Newton-Based Methods for the Numerical Solution of Risk-Averse PDE-Constrained Optimization Problems

Mae Markowski, Advisor: Matthias Heinkenschloss

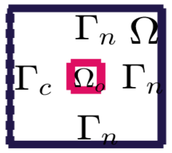
Department of Computational and Applied Mathematics, Rice University

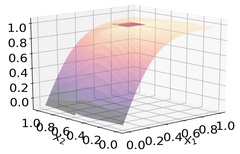
Introduction

Goal: Efficiently solve risk-averse PDE-constrained optimization problems arising in science and engineering; underlying physics uncertain
Solution: Simple but efficient modification to Newton-CG methods for solution of smoothed CVaR problems; reduces PDE solves by more than half

Deterministic Model Problem

Boundary control of Advection Diffusion Equation





$$\min_u \frac{1}{2}\int_{\Omega_o}(y(x;\cdot)-1)^2dx + \frac{10^{-2}}{2}\int_{\Gamma_c}u(x)^2dx$$
$$-\nabla\cdot(\kappa(x;\xi)\nabla y(x;\xi)) + c(x)\cdot\nabla y(x;\xi) = f(x;\xi),\quad x\in\Omega$$
$$10^{-4}(\kappa(x;\xi)\nabla y(x;\xi))\cdot n(x) + y(x;\xi) = u(x),\quad x\in\Gamma_c$$
$$\nabla y(x,\xi)\cdot n(x) = 0,\quad x\in\Gamma_n$$

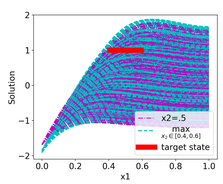
$\xi\in\Xi\subset\mathbb{R}^2$ fixed for deterministic problems.

Risk-Averse Optimization

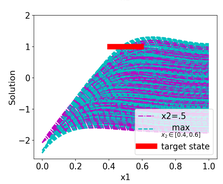
Uncertain diffusion $\kappa(x,\xi)$ and source $f(x,\xi)$ (ξ R.V.)
Cost term $s(y(u,\xi),\xi)$ is a random variable
Objective function has **risk measure** $\mathcal{R}$. Solve

$$\min_u \mathcal{R}\Big[s(y(u,\xi),\xi)\Big] + \alpha\|u\|^2_{\mathcal{U}}$$

where $s(y(u,\xi),\xi) = \max_{x\in\Omega_o}\left(s(y(u,\xi),\xi)-1\right)$

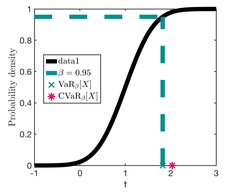


Risk netural. Minimize average loss. Many realizations result in high cost.



Risk averse, Conditional Value-at-Risk. Minimize conditional expectation of large losses that occur with small $(1-\beta)$ probability.

Conditional Value-at-Risk



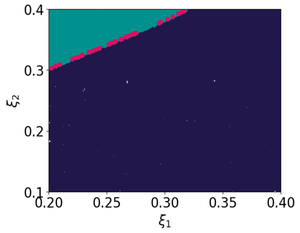
$$\mathrm{VaR}_\beta[X] = \inf_t\{t : \Pr(X>t)\le 1-\beta\}$$
$$\mathrm{CVaR}_\beta[X] = \mathbb{E}\left[X|X\ge \mathrm{VaR}_\beta[X]\right]$$
$$= \min_t\ t + \frac{1}{1-\beta}\mathbb{E}\left[(X-t)_+\right]$$

Optimization of CVaR

Challenges with CVaR minimization:

1. Only samples in risk regions contribute nonzero to CVaR

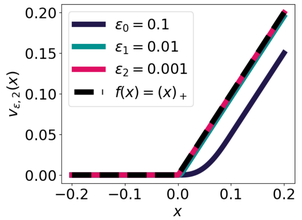
$$\{\xi\in\Xi : s(u,\xi)\ge \mathrm{VaR}_\beta[s]\}$$
$$\Rightarrow (s(u,\xi)-\mathrm{VaR}_\beta[s])_+\ne 0$$



2. CVaR is nonsmooth:

$$(x)_+ = \max\{x,0\}$$

Approximate by smoothing function v_ϵ

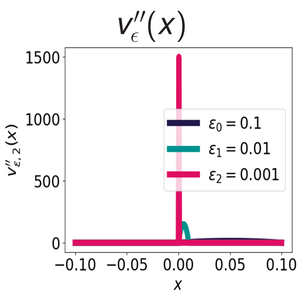


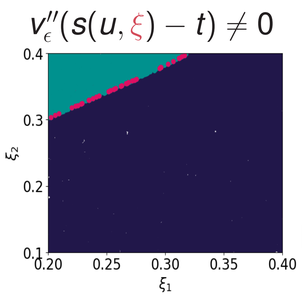
Rank Deficiency from Smoothing

Smoothing function v_ϵ allows use of Newton-CG to solve

$$\min_{u,t}\ J^\epsilon_\beta(u,t) := t + (1-\beta)^{-1}\mathbb{E}\left[v_\epsilon(s(u,\cdot)-t)\right] + \alpha\|u\|^2_{\mathcal{U}}$$

Second derivatives in quadratic model rely on $v''_\epsilon(s(u,\xi)-t)$





If no such ξ:

$$\begin{bmatrix}\nabla_{uu}J^\epsilon_\beta(u,t) & 0\\ 0 & 0\end{bmatrix}$$

Hess. Rank-deficient

Hessian near singular; quadratic model unbounded from below.
$\Rightarrow$ Poor steps, expensive line search/trust regon globalization.

Improvement of Newton-CG

At iteration k: add computationally inexpensive step

$$t_{k+1/2} = \operatorname*{argmin}_t\ J^\epsilon_\beta(u_k,t).\quad \text{(1D min, inexpensive)}$$

Minimize quadratic model around $(u_k, t_{k+1/2})$.

Advantages of improved Newton-CG:

1. Quadratic model not unbounded, since $\nabla_t J^\epsilon_\beta(u_k,t_{k+1/2}) = 0$.
2. Number of Newton-CG iterations substantially reduced.
3. Convergence theory extends to modified Newton-CG.

Numerical Results

Use structure of CVaR to further reduce PDE solves in gradient and Hessian computations; applied on Advection Diffusion model problem

Warm-started unmodified vs modified trust-region Newton-CG, $N = 500$ samples, gradient tol. = 10^{-8}, CVaR level $\beta = 0.9$.

ϵ	#N(#avg CG)		PDE solves	
1e-1	8(5)	3(9)	62,000	40,500
1e-2	21(2)	3(8)	54,822	12,328
1e-3	7(6)	2(13)	14,278	6,130
1e-4	8(3)	1(16)	12,428	3,438
1e-5	10(2)	1(15)	11,764	3,300
1e-6	0(-)	0(-)	0	0
Total	54(3)	10(10)	155,292	65,696

Acknowledgements

This research was supported by NSF Grants DMS-1522798, DMS-1819144 and Ken Kennedy Institute for Information Technology Shell Graduate Fellowship.

Contact: mae.markowski@rice.edu