Probabilistic Constraints in Optimization Problems on Flow Networks

TECHNISCHE UNIVERSITÄT DARMSTADT

Elisa Strauch, Jens Lang, Martin Gugat, Michael Schuster

Research Group Numerical Analysis and Scientific Computing, TU Darmstadt
Lehrstuhl für Angewandte Analysis, Department of Mathematics, FAU Erlangen-Nürnberg

This work is supported by the German Research Foundation (DFG) via grant SFB-Transregio 154.

Structure

Probabilistic Constraints in Optimization Problems

TECHNISCHE UNIVERSITÄT DARMSTADT

Optimization problem with a probabilistic constraint

$$\min f(x)$$
$$\text{s.t. } \mathbb{P}(g(x,\xi) \leq 0) \geq \alpha$$

with n-dimensional random variable ξ and $\alpha \in [0,1]$.

How to compute the probability for given x?

Spheric radial decomposition (SRD)

- Consider feasible set $M := \{\xi \in \mathbb{R}^n | g(x,\xi) \leq 0\} \to \mathbb{P}(\xi \in M)$
- Write $\mathbb{P}(\xi \in M)$ as integral over the unit sphere $\mathbb{S}^{n-1}$

Direct approach

- Integration of the probability density function (pdf) of $g(x,\cdot)$
- Estimation of the pdf via kernel density estimation (KDE)

Spheric Radial Decomposition

TECHNISCHE UNIVERSITÄT DARMSTADT

Theorem (SRD)

Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ be n-dimensional Gaussian distributed random variable with mean μ and positive definite correlation matrix Σ.
Then, for any Borel measurable subset $M \subset \mathbb{R}^n$ it holds that

$$\mathbb{P}(\xi \in M) = \int_{\mathbb{S}^{n-1}} \mu_\chi \{r \geq 0 | rLw + \mu \in M\} d\mu_\eta(w)$$

where $\mathbb{S}^{n-1}$ is the $(n-1)$-dimensional sphere in $\mathbb{R}^n$, μ_η is the uniform distribution on $\mathbb{S}^{n-1}$, μ_χ denotes the χ-distribution with n degrees of freedom and L is such that $\Sigma = LL^T$.

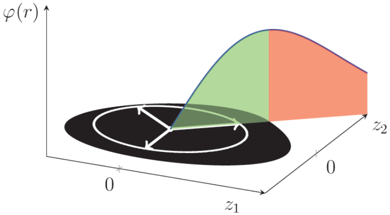
Spheric Radial Decomposition

Theorem: $\mathbb{P}(\xi \in M) = \int_{\mathbb{S}^{n-1}} \mu_\chi\{r \geq 0 | rLw + \mu \in M\} d\mu_\eta(w)$

Algorithm

Let $\xi \sim \mathcal{N}(\mu, \Sigma)$ and L such that $LL^T = \Sigma$.

1. Sample Q points $\{w_1, \dots, w_Q\}$ uniformly distributed on the sphere $\mathbb{S}^{n-1}$.
2. Compute the one-dimensional sets $M_i := \{r \geq 0 | rLw_i + \mu \in M\}$ for $i = 1, \dots, Q$.
3. Set $\mathbb{P}(\xi \in M) \approx (1/Q) \sum_{i=1}^{Q} \mu_\chi(M_i)$

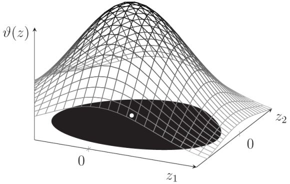


Kernel Density Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Direct approach: Integration of the probability density function ϱ_g of $g := g(x, \cdot)$

$$\mathbb{P}(g \leq 0) = \int_{-\infty}^{0} \varrho_g(z) dz.$$



Pdf ϱ is in general unknown $\rightarrow$ **KDE** provides approximation $\varrho_{g,N}$

$$\mathbb{P}_N(g \leq 0) = \int_{-\infty}^{0} \varrho_{g,N}(z) dz.$$

Kernel Density Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Definition (KDE)

- Let y be a n-dimensional real-valued random variable with an absolutely continuous distribution and probability density function ϱ.
- Let $\mathcal{Y} = \{y^{S,1}, \dots, y^{S,N}\}$ be an independent and identically distributed sampling of y.

→ The kernel density estimator $\varrho_N : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is defined as

$$\varrho_N(z) = \frac{1}{N \det(H)^{1/2}} \sum_{i=1}^{N} K(H^{\frac{1}{2}}(z - y^{S,i}))$$

with a symmetric positive definite bandwidth matrix $H \in \mathbb{R}^{n \times n}$ and a n-variate density $K : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ called kernel.

Kernel Density Estimation

KDE: $\varrho_N(z) = \frac{1}{N \det(H)^{1/2}} \sum_{i=1}^{N} K(H^{\frac{1}{2}}(z - y^{S,i}))$

Heuristic Bandwidth Matrix H

Let $\sigma_{N,i}^2$ denotes the sample variance of the i-th variable,

$$H = \left(\frac{4}{(n+2)N}\right)^{\frac{2}{n+4}} \begin{bmatrix} \sigma_{N,1}^2 & & \\ & \ddots & \\ & & \sigma_{N,n}^2 \end{bmatrix}$$

Product Gaussian Kernel

$$K(x) = \prod_{i=1}^{n} \mathcal{K}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_i}{2}\right)$$

Kernel Density Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

L^1-Convergence

$$\|\varrho_N - \varrho\|_{L^1} \xrightarrow{N\to\infty} 0 \quad \mathbb{P} - \text{almost surely}$$

Convergence of $\mathbb{P}_N(g \in D)$

Applying Scheffé's Lemma yields

$$\begin{aligned}|\mathbb{P}(g \in D) - \mathbb{P}_N(g \in D)| &= \left|\int_D \varrho(z)dz - \int_D \varrho_N(z)dz\right| \\ &\leq \frac{1}{2}\|\varrho_N - \varrho\|_{L^1} \xrightarrow{N\to\infty} 0 \quad \mathbb{P} - \text{almost surely}\end{aligned}$$

Stationary Gas Network

TECHNISCHE UNIVERSITÄT DARMSTADT

Stationary semi-linear isothermal Euler Equation

$$\partial_x q = 0,$$
$$c^2 \partial_x p = -\frac{\lambda}{2D}(R_s T)^2 \frac{|q|q}{p}$$

- Constant speed of sound c, horizontal pipes, ideal gases ($p = \rho R_s T$)
- Pressure p, gas flow q, friction coefficient λ, specific gas constant R_s, constant temperature T, pipe diameter D

Analytic solution

$$q = \text{const.},$$
$$p(L) = \sqrt{p(0)^2 - \frac{\lambda R_s^2 T^2}{c^2 D} L|q|q},$$

with L be the length of pipe.

Stationary Gas Network

Network modelling

- Tree-structured graph with only one inflow node
- Stationary Euler equation models gas flow through pipes
- Boundary and coupling conditions (conservation of mass and equality of pressure)

Random gas demand

- Sink $v_i \in \mathcal{V}_{out}$ with $|\mathcal{V}_{out}| = n$ and boundary condition $q(v_i) = b_i$
- Gaussian distributed gas demands $b \sim \mathcal{N}(\mu, \Sigma)$ with $\mu \in \mathbb{R}^n_{\geq 0}$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$
  $\rightarrow$ n-dimensional random vector $p_{out} := (p(v_1), \dots, p(v_n))$, $v_i \in \mathcal{V}_{out}$
- Motivated by the application, μ and Σ are chosen s.t. $\mathbb{P}(b > 0) \approx 1$.

Stationary Gas Network

TECHNISCHE UNIVERSITÄT DARMSTADT

- Interested in $\mathbb{P}(p_{out} \in P_{\min}^{\max})$ with $P_{\min}^{\max} = \bigotimes_{i=1}^{n}[p_i^{\min}, p_i^{\max}]$

SRD

- Define set of feasible outflows $M := \left\{b \in \mathbb{R}_{\geq 0}^{n} \middle| p_{out} \in P_{\min}^{\max}\right\}$
  $\rightarrow \mathbb{P}(b \in M) = \mathbb{P}(p_{out} \in P_{\min}^{\max})$

Direct approach - KDE

- Assumption: pressure p_{out} has a continuous distribution with probability density function ϱ_p
- Approximation of the pdf ϱ_p

TECHNISCHE UNIVERSITÄT DARMSTADT

Algorithm (SRD)

Let $b \sim \mathcal{N}(\mu, \Sigma)$ and L such that $LL^T = \Sigma$.

1. Sample Q points $\{w_1, \dots, w_Q\}$ uniformly distributed on the sphere $\mathbb{S}^{n-1}$.
2. Compute the one-dimensional sets $M_i := \{r \geq 0 | rLw_i + \mu \in M\}$ for $i = 1, \dots, Q$.
   $\longrightarrow$ Using analytic solution we get a representation of M_i as union of disjoint intervals $M_i = \bigcup_{j=1}^{t} [\underline{a}_j, \overline{a}_j]$
3. Set $\mathbb{P}(b \in M) \approx Q^{-1} \sum_{i=1}^{Q} \mu_\chi(M_i)$
   $\longrightarrow \mu_\chi(M_i) = \sum_{j=1}^{t} \mathcal{F}_\chi(\overline{a}_j) - \mathcal{F}_\chi(\underline{a}_j)$ with χ-distribution function $\mathcal{F}_\chi$

Stationary Gas Network

Kernel Density Estimator

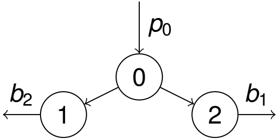
Algorithm (KDE)

1. Generate N independent and identically distributed (iid) samples $b^{S,i}$ of random outflow $b \sim \mathcal{N}(\mu, \Sigma)$
2. For every sample $b^{S,i}$ compute pressure $p_{out}(b^{S,i})$
3. Compute the kernel density estimator $\varrho_{p,N}$ for iid samples $p_{out}(b^{S,i})$ $\rightarrow$ Approximation of the exact pdf ϱ_p
4. Compute

$$\mathbb{P}(p_{out} \in P_{\text{min}}^{\text{max}}) \approx \int_{P_{\text{min}}^{\text{max}}} \varrho_{p,N}(z) dz = \frac{1}{N \prod_{j=1}^{n} h_j} \sum_{i=1}^{N} \prod_{j=1}^{n} \int_{p_j^{\text{min}}}^{p_j^{\text{max}}} \mathcal{K}\left(\frac{z_j - p_{out}(b^{S,i})}{h_j}\right) dz_j$$

Stationary Gas Network

Illustrative Example - Minimal Tree-Structured Graph



p_0	p^{min}	p^{max}	μ	Σ
60	$\begin{pmatrix} 40 \\ 30 \end{pmatrix}$	$\begin{pmatrix} 60 \\ 50 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 0.25 & 0 \\ 0 & 0.25 \end{pmatrix}$

10^5 Gaussian distributed samples for KDE, 10^4 uniform distributed samples of sphere $\mathbb{S}^1$ for SRD

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7
MC	75.07%	75.15%	75.22%	75.08%	74.88%	75.02%	74.99%
KDE	74.75%	74.84%	74.89%	74.78%	74.55%	74.76%	74.67%
SRD	74.95%						

Contamination flow in water network

For $(t, x) \in [0, T] \times [0, L]$ and constants $d < 0$, $m \leq 0$, we consider the deterministic scalar linear PDE with initial condition and boundary condition

$$\begin{aligned} \partial_t r(t, x) + d\partial_x r(t, x) &= mr(t, x), \\ r(0, x) &= r_0(x), \\ r(t, L) &= b(t). \end{aligned}$$

- Concentration r of the contamination
- Water gets polluted: $b(t) \geq 0$; water gets cleaned: $b(t) < 0$

Dynamic Flow Network

Time dependent random boundary data $b(t) \rightarrow$ representation as Fourier series

- boundary function $b_D \in L^2([0, T])$

$$b_D(t) = \sum_{m=0}^{\infty} a_m^0 \psi_m(t).$$

with

$$a_m^0 := \int_0^T b_D(t)\psi_m(t)dt, \quad \psi_m(t) := \frac{\sqrt{2}}{\sqrt{T}} \sin\left(\left(\frac{\pi}{2} + m\pi\right)\frac{t}{T}\right),$$

- For $a_m \sim \mathcal{N}(\mu, \sigma^2)$ consider random boundary data

$$b(t) = \sum_{m=0}^{\infty} a_m a_m^0 \psi_m(t) \qquad \in L^2([0, T]) \quad \mathbb{P}\text{-almost surely}$$

Dynamic Flow Network

Probabilistic Constraints in a Time-Dependent Setting

TECHNISCHE UNIVERSITÄT DARMSTADT

What is the probability that the contamination concentration r_1 at the inflow node v_1 meets the bounds $B := [r^{\min}, r^{\max}]$ during the whole time?

- With $M(t) := \{b \in L^2([0, T]; \mathbb{R}^n_{\leq 0}) | r_1(t, b) \in B\}$ we can write

$$\mathbb{P}\left(b \in M(t) \quad \forall t \in [0, T]\right)$$

- Set $\underline{t} := \operatorname{argmin}_{t \in [0,T]} r_1(t), \bar{t} := \operatorname{argmax}_{t \in [0,T]} r_1(t)$

$$\mathbb{P}(b \in M(\underline{t}) \text{ and } b \in M(\bar{t})) \qquad \longrightarrow \text{SRD?}$$

- Minimal and maximal value of $r_1(t)$ in $t \in [0, T]$ meet the bounds B

$$\mathbb{P}\left(\left(\min_{t \in [0,T]} r_1(t), \max_{t \in [0,T]} r_1(t)\right)^T \in B^2\right)$$

Dynamic Flow Network
KDE

Assumption

Random vector $R := (\min_{t\in[0,T]} r_1(t,b), \max_{t\in[0,T]} r_1(t,b))^T$ has an absolutely continuous distribution with pdf ϱ_R

Algorithm KDE

- Generate N iid samples $a_m^{S,i}$ of $a_m \sim \mathcal{N}(\mu,\Sigma)$ for $m = 1, \dots, N_F$
- Compute corresponding sample $b^{S,i}$ of truncated random boundary function b
- For every sample $b^{S,i}$ compute

$$\underline{r_1}^{S,i} := \min_{t\in[0,T]} r_1(t, b^{\mathcal{S},i}), \qquad \overline{r_1}^{S,i} := \max_{t\in[0,T]} r_1(t, b^{\mathcal{S},i}).$$

- Compute KDE $\varrho_{R,N}$ for 2-dim. samples $R^{S,i} = (\underline{r_1}^{S,i}, \overline{r_1}^{S,i})^T$ and approximate

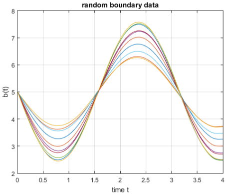
$$\mathbb{P}(R \in [r^{\min}, r^{\max}]^2) \approx \int_{[r^{\min}, r^{\max}]^2} \varrho_{R,N}(z)dz$$

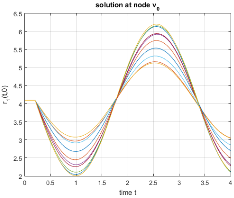
Dynamic Flow Network
Illustrative Example - One Pipe

TECHNISCHE UNIVERSITÄT DARMSTADT

$r^{\min}$	$r^{\max}$	μ	σ	d	m	L	T
2	6	1	0.25	-5	-1	1	4

Further: 101 time discretization points, 30 terms in the Fourier series, sampling of 10^5 boundary functions





	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7
MC	74.32%	74.39%	74.22%	74.32%	74.24%	74.33%	74.51%
KDE	74.33%	74.39%	74.21%	74.31%	74.24%	74.32%	74.51%

Conclusion

TECHNISCHE UNIVERSITÄT DARMSTADT

- Compare SRD and KDE on a minimal tree-structured graph
- Extend probabilistic constraints to a time dependent setting
- Illustrative example for KDE in dynamic setting

Outlook

- Exact solution unknown $\rightarrow$ Need numerical method for approximation $\rightarrow$ Reduce sampling cost for KDE using adaptive stochastic collocation method
- Analysis of necessary optimality conditions

Reference

- Devroye and Gyorfi
  Nonparametric density estimation: the L1 view, 1985
- Farshbaf-Shaker, Gugat, Heitsch and Henrion
  Optimal Neumann Boundary Contol of a Vibrating String with Uncertain Initial Data and Probabilistic Terminal Constraints, 2019
- Gugat and Schuster
  Stationary Gas Networks with Compressor Control and Random Loads: Optimization with Probabilistic Constraints, 2018
- Gramacki
  Nonparametric Kernel Density Estimation and Its Computational Aspects, 2018
- Van Ackooij and Henrion
  Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions, 2014