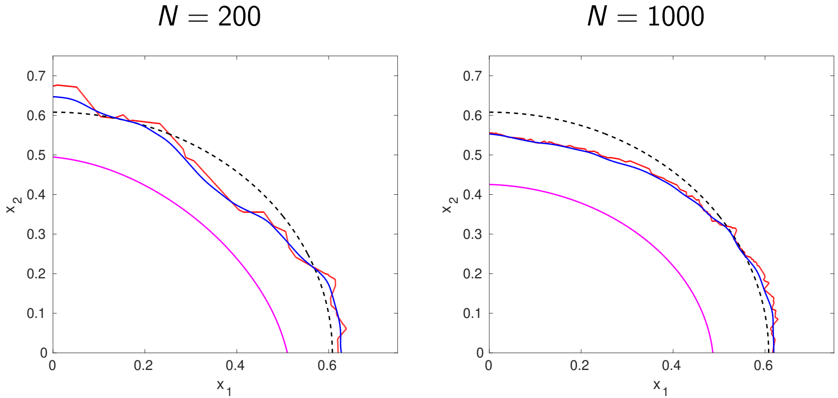
$$\bar{q}^{N,\epsilon}(x) = (\bar{\Phi}^{N,\epsilon})^{-1}(1-\alpha; x)$$

- Compute $\bar{q}^{N,\epsilon}(x)$ as solution t^* of

$$1-\alpha = \bar{\Phi}^{N,\epsilon}(t; x) = \frac{1}{N}\sum_{i=1}^{N} \bar{\mathbb{1}}^{\epsilon}(c(x,\xi_i) - t)$$

- Implicit function theorem:
  - $\bar{q}^{N,\epsilon}(x)$ exists and is differentiable with respect to x.

Feasible Region with Smoothed Quantile



$$c(x, \xi) = \xi_1 x_1 + \xi_2 x_2 - 1 \qquad \xi_1, \xi_1 \sim \mathcal{N}(0, 1)$$

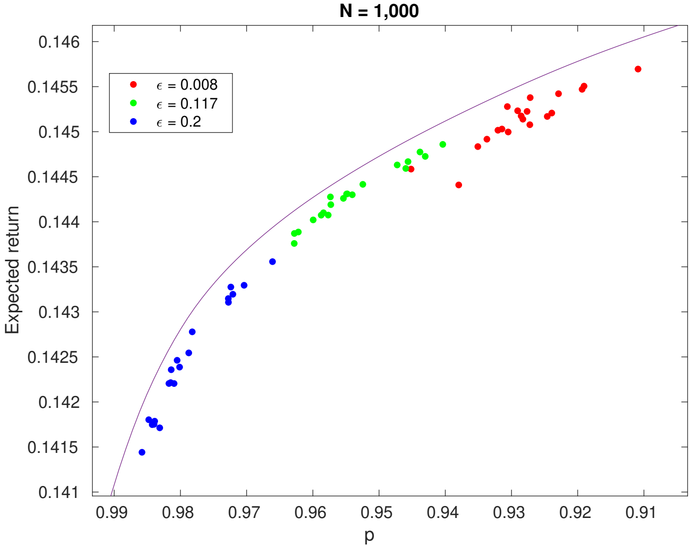
- Empirical
- $\epsilon = 0.1$
- $\epsilon = 1$

Example: Portfolio Optimization

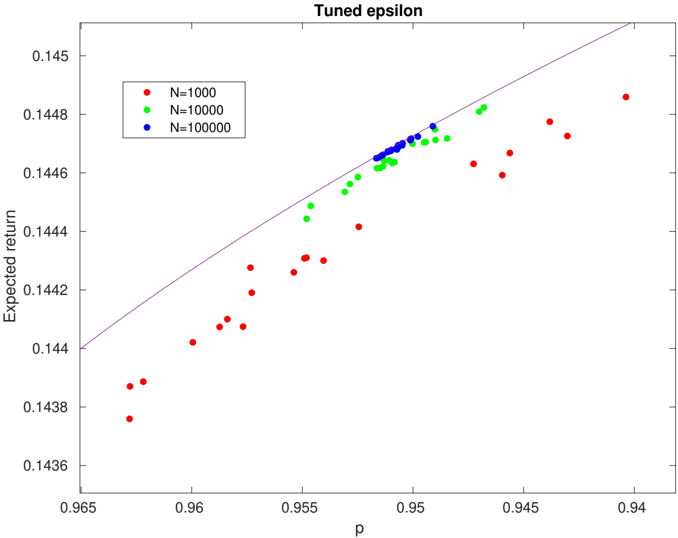
$$\max_{x \in \mathbb{R}^n} \quad \mathbb{E}[r^T x] = \mu^T x$$
$$\text{s.t.} \quad \mathbb{P}\{r^T x \geq -0.05\} \geq 0.95$$
$$\sum_{i=1}^{n} x_i = 1, \qquad x \geq 0$$

- Dimension $n = 100$ $\qquad (r \sim \mathcal{N}(\mu, \Sigma))$
- Solve resulting NLP with Knitro (general-purpose NLP solver)
- Vary number of scenarios N and smoothing parameter ϵ
- 20 replications (different draws) for each combination

Solution Quality



Solution Quality



Observations

- Even for $N = 100,000$ scenarios, Knitro's computation time is about 30 secs.
- Choice of smoothing parameter ϵ impacts true probability of constraint satisfaction (bias)
- Finding a good value of ϵ requires some tuning
  - Perform binary search
  - Estimate "true" probability using Monte-Carlo approximation
- Very good solution quality for large N
  - Close to best possible objective for achieved feasibility level
- Less variance in solution for large N

Asymptotic Convergence

Assumptions

There exists an optimal solution of the true problem, x^*, such that for any $\delta > 0$ there is $x \in X$ such that $\|x - x^*\| \leq \delta$ and $\Phi(x) := \mathbb{P}\{c(x, \xi) \leq 0\} > 1 - \alpha$.

Theorem

Suppose: X is compact, $f(x)$ is continuous, and $c(x, \xi)$ is a Carathéodory function. Then $opt_\epsilon^N \to opt^*$ and $\mathbb{D}(S_\epsilon^N, S) \to 0$ w.p.1 as $N \to \infty$ and $\epsilon \to 0$.

Finite Sample Feasibility

Theorem (Feasibility of single point)

Let $x \in X$ be such that the pdf of the random variable $c(x,\xi)$, is strictly decreasing in the interval $(-\epsilon,\epsilon)$. If $\bar{\Phi}^{N,\epsilon}(x) \geq 1-\alpha$, then

$$\mathbb{P}\left\{x \in X_\alpha\right\} \geq 1 - \exp\left\{-2N\beta_x^2\right\}$$

where $\beta_x = \Phi(x) - \Phi^\epsilon(x) > 0$.

Finite Sample Feasibility

Theorem (Feasibility of single point)

Let $x \in X$ be such that the pdf of the random variable $c(x, \xi)$, is strictly decreasing in the interval $(-\epsilon, \epsilon)$. If $\bar{\Phi}^{N,\epsilon}(x) \geq 1 - \alpha$, then

$$\mathbb{P}\left\{x \in X_\alpha\right\} \geq 1 - \exp\left\{-2N\beta_x^2\right\}$$

where $\beta_x = \Phi(x) - \Phi^\epsilon(x) > 0$.

- This shows that our approximation is asymptotically conservative.

Finite Sample Feasibility

Theorem (Feasibility of single point)

Let $x \in X$ be such that the pdf of the random variable $c(x, \xi)$, is strictly decreasing in the interval $(-\epsilon, \epsilon)$. If $\bar{\Phi}^{N,\epsilon}(x) \geq 1 - \alpha$, then

$$\mathbb{P}\{x \in X_\alpha\} \geq 1 - \exp\left\{-2N\beta_x^2\right\}$$

where $\beta_x = \Phi(x) - \Phi^\epsilon(x) > 0$.

- This shows that our approximation is asymptotically conservative.

Theorem (Feasibility of all points)

(Definitions and assumptions. . .)

$$\mathbb{P}\{X_{\epsilon,\alpha}^{N,t} \subseteq X_\alpha\} \geq 1 - \lceil 1/\beta \rceil \lceil 2LD/t \rceil^n \exp\left\{-2N(M - \beta)^2\right\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}\{c(x,\xi) \leq 0\} \geq 1-\alpha$$

- Sample average approximation

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(c(x,\xi_i)) \leq 0$$

$$\mathcal{Y} = \{c(x,\xi_1),\ldots,c(x,\xi_N)\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}\{c_j(x,\xi) \le 0,\ j = 1,\dots,m\} \ge 1 - \alpha$$

- Sample average approximation

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(c(x,\xi_i)) \le 0$$

$$\mathcal{Y} = \{c(x,\xi_1),\dots,c(x,\xi_N)\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x) \quad \text{s.t. } \mathbb{P}_\xi\{\hat{c}(x,\xi) \le 0\} \ge 1-\alpha$$

$$\hat{c}(x,\xi) := \max_j c_j(x,\xi)$$

- Sample average approximation

$$\min_{x \in X} f(x) \quad \text{s.t. } \bar{q}^{N,\epsilon}(c(x,\xi_i)) \le 0$$

$$\mathcal{Y} = \{c(x,\xi_1),\ldots,c(x,\xi_N)\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x) \quad \text{s.t. } \mathbb{P}_\xi\{\hat{c}(x,\xi) \le 0\} \ge 1-\alpha$$

$$\hat{c}(x,\xi) := \max_j c_j(x,\xi)$$

- Sample average approximation

$$\min_{x \in X} f(x) \quad \text{s.t. } \bar{q}^{N,\epsilon}(\hat{c}(x,\xi_i)) \le 0$$

$$\mathcal{Y} = \{\hat{c}(x,\xi_1),\ldots,\hat{c}(x,\xi_N)\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_\xi\{\widehat{c}(x,\xi) \leq 0\} \geq 1-\alpha$$

$$\widehat{c}(x,\xi) := \max_j c_j(x,\xi)$$

- Sample average approximation

$$\min_{x \in X, z} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(z_i) \leq 0$$
$$c_j(x,\xi_i) \leq z_i \text{ for all } j \text{ and } i$$

$$\mathcal{Y} = \{z_1, \ldots, z_N\}$$

Joint Chance Constraints

- Original Problem

$$\min_{x \in X} f(x)$$
$$\text{s.t. } \mathbb{P}_\xi\{\hat{c}(x,\xi) \le 0\} \ge 1 - \alpha$$

$$\hat{c}(x,\xi) := \max_j c_j(x,\xi)$$

- Sample average approximation

$$\min_{x \in X, z} f(x)$$
$$\text{s.t. } \bar{q}^{N,\epsilon}(z_i) \le 0$$
$$c_j(x,\xi_i) \le z_i \text{ for all } j \text{ and } i$$

$$\mathcal{Y} = \{z_1, \ldots, z_N\}$$

- We developed specialized SQP-type trust region algorithm
  - Converges to stationary point of SAA problem

Joint Chance Constraints: DC-OPF

$$\min_{g,\alpha} f(g)$$

$$\text{s.t. } \mathbb{P}_\xi \left\{ \begin{array}{ll} g_i^L \le g_i + \alpha_i \Delta \le g_i^U & \forall_i \\ p_{ij}^L \le m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \le p_{ij}^U & \forall_{i,j} \end{array} \right\} \ge 0.95$$

$$\sum_i g_i = \sum_i d_i, \quad \alpha \ge 0, \quad \sum_i \alpha_i = 1$$

- MatPower `case118_bis` (line limits from PGlib)
  - 118 buses, 54 generators, 186 lines
  - 480 joint chance constraints

Joint Chance Constraints: DC-OPF

$$
\begin{aligned}
&\min_{g,\alpha} \; f(g) \\
&\text{s.t. } \mathbb{P}_{\xi}\left\{ \begin{array}{ll} g_i^L \le g_i + \alpha_i \Delta \le g_i^U & \forall_i \\ p_{ij}^L \le m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \le p_{ij}^U & \forall_{i,j} \end{array} \right\} \ge 0.95 \\
&\qquad \sum_i g_i = \sum_i d_i, \quad \alpha \ge 0, \quad \sum_i \alpha_i = 1
\end{aligned}
$$

- MatPower `case118_bis` (line limits from PGlib)
  - 118 buses, 54 generators, 186 lines
  - 480 joint chance constraints
- Matlab implementation of SQP trust-region algorithm:
  - CPLEX to solve QP subproblems: $213 + N$ vars, $480 \times N$ cons

Joint Chance Constraints: DC-OPF

$$\min_{g,\alpha} f(g)$$
$$\text{s.t. } \mathbb{P}_{\xi}\left\{ \begin{array}{ll} g_i^L \le g_i + \alpha_i \Delta \le g_i^U & \forall_i \\ p_{ij}^L \le m_{ij}(g_i + \alpha_i \Delta - d_i - \xi_i) \le p_{ij}^U & \forall_{i,j} \end{array} \right\} \ge 0.95$$
$$\sum_i g_i = \sum_i d_i, \quad \alpha \ge 0, \quad \sum_i \alpha_i = 1$$

- MatPower `case118_bis` (line limits from PGlib)
  - 118 buses, 54 generators, 186 lines
  - 480 joint chance constraints
- Matlab implementation of SQP trust-region algorithm:
  - CPLEX to solve QP subproblems: $213 + N$ vars, $480 \times N$ cons
  - Speedup: Give only small subset of constraints to QP solver

Numerical Results $N = 100$

	total it	total CPU	$f(x^*)$	$p(x^*)$	ϵ
1	44	89.95	126.3310	0.9499	0.0919
2	35	95.10	126.4500	0.9501	0.1182
3	5	33.28	126.4371	0.9499	0.1075
4	17	48.29	126.6808	0.9500	0.1441
5	54	113.28	126.4482	0.9600	0.1075
6	113	222.68	126.1985	0.9499	0.0401
7	115	245.82	126.2339	0.9501	0.0518
8	34	88.24	126.3858	0.9500	0.1079
9	31	96.37	126.3819	0.9500	0.1029
10	50	114.17	126.4303	0.9504	0.1189
min	5.0	33.28	126.1985	0.9499	0.0401
mean	49.8	114.72	126.3977	0.9510	0.0991
max	115.0	245.82	126.6808	0.9600	0.1441

Numerical Results $N = 200$

	total it	total CPU	$f(x^*)$	$p(x^*)$	ϵ
1	127	373.50	126.2354	0.9517	0.0102
2	65	341.38	126.2847	0.9500	0.0637
3	64	324.84	126.3637	0.9500	0.0851
4	73	410.22	126.2728	0.9500	0.0694
5	74	404.83	126.3427	0.9499	0.0770
6	42	249.00	126.2816	0.9500	0.0641
7	73	427.45	126.3206	0.9500	0.0797
8	98	515.27	126.3200	0.9500	0.0808
9	99	542.51	126.2898	0.9501	0.0713
10	32	222.67	126.4855	0.9500	0.1136
min	32.0	222.67	126.2354	0.9499	0.0102
mean	74.7	381.17	126.3197	0.9502	0.0715
max	127.0	542.51	126.4855	0.9517	0.1136

Numerical Results $N = 500$

	total it	total CPU	$f(x^*)$	$p(x^*)$	ϵ
1	80	1357.93	126.2396	0.9499	0.0538
2	84	1178.41	126.2889	0.9500	0.0725
3	70	966.20	126.2373	0.9500	0.0549
4	172	1923.39	126.2334	0.9510	0.0344
5	92	1423.59	126.1936	0.9500	0.0351
6	66	1026.44	126.2719	0.9500	0.0687
7	147	2268.50	126.2612	0.9499	0.0603
8	222	3238.55	126.2770	0.9502	0.0709
9	137	2096.50	126.2569	0.9511	0.0588
10	87	1328.05	126.1991	0.9500	0.0374
min	66.0	966.20	126.1936	0.9499	0.0344
mean	115.7	1680.76	126.2459	0.9502	0.0547
max	222.0	3238.55	126.2889	0.9511	0.0725

Numerical Results $N = 1000$

	total it	total CPU	$f(x^*)$	$p(x^*)$	ϵ
1	117	3731.60	126.1978	0.9499	0.0367
2	67	2103.88	126.1956	0.9500	0.0374
3	118	3592.03	126.2019	0.9500	0.0401
4	91	2749.62	126.2003	0.9499	0.0405
5	107	2893.31	126.2285	0.9500	0.0546
6	131	3749.32	126.2093	0.9500	0.0435
7	193	4447.18	126.2359	0.9508	0.0344
8	116	3552.09	126.2115	0.9500	0.0431
9	79	2487.61	126.2543	0.9500	0.0633
10	212	4687.62	126.2177	0.9506	0.0222
min	67.0	2103.88	126.1956	0.9499	0.0222
mean	123.1	3399.43	126.2153	0.9501	0.0416
max	212.0	4687.62	126.2543	0.9508	0.0633

Wish List Revisited

Develop approximation technique for probabilistic constraints that

- results in a single smooth constraint
- can be handled by standard continuous nonlinear programming techniques and solvers
- can be added to any nonlinear optimization problem
- does not require knowledge of distribution function
- is not restricted to convex functions
- avoids combinatorial complexity
- can be extended to handle joint chance constraints

Conclusions

Contributions:

- New nonlinear programming approximation of chance constraints
  - Pose constraint on quantile, not probability
  - Key: Apply implicit function theorem to smoothed empirical CDF
- Stochastic convergence guarantees
- Appears to perform better than state-of-the-art MILP approaches
- Specialized trust-region solver for joint chance constraints
- Solved DC-OPF 118 bus system with 480 joint chance constraints and up to $N = 1000$ scenarios

Current work

- Extend to AC-OPF