Nonlinear Programming Formulation of Chance-Constraints

Andreas Wächter
joint with Alejandra Peña-Ordieres and Jim Luedtke

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

US & Mexico Workshop on Optimization and Its Applications
Huatulco, Mexico

January 9, 2018

Problem Statement

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \leq 0] \geq 1 - \alpha \end{aligned}}$$

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

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

Problem Statement

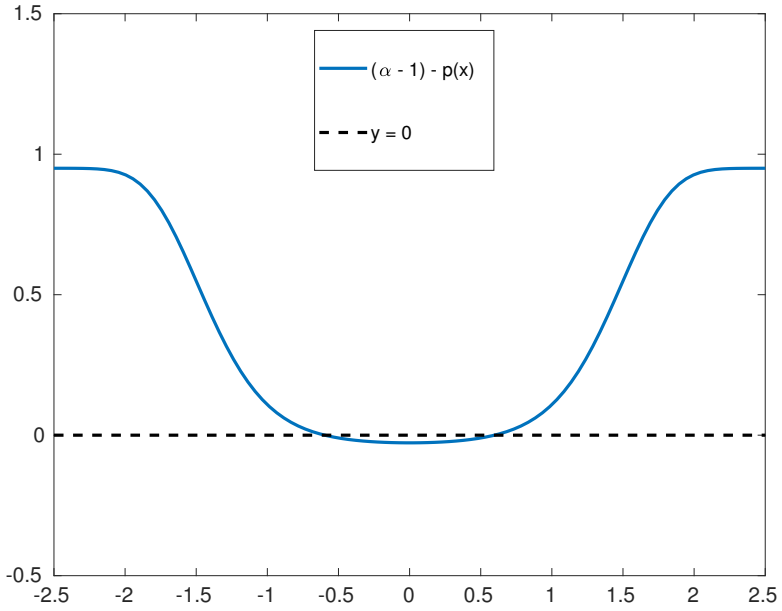
$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \leq 0] \geq 1 - \alpha \end{aligned}}$$

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

- Random variable ξ with support Ξ
- Assume $f(x)$ and $c(x,\xi)$ are sufficiently smooth for all $\xi \in \Xi$
- $X \subseteq \mathbb{R}^n$: Captures additional constraints
- Goal: Formulate as continuous NLP
  - Do not want to assume particular probability distribution
  - Do not want to assume convexity (or convex approximations)
  - Want to avoid combinatorial approach
  - Use powerful NLP algorithms and techniques

Constraint in Probability Space

Impose $p(x) = \mathbb{P}[c(x,\xi) \leq 0] \geq 1 - \alpha$

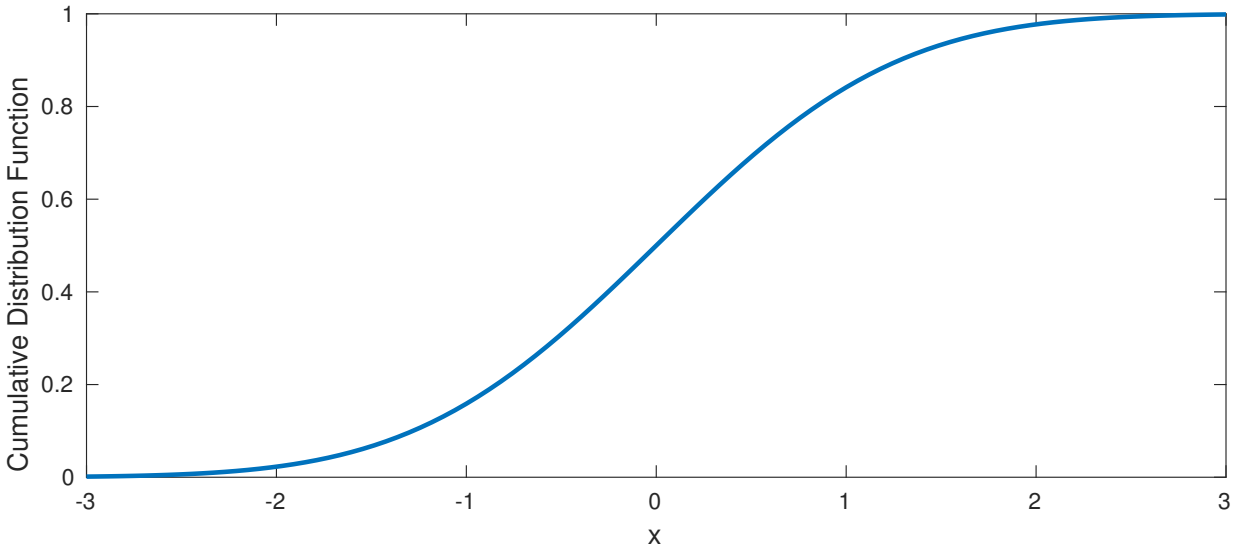


$$c(x,\xi) = x^2 - 2 + \xi$$
$$\xi \sim \mathcal{N}(0,1)$$

Issues

- Linearization is poor approximation
- Always nonconvex
- Used in [Hu, Hong, Zhang 13], [Bremer, Henrion, Möller 15]

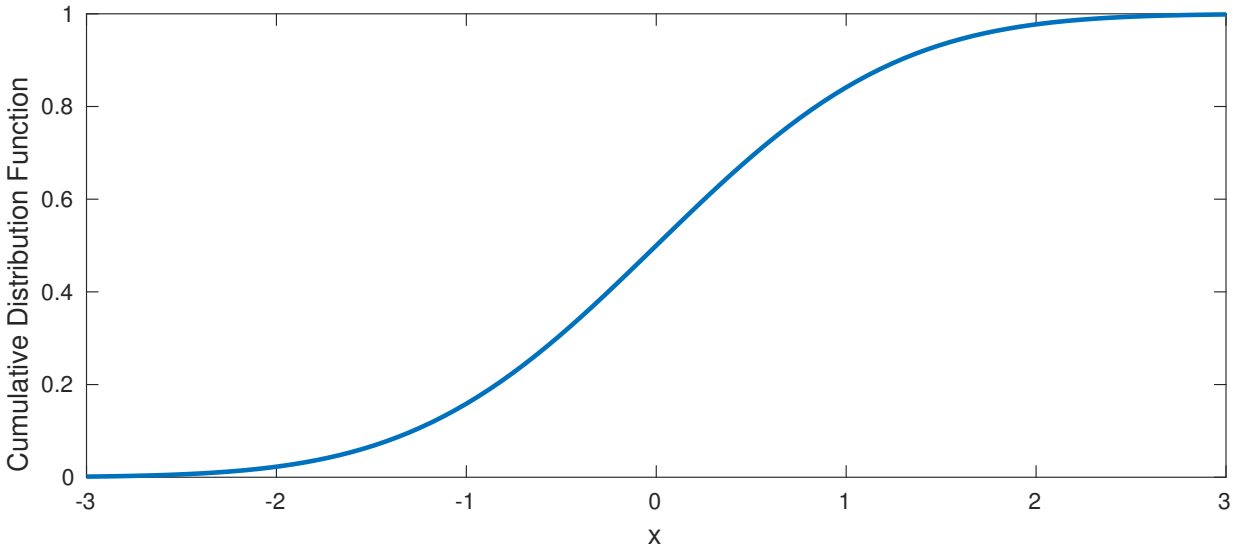
Quantile Formulation



- Let Y be a real-valued random variable
- $(1-\alpha)$-quantile:

$$\boxed{Q^Y_{1-\alpha} = \inf\{y \in \mathbb{R} : \mathbb{P}[Y \leq y] \geq 1-\alpha\}}$$

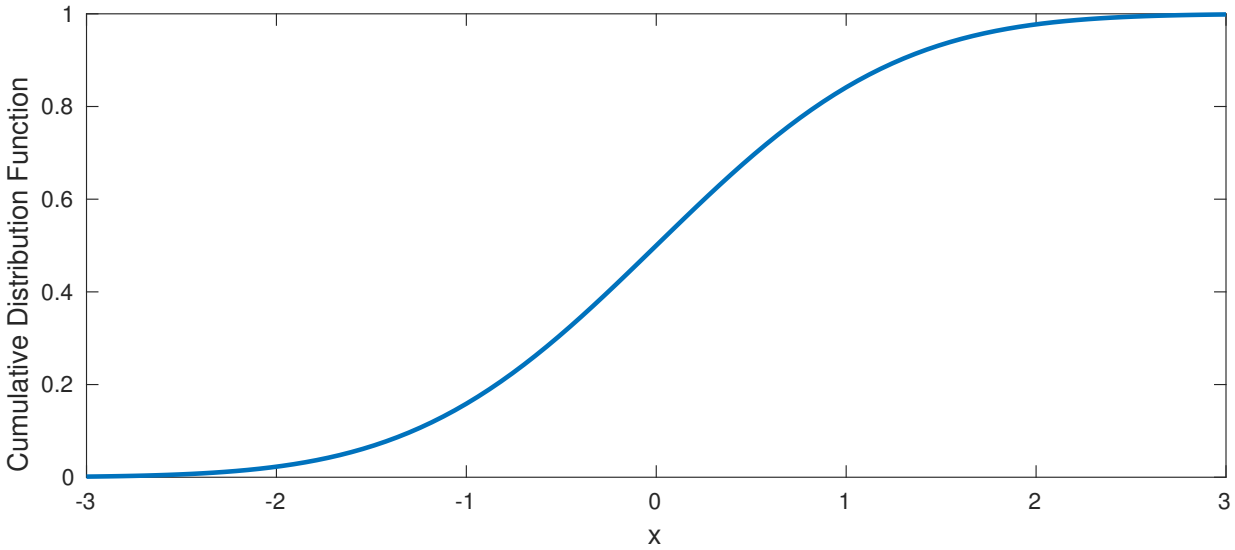
Quantile Formulation



- Let Y be a real-valued random variable ($Y = c(x, \xi)$)
- $(1-\alpha)$-quantile:

$$\boxed{Q_{1-\alpha}^{c(x,\xi)} = \inf\{y \in \mathbb{R} : \mathbb{P}[c(x,\xi) \leq y] \geq 1-\alpha\}}$$

Quantile Formulation



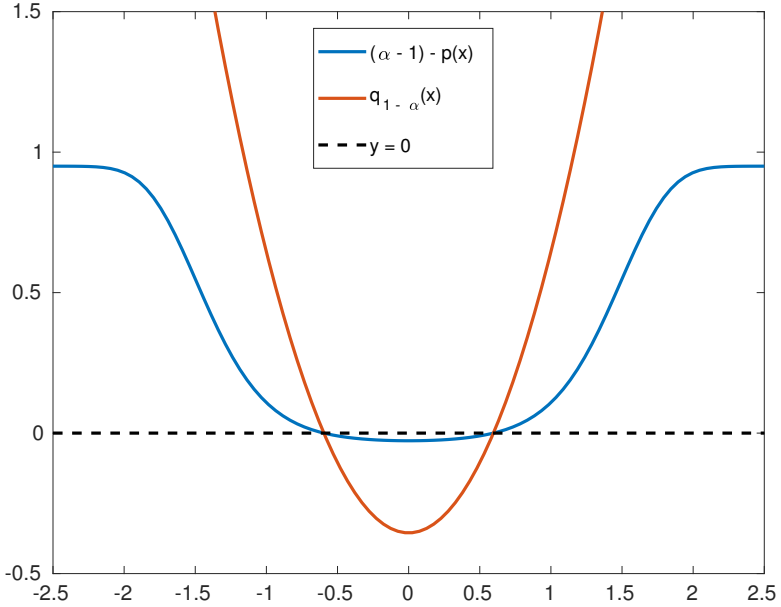
- Let Y be a real-valued random variable ($Y = c(x, \xi)$)
- $(1-\alpha)$-quantile:

$$\boxed{Q_{1-\alpha}^{c(x,\xi)} = \inf\{y \in \mathbb{R} : \mathbb{P}[c(x,\xi) \leq y] \geq 1-\alpha\}}$$

$$p(x) \geq 1-\alpha \qquad \Longleftrightarrow \qquad q_{1-\alpha}(x) = Q_{1-\alpha}^{c(x,\xi)} \leq 0$$

Choice of Formulation

$$p(x) \geq 1 - \alpha \qquad \Longleftrightarrow \qquad q_{1-\alpha}(x) \leq 0$$



- Quantile formulation more suitable for NLP solver

Sample Average Approximation (SAA)

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \le 0] \ge 1 - \alpha \end{aligned}}$$

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

- Finite scenario set $\widehat{\Xi}_N = \{\widehat{\xi}_1, \ldots, \widehat{\xi}_N\}$ with $\widehat{\xi}_i \in \Xi$ chosen i.i.d.
- For simplicity: $c_i(x) = c(x, \widehat{\xi}_i)$

Sample Average Approximation (SAA)

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

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

- Finite scenario set $\widehat{\Xi}_N = \{\widehat{\xi}_1, \ldots, \widehat{\xi}_N\}$ with $\widehat{\xi}_i \in \Xi$ chosen i.i.d.
- For simplicity: $c_i(x) = c(x, \widehat{\xi}_i)$

SAA approximation of quantile

- Ensure that $M = \lceil (1-\alpha)N \rceil$ constraints are satisfied

Sample Average Approximation (SAA)

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

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

- Finite scenario set $\widehat{\Xi}_N = \{\widehat{\xi}_1, \ldots, \widehat{\xi}_N\}$ with $\widehat{\xi}_i \in \Xi$ chosen i.i.d.
- For simplicity: $c_i(x) = c(x, \widehat{\xi}_i)$

SAA approximation of quantile

- Ensure that $M = \lceil (1-\alpha)N \rceil$ constraints are satisfied
- Scenario vector: $C(x) = (c_1(x), \ldots, c_N(x))$

Sample Average Approximation (SAA)

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

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

- Finite scenario set $\widehat{\Xi}_N = \{\widehat{\xi}_1, \ldots, \widehat{\xi}_N\}$ with $\widehat{\xi}_i \in \Xi$ chosen i.i.d.
- For simplicity: $c_i(x) = c(x, \widehat{\xi}_i)$

SAA approximation of quantile

- Ensure that $M = \lceil (1-\alpha)N \rceil$ constraints are satisfied
- Scenario vector: $C(x) = (c_1(x), \ldots, c_N(x))$
- Order constraint values: $c_{[1]}(x) \le c_{[2]}(x) \le \ldots \le c_{[N]}(x)$

Sample Average Approximation (SAA)

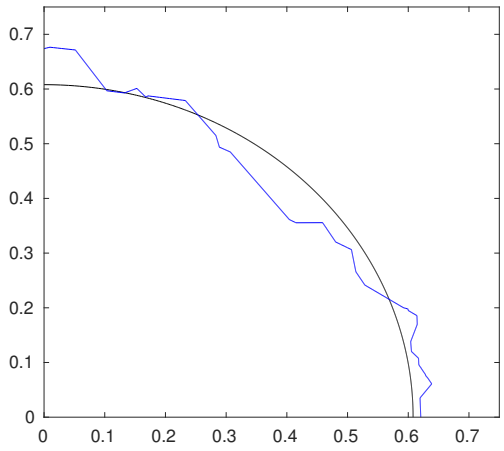
$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

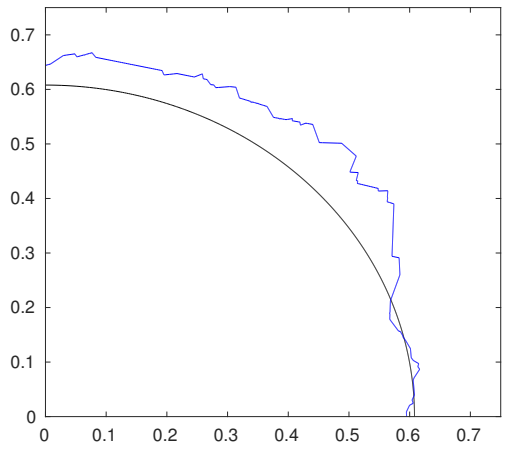
$$f(x) : \mathbb{R}^n \to \mathbb{R}$$
$$c(x,\xi) : \mathbb{R}^n \times \Xi \to \mathbb{R}$$

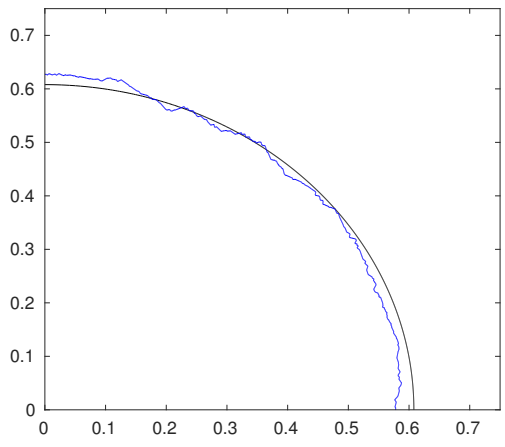
- Finite scenario set $\widehat{\Xi}_N = \{\widehat{\xi}_1, \ldots, \widehat{\xi}_N\}$ with $\widehat{\xi}_i \in \Xi$ chosen i.i.d.
- For simplicity: $c_i(x) = c(x, \widehat{\xi}_i)$

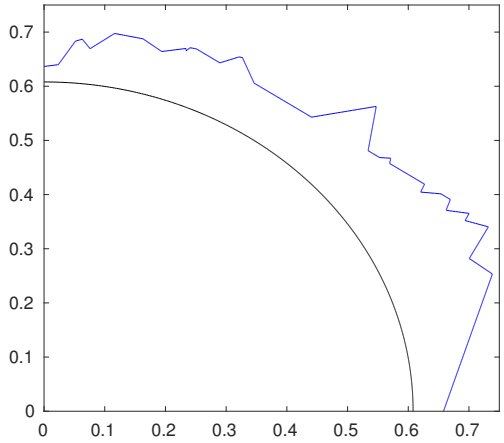
SAA approximation of quantile

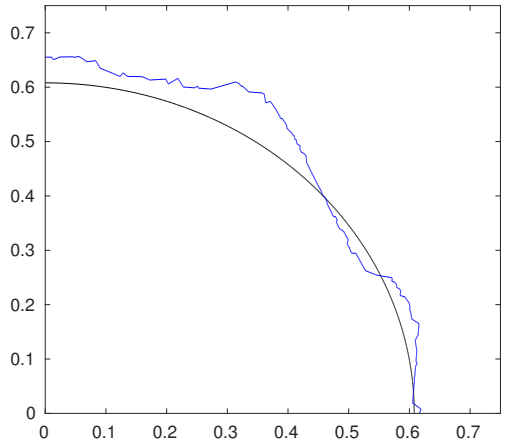
- Ensure that $M = \lceil (1-\alpha)N \rceil$ constraints are satisfied
- Scenario vector: $C(x) = (c_1(x), \ldots, c_N(x))$
- Order constraint values: $c_{[1]}(x) \le c_{[2]}(x) \le \ldots \le c_{[N]}(x)$
- Empirical quantile: $\widetilde{q}_{1-\alpha}(C(x)) = c_{[M]}(x) \le 0$

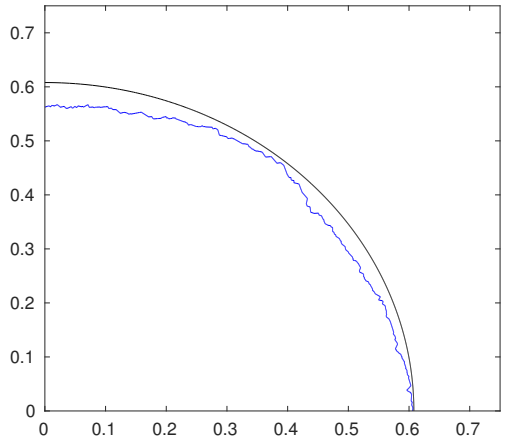












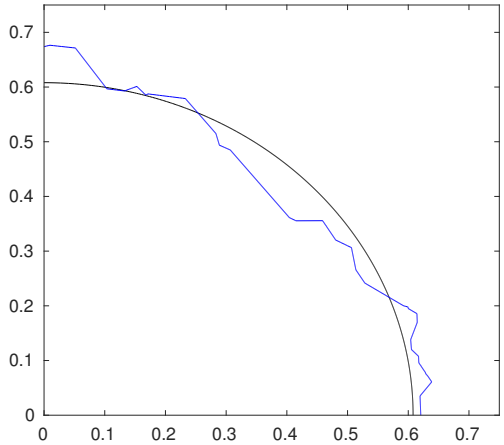
Empirical Quantile

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

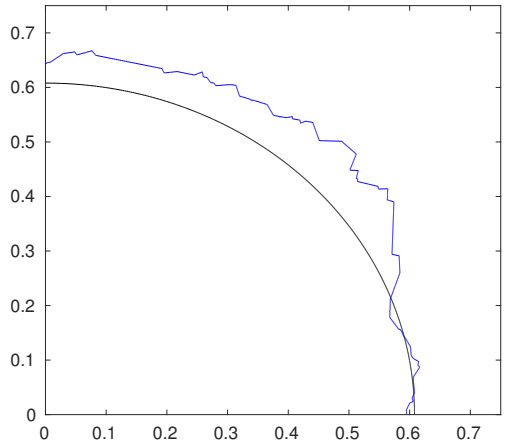
$N = 200$ $N = 500$ $N = 2000$

Empirical Quantile

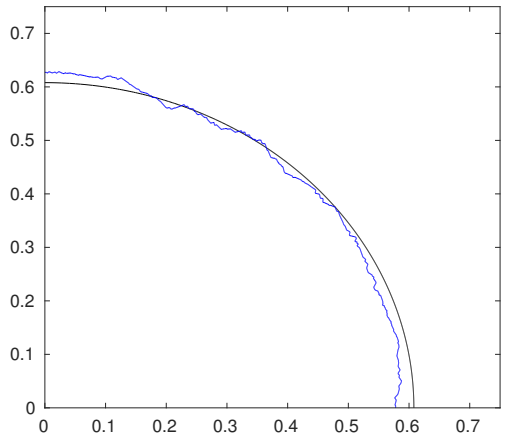
- Feasible region for $c(x, \xi) = \xi_1 x_1 + \xi_2 x_2 - 1 \qquad \xi_1, \xi_1 \sim \mathcal{N}(0, 1)$



$N = 200$



$N = 500$

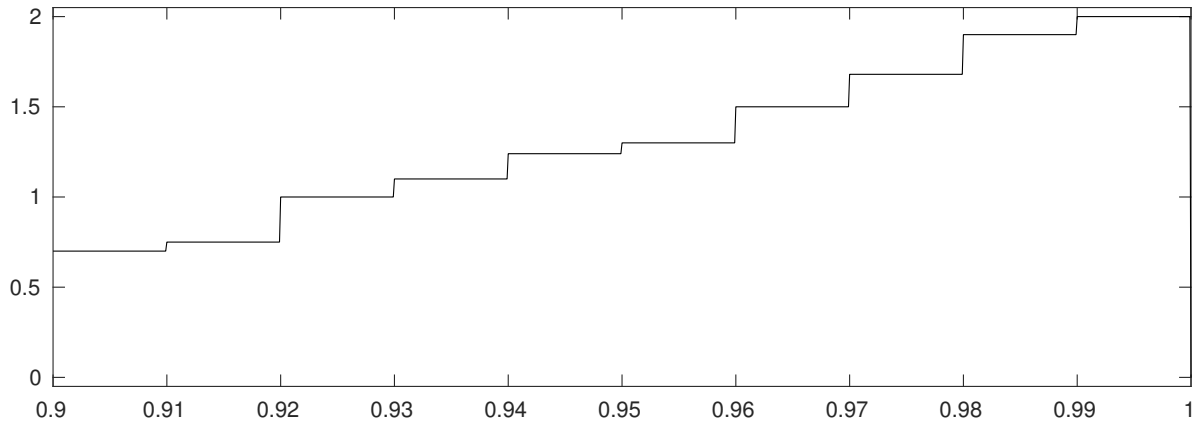


$N = 2000$

Observations:

- Approximation improves as N increases
- Rough boundary of feasible region results in spurious local minima
- A lot of variance for small N as ξ_i are resampled
- $\widetilde{q}_{1-\alpha}(C(x))$ is non-differentiable

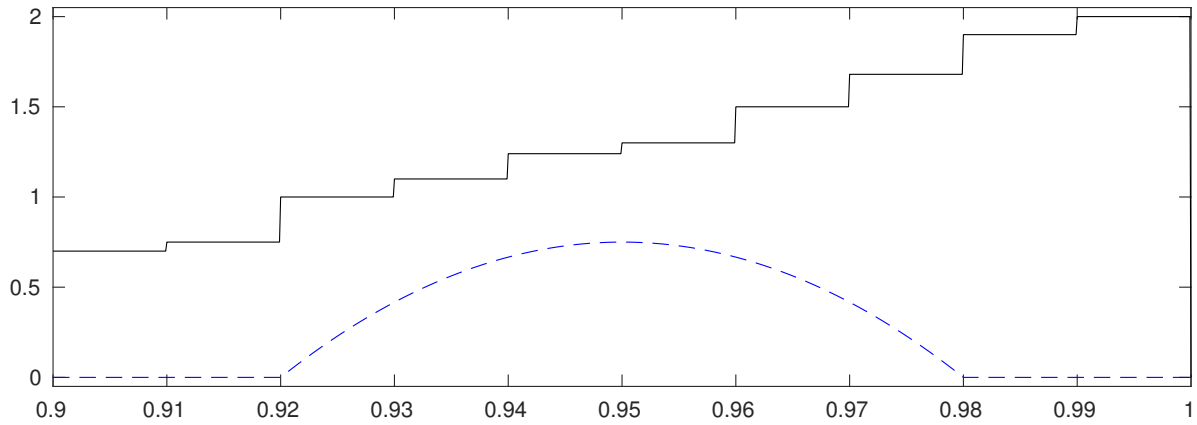
Variance Reduction using Kernels



- $Y_{[1]}, \ldots, Y_{[N]}$ ordered realizations of a random variable $(c(x, \xi_i))$
- Empirical quantiles: For $p \in [0, 1]$ define

$$\widetilde{Q}_p^Y = Y_{[j]}, \text{ where } j \in \{1, .., N\} \text{ with } \frac{j-1}{N} < p \leq \frac{j}{N}$$

Variance Reduction using Kernels



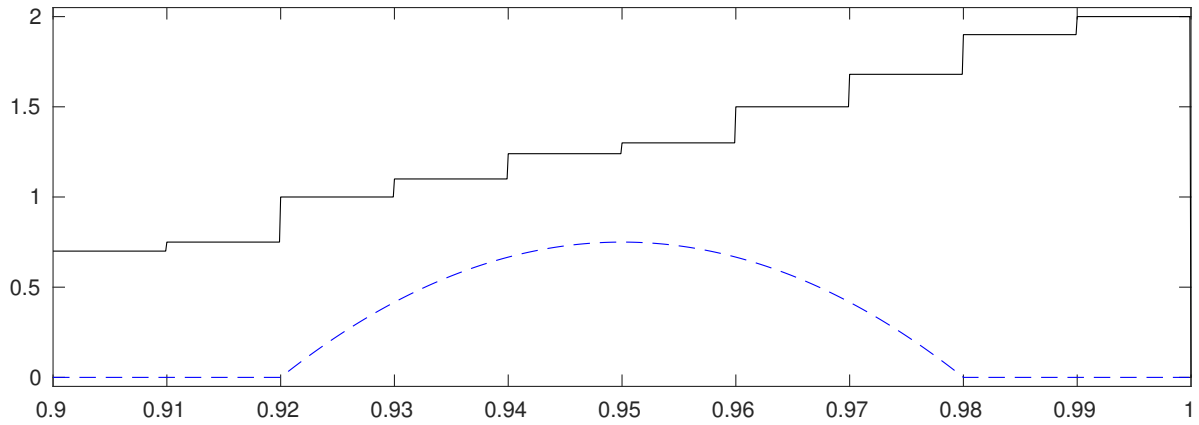
- $Y_{[1]}, \ldots, Y_{[N]}$ ordered realizations of a random variable $(c(x, \xi_i))$
- Empirical quantiles: For $p \in [0, 1]$ define

$$\widetilde{Q}_p^Y = Y_{[j]}, \text{ where } j \in \{1, .., N\} \text{ with } \frac{j-1}{N} < p \leq \frac{j}{N}$$

- Kernel smoothing [Parson 79]

$$Q_{1-\alpha}^{Y,N,h} = \int_0^1 \widetilde{Q}_p^Y \frac{1}{h} K\left(\frac{\alpha - p}{h}\right) dp$$

Variance Reduction using Kernels



- $Y_{[1]}, \ldots, Y_{[N]}$ ordered realizations of a random variable $(c(x, \xi_i))$
- Empirical quantiles: For $p \in [0, 1]$ define

$$\widetilde{Q}_p^Y = Y_{[j]}, \text{ where } j \in \{1, .., N\} \text{ with } \frac{j-1}{N} < p \leq \frac{j}{N}$$

- Kernel smoothing [Parson 79]

$$Q_{1-\alpha}^{Y,N,h} = \int_0^1 \widetilde{Q}_p^Y \frac{1}{h} K\left(\frac{\alpha - p}{h}\right) dp = \sum_{j=1}^{N} Y_{[j]} \overbrace{\int_{\frac{j-1}{N}}^{\frac{j}{N}} \frac{1}{h} K\left(\frac{\alpha - p}{h}\right) dp}^{=: w_j^{N,h}}$$

Kernel Estimation of Quantiles

- Approximate $q_{1-\alpha}(C(x)) = Q^Y_{1-\alpha}$ with $Y_i = c(x;\xi_i)$ by kernel estimate

$$\boxed{q^{N,h}_{1-\alpha}(C(x)) := \sum_{i=1}^{N} w_i^{N,h} c_{[i]}(x) \leq 0}$$

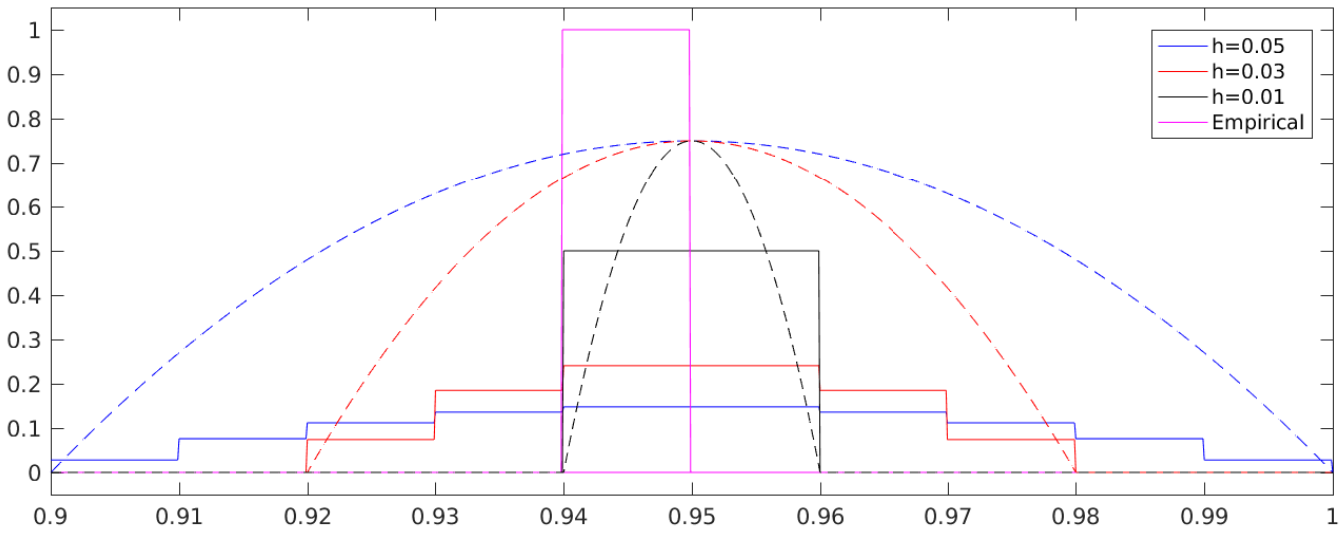
- The weights depend only on N and h and can be precomputed

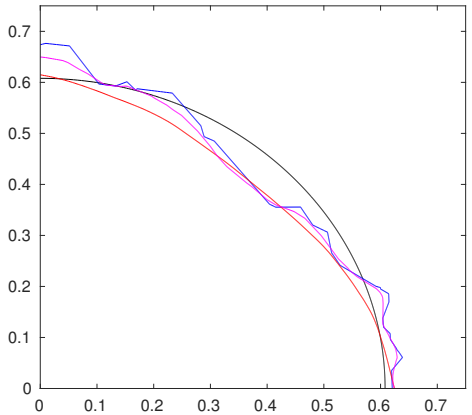
Kernel Estimation of Quantiles

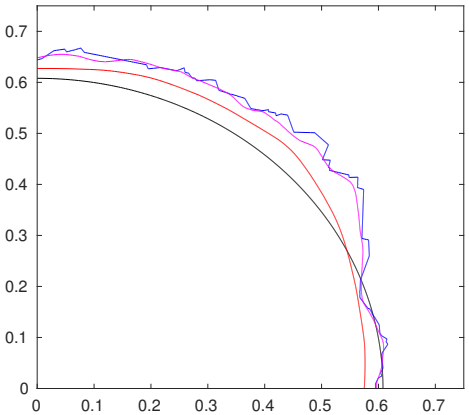
- Approximate $q_{1-\alpha}(C(x)) = Q^Y_{1-\alpha}$ with $Y_i = c(x;\xi_i)$ by kernel estimate

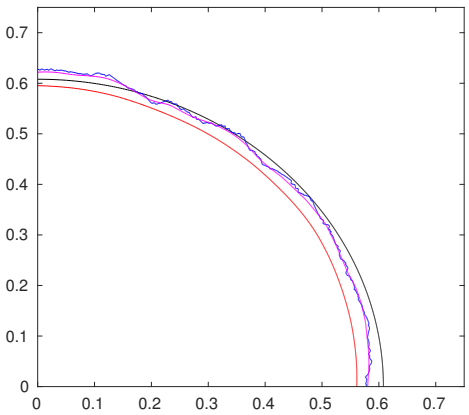
$$\boxed{q^{N,h}_{1-\alpha}(C(x)) := \sum_{i=1}^{N} w_i^{N,h} c_{[i]}(x) \leq 0}$$

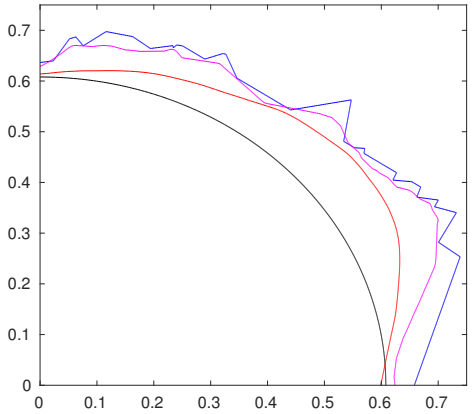
- The weights depend only on N and h and can be precomputed
- Epanechnikov kernel: $K(u) = \frac{3}{4}(1-u^2)\mathbb{1}_{|u|<1}$

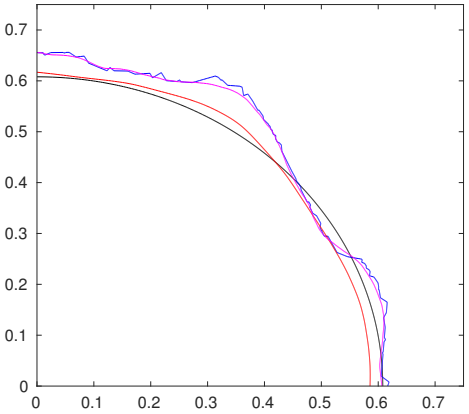


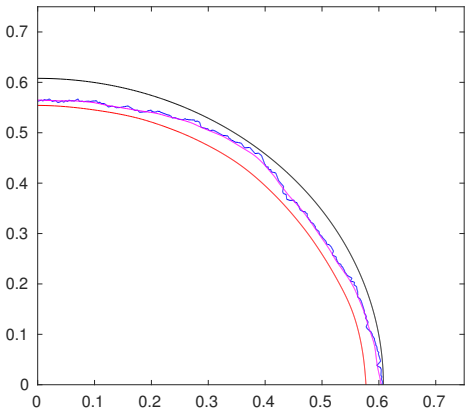












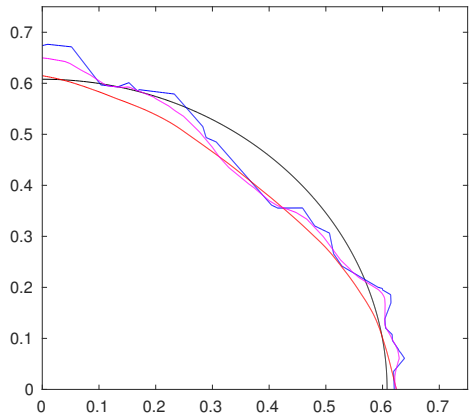
Feasible Region of Kernel-Quantile Formulation

$c(x,\xi) = \xi_1 x_1 + \xi_2 x_2 - 1 \qquad \xi_1, \xi_1 \sim \mathcal{N}(0,1) \qquad h \in \{0.05, 0.01\}$

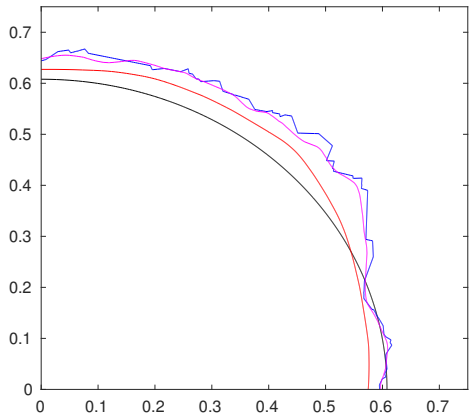
$N = 200 \qquad N = 500 \qquad N = 2000$

Feasible Region of Kernel-Quantile Formulation

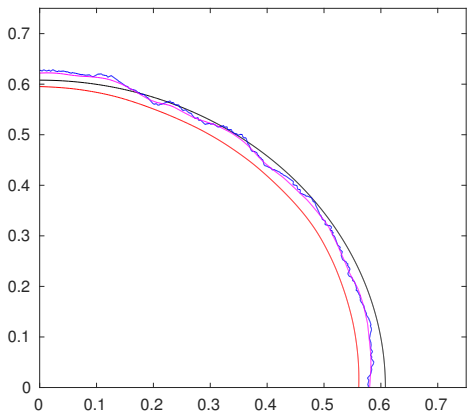
$$c(x,\xi) = \xi_1 x_1 + \xi_2 x_2 - 1 \qquad \xi_1, \xi_1 \sim \mathcal{N}(0,1) \qquad h \in \{0.05, 0.01\}$$



$N = 200$



$N = 500$

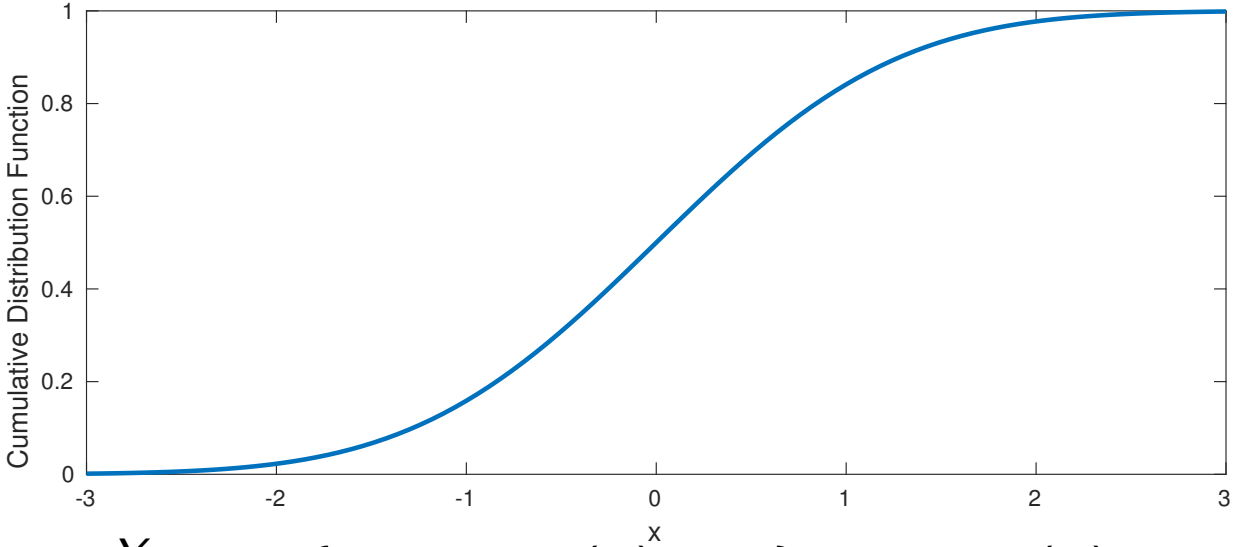


$N = 2000$

Observations:

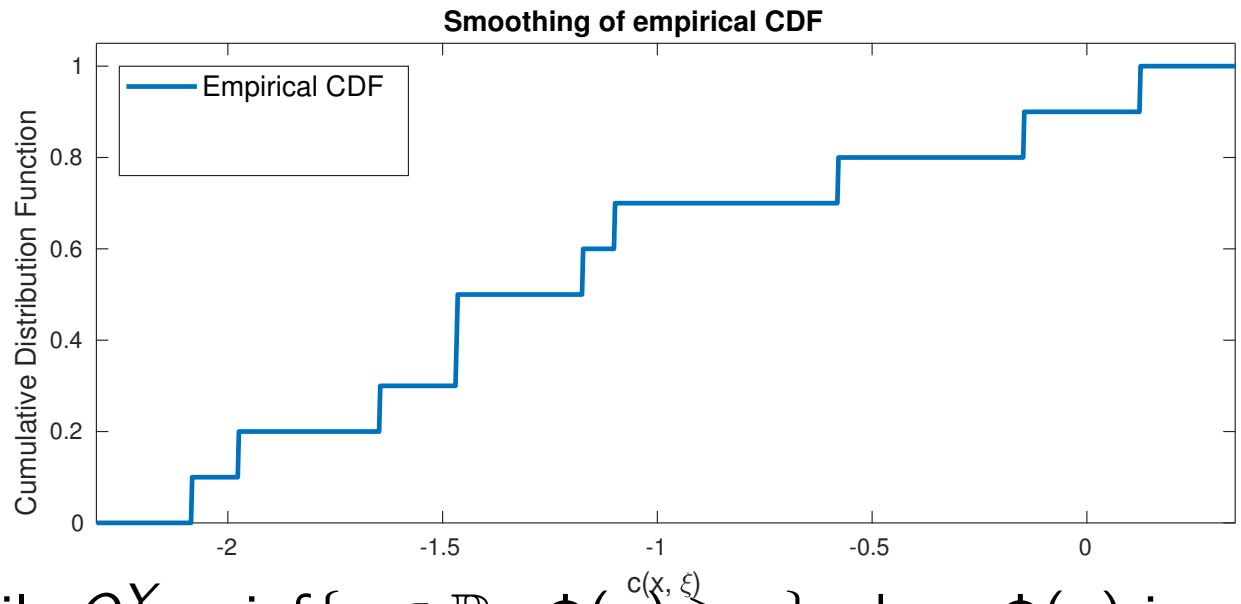
- "Looks" smoother and convex
- Reduces variance as ξ_i are resampled
- Large h creates bias
- $q_{1-\alpha}^{N,h}(C(x))$ has more points of non-differentiability
- Usually "solved" well with Knitro, but no proper termination

Smoothed Empirical CDF



- Quantile $Q_p^Y = \inf\{y \in \mathbb{R} : \Phi(y) \geq p\}$ where $\Phi(y)$ is cdf of Y

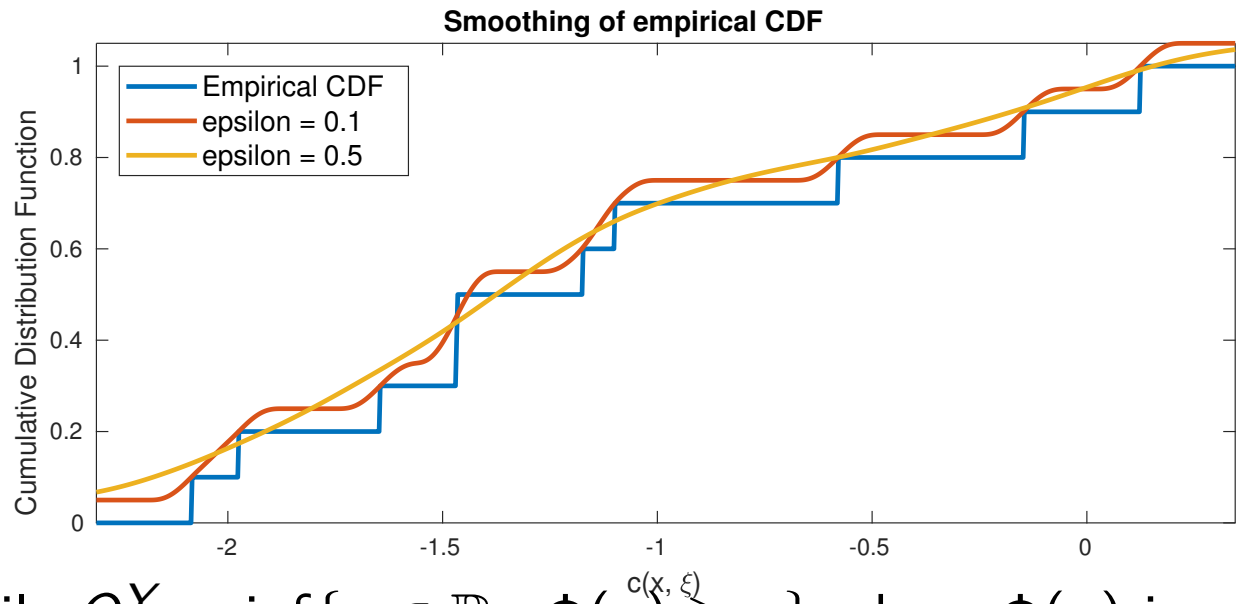
Smoothed Empirical CDF



- Quantile $Q_p^Y = \inf\{y \in \mathbb{R} : \Phi(y) \geq p\}$ where $\Phi(y)$ is cdf of Y
- $Y_1, \ldots, Y_N$ (unordered) realizations of a random variable
- Empirical cdf:

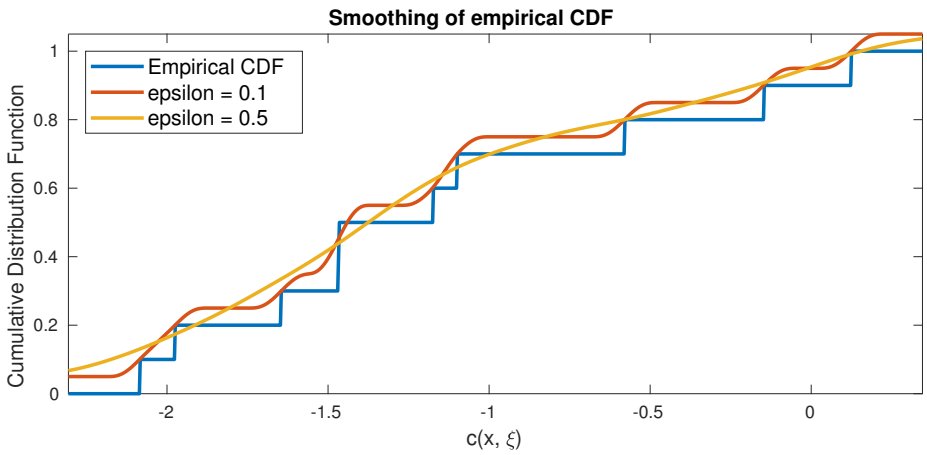
$$\widetilde{\Phi}(y) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\geq 0}(y - Y_i)$$

Smoothed Empirical CDF



- Quantile $Q_p^Y = \inf\{y \in \mathbb{R} : \Phi(y) \geq p\}$ where $\Phi(y)$ is cdf of Y
- $Y_1, \ldots, Y_N$ (unordered) realizations of a random variable
- Empirical cdf:
$$\widetilde{\Phi}(y) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\geq 0}(y - Y_i)$$
- Smoothed cdf [Azzalini 81]: ($K_\epsilon(t) \approx \mathbb{1}_{\geq 0}(t)$ differentiable)
$$\Phi^{N,\epsilon}(y) = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - Y_i)$$

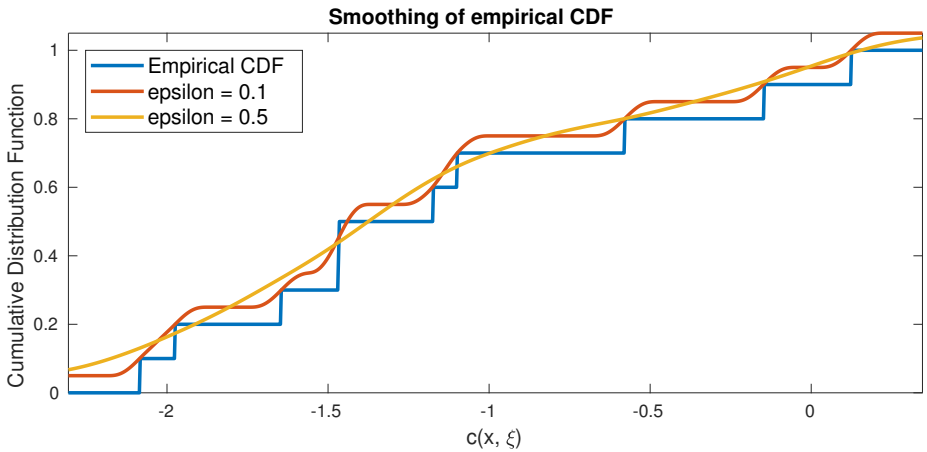
Differentiable Quantile Estimates



- $Y_i = c(x, \xi_i)$
- Smoothed cdf: $\Phi^{N,\epsilon}(y) = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - c_i(x))$
- Resulting quantile estimate $\widehat{q}_p^{N,\epsilon}(C(x))$ is the solution y^* of

$$\boxed{p = \Phi^{N,\epsilon}(y)}$$

Differentiable Quantile Estimates

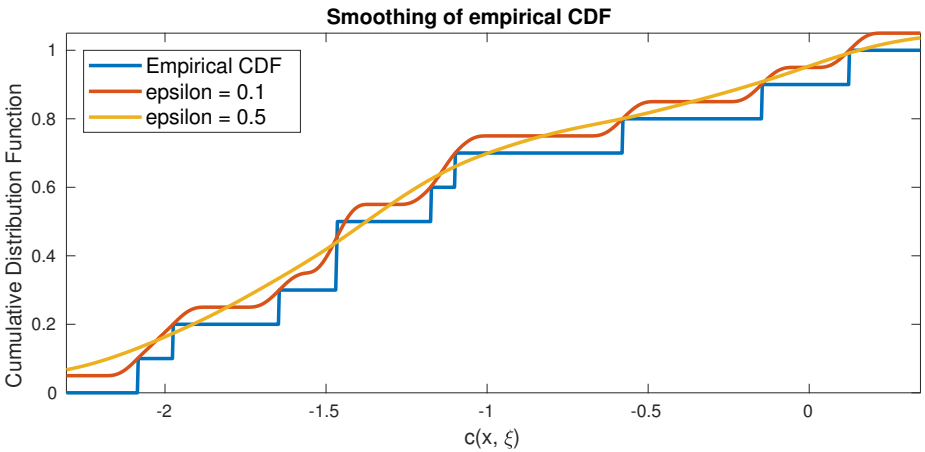


- $Y_i = c(x, \xi_i)$
- Smoothed cdf: $\Phi^{N,\epsilon}(y) = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - c_i(x))$
- Resulting quantile estimate $\widehat{q}_p^{N,\epsilon}(C(x))$ is the solution y^* of

$$\boxed{p = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - c_i(x))}$$

- Implicit function theorem guarantees that $\widehat{q}_p^{N,\epsilon}(C(x))$ exists and is differentiable

Differentiable Quantile Estimates



- $Y_i = c(x, \xi_i)$
- Smoothed cdf: $\Phi^{N,\epsilon}(y) = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - c_i(x))$
- Resulting quantile estimate $\widehat{q}_p^{N,\epsilon}(C(x))$ is the solution y^* of

$$\boxed{p = \frac{1}{N}\sum_{i=1}^{N} K_\epsilon(y - c_i(x)) + \frac{1}{2N}}$$

- Implicit function theorem guarantees that $\widehat{q}_p^{N,\epsilon}(C(x))$ exists and is differentiable (for $p \in \{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}\}$)

Final Estimate

1. Kernel estimate using empirical quantiles $c_{[i]}(x)$

$$\boxed{q_{1-\alpha}^{N,h}(C(x)) = \sum_{i=1}^{N} w_i^{N,h} c_{[i]}(x)}$$

Final Estimate

1. Kernel estimate using empirical quantiles $c_{[i]}(x)$

$$q_{1-\alpha}^{N,h}(C(x)) = \sum_{i=1}^{N} w_i^{N,h} c_{[i]}(x)$$

2. Smoothed empirical quantile $\widehat{q}_p^{N,\epsilon}(C(x))$, solution of

$$p = \frac{1}{N} \sum_{i=1}^{N} K_\epsilon(y - c_i(x)) + \frac{1}{2N}$$

Final Estimate

1. Kernel estimate using empirical quantiles $c_{[i]}(x)$

$$q_{1-\alpha}^{N,h}(C(x)) = \sum_{i=1}^{N} w_i^{N,h} c_{[i]}(x)$$

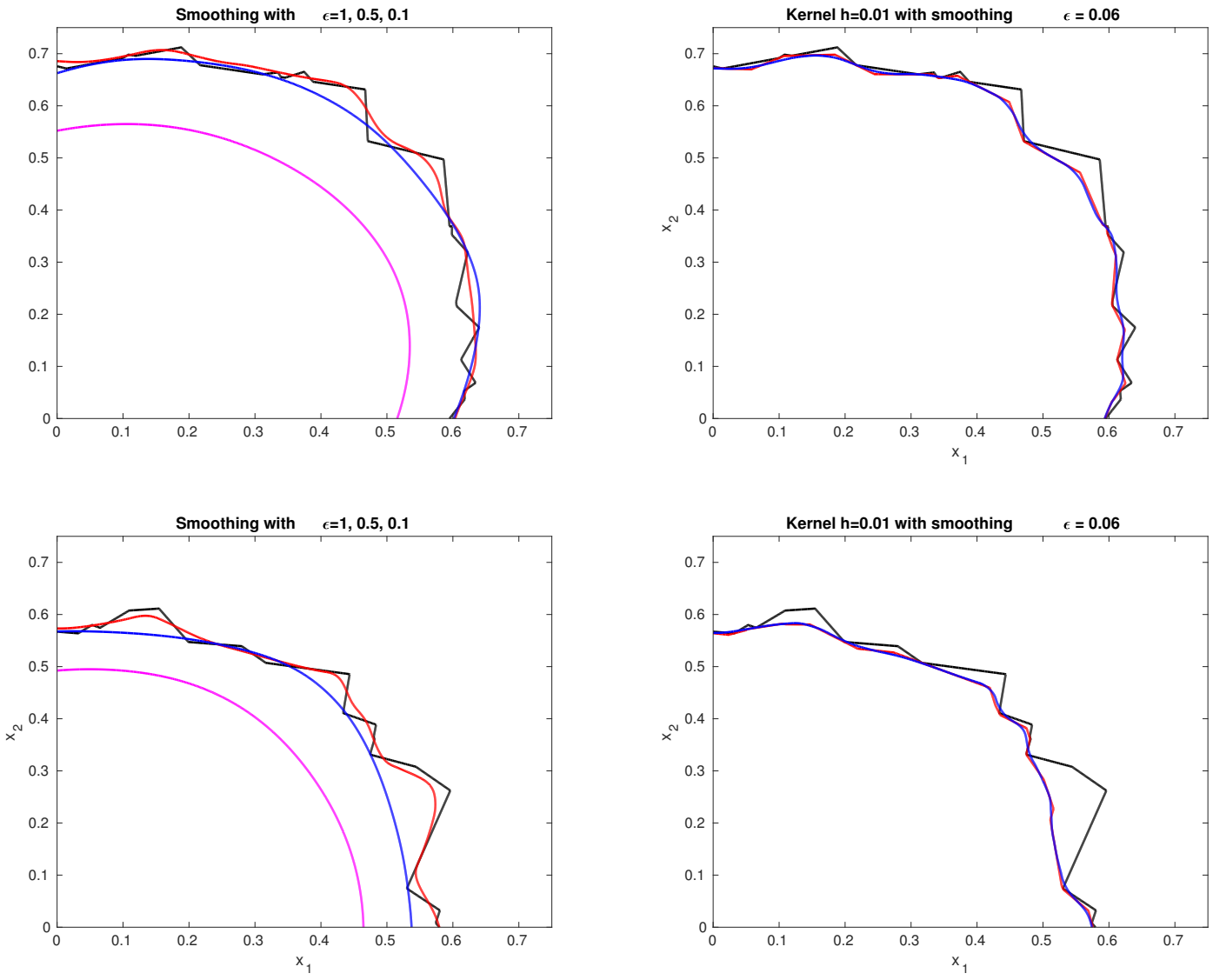
2. Smoothed empirical quantile $\widehat{q}_p^{N,\epsilon}(C(x))$, solution of

$$p = \frac{1}{N} \sum_{i=1}^{N} K_\epsilon(y - c_i(x)) + \frac{1}{2N}$$

3. Final estimate

$$q_{1-\alpha}^{N,h,\epsilon}(C(x)) := \sum_{i=1}^{N} w_i^{N,h} \, \widehat{q}_{\frac{i}{N}}^{N,\epsilon}(C(x))$$

Feasible Region with Smoothing ($N = 100$)



Example: Portfolio Optimization

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

- $r \sim \mathcal{N}(\mu, \Sigma)$, fixed $\mu \in \mathbb{R}^{100}$ and $\Sigma \in \mathbb{R}^{100 \times 100}$ pos. def.
- 100 replications

Example: Portfolio Optimization

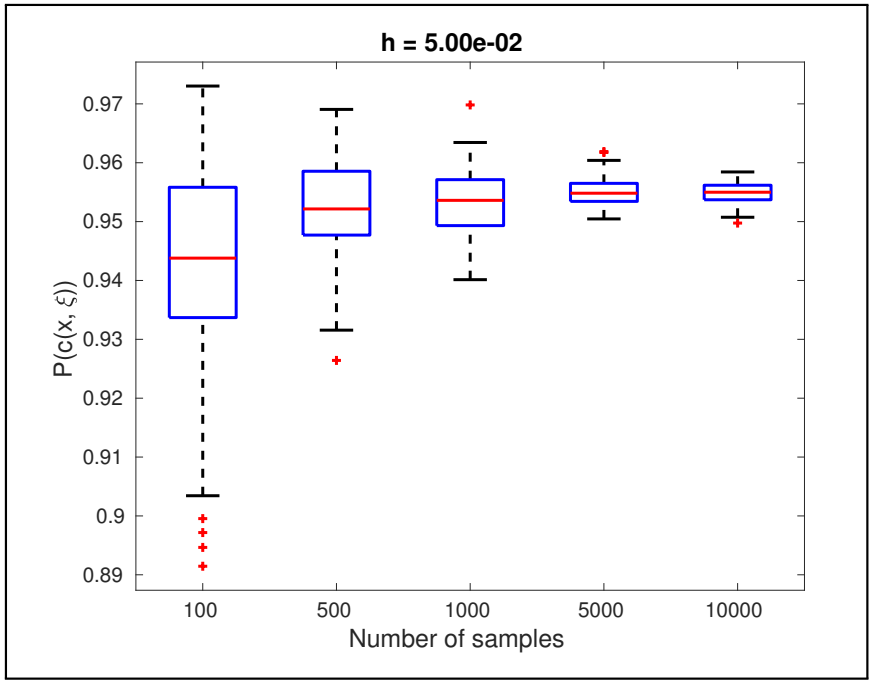
$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & \mathbb{E}[r^T x] = \mu^T x \\ \text{s.t.} \quad & \mathbb{P}_r[r^T x \geq -0.05] \geq 0.95 \\ & \sum_{i=1}^{n} x_i = 1, \qquad x \geq 0 \end{aligned}$$

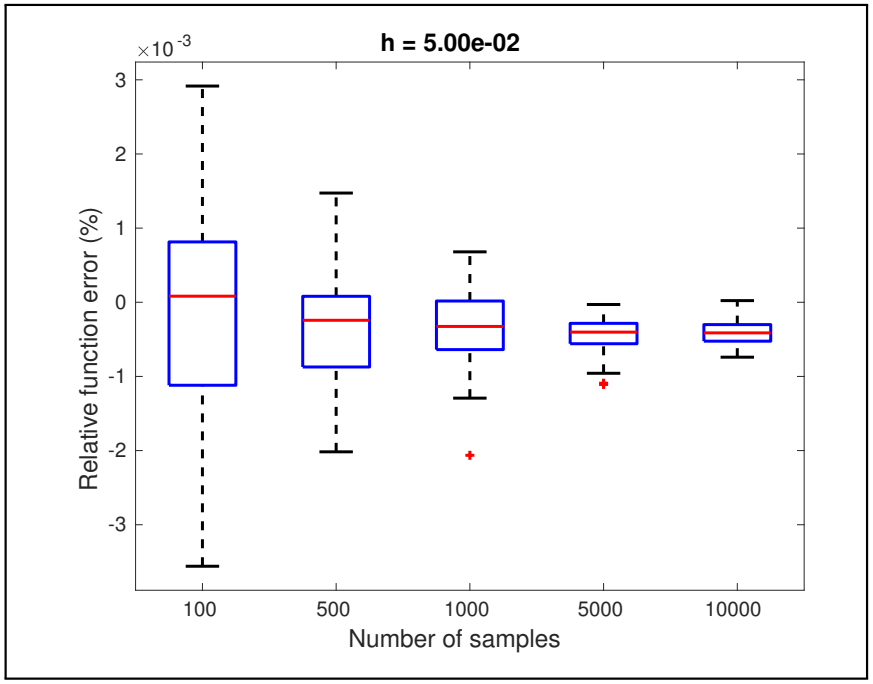
- $r \sim \mathcal{N}(\mu, \Sigma)$, fixed $\mu \in \mathbb{R}^{100}$ and $\Sigma \in \mathbb{R}^{100 \times 100}$ pos. def.
- 100 replications

Iterations with Knitro

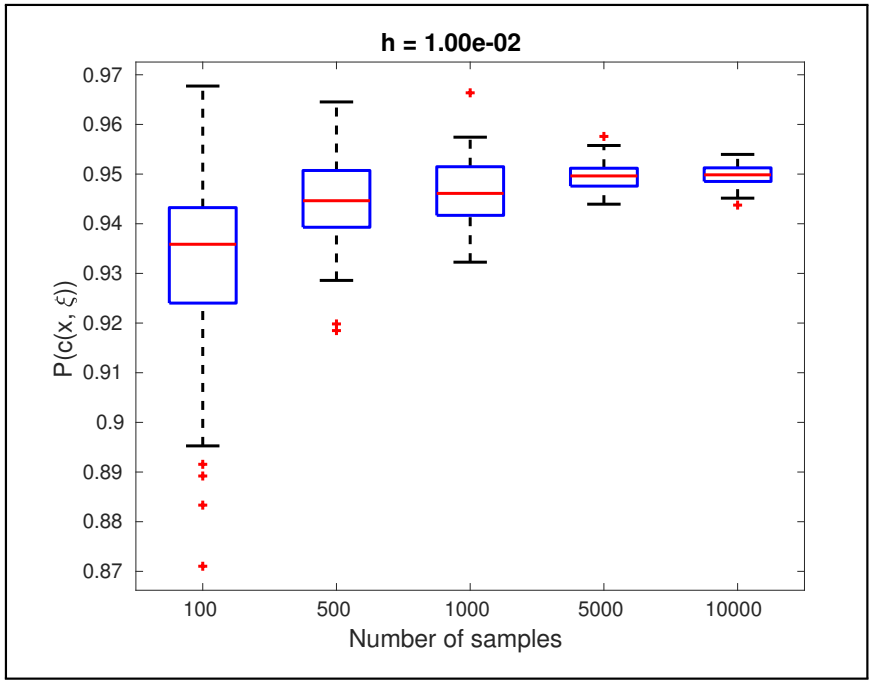
N	h	min iter	mean iter	max iter
500	0.05	15	26.7	54
10,000	0.005	17	29.3	45

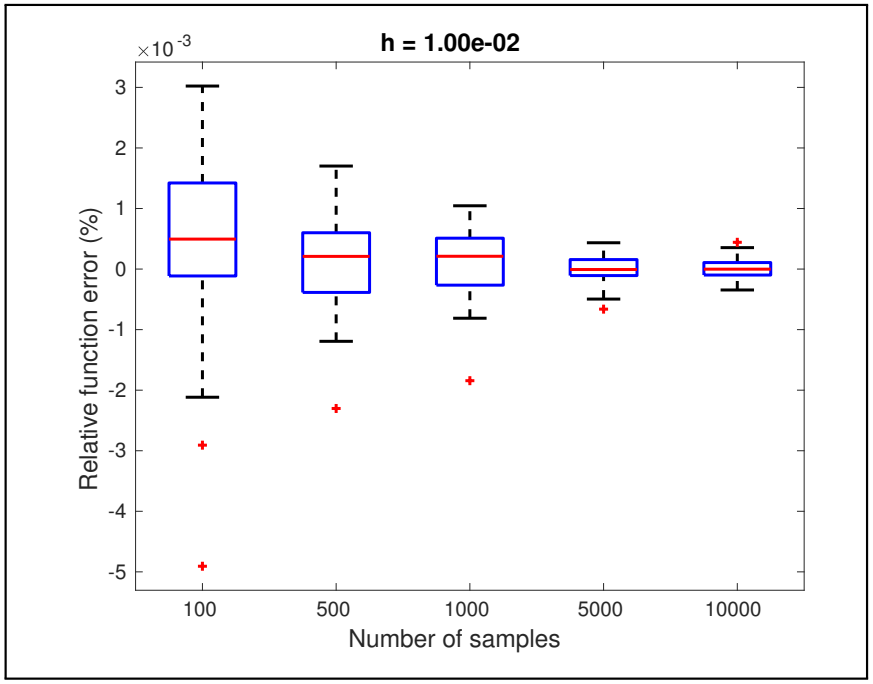
Portfolio Optimization



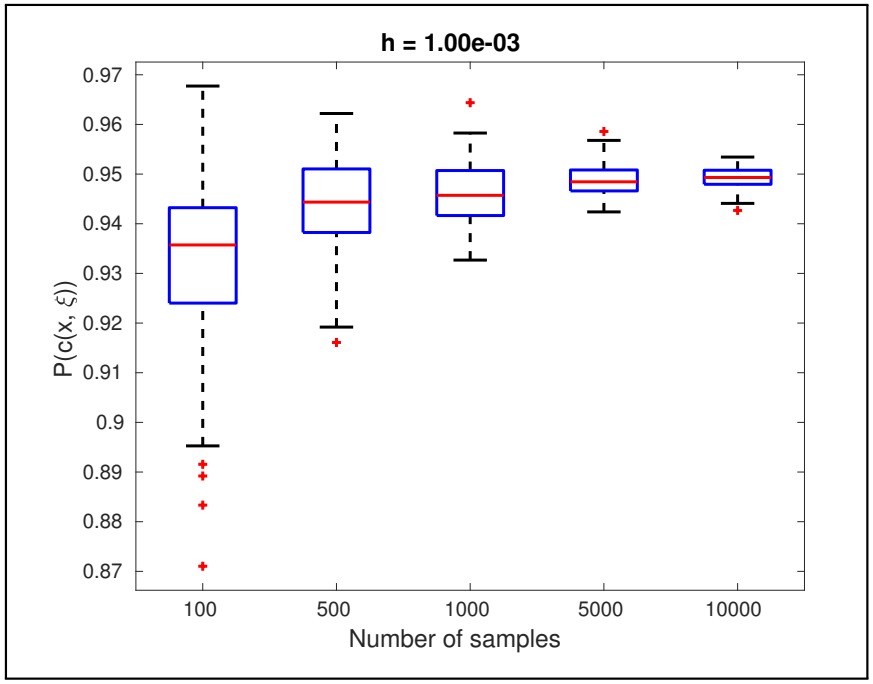


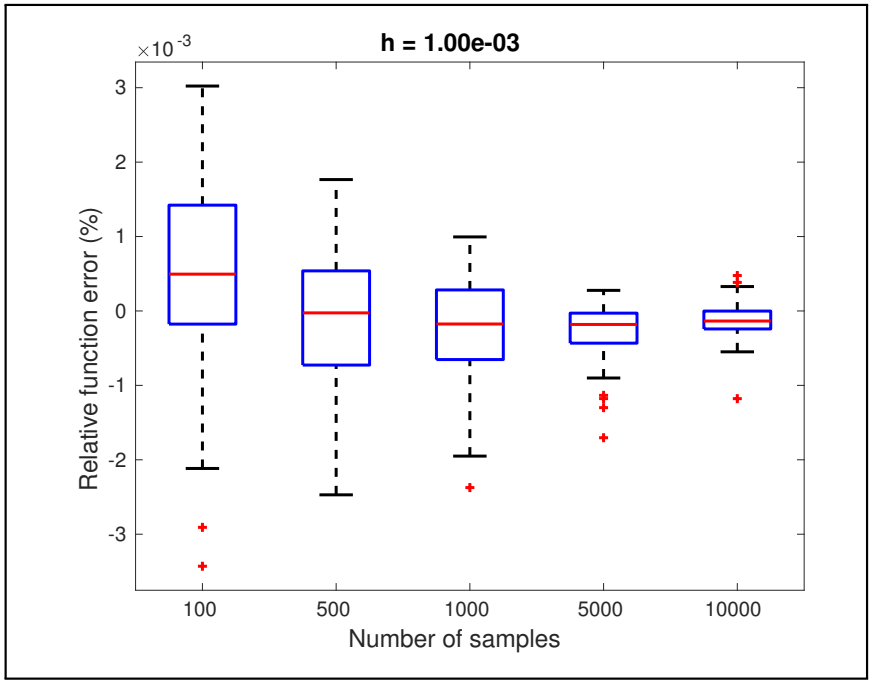
Portfolio Optimization



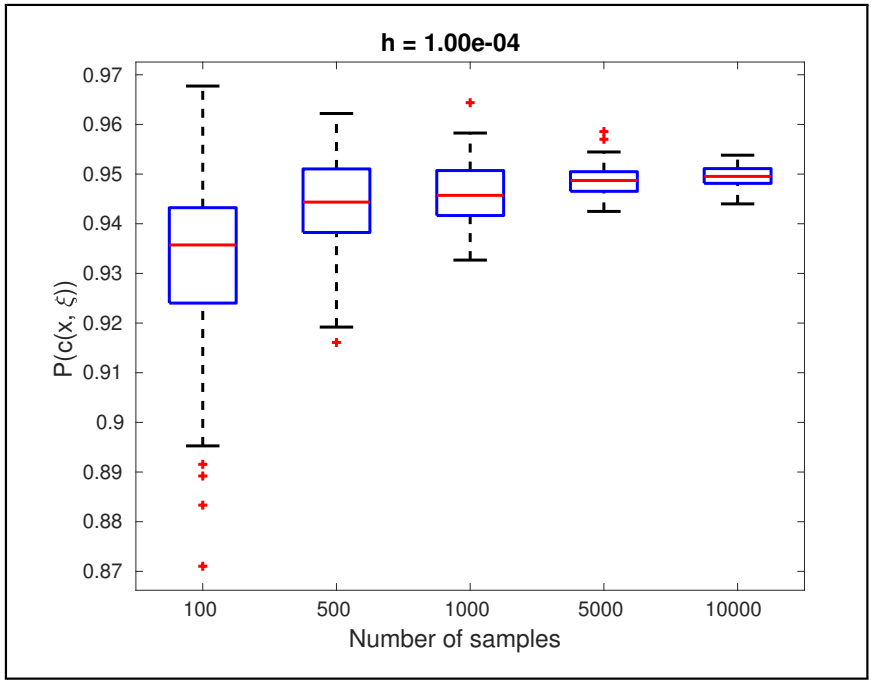


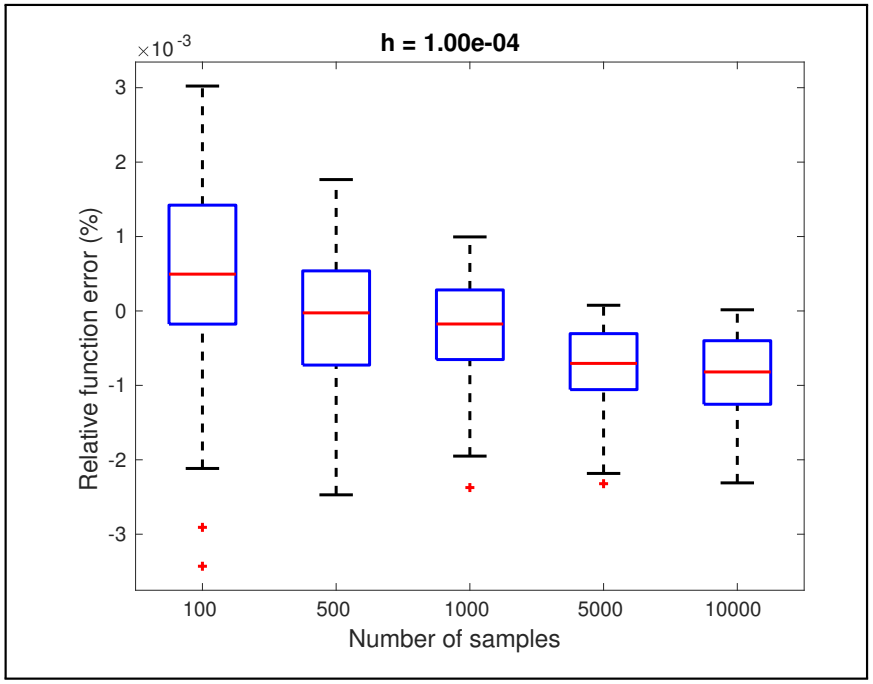
Portfolio Optimization





Portfolio Optimization





Joint Chance Constraints

- Original Problem

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[c(x,\xi) \leq 0] \geq 1 - \alpha \end{aligned}}$$

- Scenario vector

$$C(x) = (c(x,\xi_1), \ldots, c(x,\xi_N))$$

- NLP formulation

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } q_{1-\alpha}^{N,h,\epsilon}(C(x)) \leq 0 \end{aligned}}$$

Joint Chance Constraints

- Original Problem

$$\begin{aligned} \min_{x \in X} \; & f(x) \\ \text{s.t. } & \mathbb{P}_\xi[c_j(x,\xi) \le 0, \; j = 1, \ldots, m] \ge 1 - \alpha \end{aligned}$$

- Scenario vector

$$C(x) = (c(x,\xi_1), \ldots, c(x,\xi_N))$$

- NLP formulation

$$\begin{aligned} \min_{x \in X} \; & f(x) \\ \text{s.t. } & q_{1-\alpha}^{N,h,\epsilon}(C(x)) \le 0 \end{aligned}$$

Joint Chance Constraints

- Original Problem

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{P}_\xi[\widehat{c}(x,\xi) \leq 0] \geq 1-\alpha \end{aligned}}$$

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

- Scenario vector

$$C(x) = (c(x,\xi_1), \ldots, c(x,\xi_N))$$

- NLP formulation

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } q_{1-\alpha}^{N,h,\epsilon}(C(x)) \leq 0 \end{aligned}}$$

Joint Chance Constraints

- Original Problem

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } \mathbb{P}_\xi[\widehat{c}(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

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

- Scenario vector

$$\widehat{C}(x) = (\widehat{c}(x,\xi_1), \ldots, \widehat{c}(x,\xi_N))$$

- NLP formulation

$$\boxed{\begin{aligned} &\min_{x \in X} \; f(x) \\ &\text{s.t. } q_{1-\alpha}^{N,h,\epsilon}(\widehat{C}(x)) \le 0 \end{aligned}}$$

Note: Constraint no longer differentiable

Joint Chance Constraints

- Original Problem

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } \mathbb{P}_\xi[\widehat{c}(x,\xi) \le 0] \ge 1-\alpha \end{aligned}}$$

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

- Scenario vector

$$\widehat{C}(x) = (\widehat{c}(x,\xi_1), \ldots, \widehat{c}(x,\xi_N))$$

- NLP formulation

$$\boxed{\begin{aligned} &\min_{x \in X} f(x) \\ &\text{s.t. } q(\widehat{C}(x)) \le 0 \end{aligned}}$$

$$q(\widehat{C}(x)) := q_{1-\alpha}^{N,h,\epsilon}(\widehat{C}(x))$$

Note: Constraint no longer differentiable

Exact Penalty Function

$$\min_{x \in X} f(x)$$
$$\text{s.t. } q(\widehat{C}(x)) \leq 0$$

- Exact penalty function:

$$\phi_\rho(x) = f(x) + \rho \max\{q(\widehat{C}(x)), 0\} \quad (+\text{penalty for } X)$$

Exact Penalty Function

$$\min_{x \in X} f(x)$$
$$\text{s.t. } q(\widehat{C}(x)) \leq 0$$

- Exact penalty function:

$$\phi_\rho(x) = f(x) + \rho \max\{q(\widehat{C}(x)), 0\} \quad (+\text{penalty for } X)$$

- Model of scenario vector $\widehat{C}(x) = (\ldots, \max_j c_j(x, \xi_i), \ldots)$

$$m_{\widehat{C}}(x_k, d) = \left(\ldots, \max_j \{c_j(x_k, \xi_i) + \nabla c_j(x_k, \xi_i)^T d\}, \ldots\right)$$

Exact Penalty Function

$$\min_{x \in X} f(x)$$
$$\text{s.t. } q(\widehat{C}(x)) \leq 0$$

- Exact penalty function:

$$\phi_\rho(x) = f(x) + \rho \max\{q(\widehat{C}(x)), 0\} \quad (+\text{penalty for } X)$$

- Model of scenario vector $\widehat{C}(x) = (\ldots, \max_j c_j(x, \xi_i), \ldots)$

$$m_{\widehat{C}}(x_k, d) = \left(\ldots, \max_j \{c_j(x_k, \xi_i) + \nabla c_j(x_k, \xi_i)^T d\}, \ldots\right)$$

- Model of penalty function

$$\begin{aligned} m_\phi(x_k, d) = & f(x_k) + \nabla f(x_k)^T d + \tfrac{1}{2} d^T B_k d + \\ & \rho \max\{q(\widehat{C}(x_k)) + \nabla q(\widehat{C}(x_k))^T (m_{\widehat{C}}(x_k, d) - \widehat{C}(x_k)), 0\} \end{aligned}$$

Trust Region Sℓ_1QP Algorithm

$$\boxed{\begin{aligned} m_\phi(x_k, d) =& f(x_k) + \nabla f(x_k)^T d + \tfrac{1}{2} d^T B_k d + \\ & \rho \max\{q(\widehat{C}(x_k)) + \nabla q(\widehat{C}(x_k))^T (m_{\widehat{C}}(x_k, d) - \widehat{C}(x_k)), 0\} \end{aligned}}$$

- Can show: $|\phi_\rho(x_k) - m_\phi(x_k, d)| = O(\|d\|^2)$
- Standard Sℓ_1QP-type trust region algorithm can be applied

Trust Region Sℓ_1QP Algorithm

$$m_\phi(x_k, d) = f(x_k) + \nabla f(x_k)^T d + \tfrac{1}{2} d^T B_k d + \rho \max\{q(\widehat{C}(x_k)) + \nabla q(\widehat{C}(x_k))^T (m_{\widehat{C}}(x_k, d) - \widehat{C}(x_k)), 0\}$$

- Can show: $|\phi_\rho(x_k) - m_\phi(x_k, d)| = O(\|d\|^2)$
- Standard Sℓ_1QP-type trust region algorithm can be applied
- QP Subproblem

$$\begin{aligned} \min_{d,z,t} \; & f(x_k) + \nabla f(x_k)^T d + \tfrac{1}{2} d^T B_k d + \rho\, t \\ \text{s.t.} \; & q(\widehat{C}(x_k)) + \nabla q(\widehat{C}(x_k))^T Z \le t \qquad [Z = (z_1, \ldots, z_n)] \\ & c_j(x_k, \xi_i) + \nabla c_j(x_k, \xi_i)^T d \le z_i \qquad \text{for all } i, j \\ & t \ge 0, \quad x_k + d \in \tilde{X}, \qquad \|d\|_\infty \le \Delta_k \end{aligned}$$

- Size of QP proportional to number of nonzeros in $\nabla q(\widehat{C}(x_k))$, not N

Summary

- Goal: Formulate chance-constrained problem as NLP
- Constrain the quantile $q_{1-\alpha}(x) \leq 0$
  - Better linearization than $p(x) \geq (1-\alpha)$
- Sample-based empirical quantile
  - No need to know probability distribution
- Variance reduction using kernel
  - Improves approximation of feasible region, "smoother boundary"
- Quantile estimates using smoothed empirical cdf
  - Constraint becomes differentiable
- Joint chance constraints
  - Exact merit function
  - Sℓ_1QP-type trust region algorithm