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

\[
\begin{align*}
\min \frac{1}{2} \int_{\Omega} (u(x) - 1)^2 dx + \frac{10^{-2}}{2} \int_{\Omega} (u(x))^2 dx \\
-\nabla \cdot \left( \kappa(x, \xi) \nabla y(x; \xi) \right) + c(x) \cdot \nabla y(x; \xi) = f(x; \xi), \quad x \in \Omega \\
10^{-4} |\kappa(x, \xi) \nabla y(x; \xi) \cdot n(x) + y(x; \xi) - u(x)| = 0, \quad x \in \Gamma_e \\
\nabla y(x; \xi) \cdot n(x) = 0, \quad x \in \Gamma_n
\end{align*}
\]

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

Risk-Averse Optimization

Uncertain diffusion \( \kappa(x, \xi) \) and source \( f(x, \xi) \) \((\xi \text{ R.V.})\)

Cost term \( s(y(u, \xi), \xi) \) is a random variable

Objective function has risk measure \( \mathcal{R} \). Solve

\[
\begin{align*}
\min_{u, t} \mathcal{R} \left( s(y(u, \xi), \xi) \right) + \alpha \|u\|^2_{H^1}
\end{align*}
\]

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

Risk neutral. 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

\[
\begin{align*}
\text{VaR}_\beta[X] &= \inf \{ t : \Pr(X > t) \leq 1 - \beta \} \\
\text{CVaR}_\beta[X] &= \mathbb{E}[X | X \geq \text{VaR}_\beta[X]] \\
&= \min_{t} t + \frac{1}{1 - \beta} \mathbb{E}[(X - t)_+]
\end{align*}
\]

Optimization of CVaR

Challenges with CVaR minimization:

1. Only samples in risk regions contribute nonzero to CVaR
2. CVaR is nonsmooth:

\[
(\xi) = \max\{x, 0\}
\]

Approximate by smoothing function \( \nu \).

Smoothing function \( \nu \) allows use of Newton-CG to solve

\[
\min_{u, t} J_\beta(u, t) := t + (1 - \beta)^{-1} \mathbb{E}[\nu'(s(u, \cdot) - t)] + \alpha \|u\|^2_{H^1}
\]

Second derivatives in quadratic model rely on \( \nu''(s(u, \xi) - t) \)

\[
\nu''(s(u, \xi) - t) \neq 0
\]

If no such \( \xi \):

\[
\begin{bmatrix}
\nabla \nu \nu'(s(u, \cdot) - t) & 0 \\
0 & 0
\end{bmatrix}
\]

Hess. Rank-deficient

Hessian near singular; quadratic model unbounded from below.

\( \Rightarrow \) Poor steps, expensive line search/trust region globalization.

Improvement of Newton-CG

At iteration \( k \): add computationally inexpensive step

\[
t_{k+1/2} = \arg\min_{t} J_\beta(u_k, t). \quad (1D \text{ min, inexpensive})
\]

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

Advantages of improved Newton-CG:

1. Quadratic model not unbounded, since \( \nabla v J_\beta(u_k, t_{k+1/2}) = 0 \).
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 \( \beta = 0.9 \).

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>#N/#avg CG</th>
<th>PDE solves</th>
</tr>
</thead>
<tbody>
<tr>
<td>1e-1</td>
<td>8(5) \quad 3(9)</td>
<td>62,000 \quad 40,500</td>
</tr>
<tr>
<td>1e-2</td>
<td>21(2) \quad 3(8)</td>
<td>54,822 \quad 12,328</td>
</tr>
<tr>
<td>1e-3</td>
<td>7(6) \quad 2(13)</td>
<td>14,278 \quad 6,130</td>
</tr>
<tr>
<td>1e-4</td>
<td>8(3) \quad 1(16)</td>
<td>12,428 \quad 3,438</td>
</tr>
<tr>
<td>1e-5</td>
<td>10(2) \quad 1(15)</td>
<td>11,764 \quad 3,300</td>
</tr>
<tr>
<td>1e-6</td>
<td>0(-) \quad 0(-)</td>
<td>0 \quad 0</td>
</tr>
<tr>
<td>Total</td>
<td>54(3) \quad 10(10)</td>
<td>155,292 \quad 65,696</td>
</tr>
</tbody>
</table>

Acknowledgements

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