Summer workshop IGDK

Proposal: Semismooth Newton for control constrained optimization in H^1

Konstantin Pieper
Moritz Keuthen

May 14, 2013

1 Motivation

Consider an optimal control problem of the following structure:

$$\begin{aligned} \text{Minimize} \quad & j(u) + \tfrac{\alpha}{2}\|u\|_X^2 \\ \text{s. t.} \quad & u \le u_b, \end{aligned} \tag{1}$$

where $j\colon X \to \mathbb{R}$ is a smooth objective functional depending on the variable $u \in X$ where is $X \hookrightarrow L^2(\Omega)$ an (infinite dimensional) Hilbert space. The optimality conditions are

$$(Dj(u) + \alpha u,\, \tilde{u} - u)_X \ge 0 \quad \text{for all } \tilde{u} \in X \cap U_{ad}.$$

with the admissible set given by $U_{ad} := \{u \in X \mid u \le u_b\}$ and the X-gradient of j

$$(Dj(u),\, \delta u)_X = j'(u)(\delta u).$$

We can rewrite this as

$$u = P_{ad}(-\tfrac{1}{\alpha} Dj(u)),$$

where P_{ad} is the X-projection onto U_{ad},

$$u = \operatorname*{argmin}_{X \cap U_{ad}} \|u + \tfrac{1}{\alpha} Dj(u)\|_X^2.$$

We want to consider semismooth Newton methods for (1), which can be analyzed in a Banach space setting, i.e. without discretizing X. Such methods are known to behave in a mesh independent way after discretization. By this, we mean that we want to apply Newtons method to the equation

$$G(u) = u - P_{ad}(-\tfrac{1}{\alpha} Dj(u)) = 0.$$

For $X = L^2(\Omega)$ this is well understood (see [1, 2, 3]). In this case P_{ad} is the L^2-projection onto U_{ad}, which can be written as the superposition operator $P_{ad}(\cdot) = \min\{\cdot\,, u_b\}$. Since this projection operates pointwise in space, it can be efficiently computed, and we can compute a linearization of $G(\cdot)$ using the concept of a Newton-derivative for superposition operators. It is well known, that a Newton method based on this can be analyzed in a Banach space setting under a smoothing assumption on the operator $v \mapsto Dj(v)$. We obtain superlinear convergence of the method if the Newton-derivatives of $G(\cdot)$ are uniformly invertible, see [1, 2, 3].

Consider now the case $X = H_0^1(\Omega)$ with $\|u\|_X^2 = \|\nabla u\|^2$ as another example. In this case the gradient fulfills the condition

$$(\nabla Dj(u), \nabla \delta u) = j'(u)(\delta u)$$

and the projection maps any $w \in H_0^1$ to $v = P_{ad}(w) \in H_0^1 \cap U_{ad}$, defined as the solution of

$$(\nabla v, \nabla(\tilde{v} - v)) \geq f(\tilde{v} - v) \quad \text{for all } \tilde{v} \in H_0^1 \cap U_{ad}. \tag{2}$$

where $f(\delta v) = (\nabla w, \nabla \delta v)$, which is an obstacle problem for v. In fact we have

$$u = P_{ad}(-\tfrac{1}{\alpha} Dj(u)) = S_o(-\tfrac{1}{\alpha} j'(u)) \tag{3}$$

where $S_o \colon H^{-1} \to H_0^1, f \mapsto v$ is the solution operator of the obstacle problem. Even though a closed formula for P_{ad} does not exist, contrary to the L^2-case, it is known, that this solution operator can be realized for a suitable discretization in (near) linear time.

To the best of the authors' knowledge, a Newton method based on this equation has not been derived before. Therefore we would need a representation formula for the sensitivities of $S_o(f)$ with respect to f, and an efficient way to compute this in practice. Also we need to find a suitable concept and conditions for differentability.

2 Outline

The questions we will try to address in this project are:

- What is a suitable concept for a Newton-derivative of P_{ad} in the H_0^1-context? For starters, we have to check [5, 6, 7] for known continuity and differentability results.
- Does the derivative have a realization suitable for (efficient) practical computations?
- Can a Newton method based on (3) be analyzed in a Banach space setting? Do we observe mesh-independence in practice?
- What is the appropriate smoothing condition on $Dj(\cdot)$ in this context?

The practical viability of the method necessitates an efficient solution method for (2). Ideally it should have linear complexity in the number of variables after discretization. So additional questions are:

- What is known about efficient solution strategies for (2)? Possible references are [3, 4], i.e., PDAS in combination with a grid hierarchy or continuation strategy. What about monotonous multigrid, i.e., multigrid for the variational inequality (2)?
- How does a practical realization of the new method compare to existing strategies?
  - continuation methods ...
  - SSN based on the L^2-projection (after discretization, since no Banach-space analysis is possible)

In connection to the last questions, we will implement the algorithm for a model optimal control problem, after discretizing X with finite elements.

References

[1] M. Ulbrich, *Semismooth Newton Methods for Variational Inequalities and Constrained Optimization Problems in Function Spaces*, SIAM, (2011).

[2] A. Schiela, *A Simplified Approach to Semismooth Newton Methods in Function Space*, SIAM J. Opt., (2008).

[3] M. Hintermüller, K. Ito, K. Kunisch, *The Primal-Dual Active Set Strategy as a Semismooth Newton Method*, SIAM J. Opt., (2002).

[4] G. Stadler, *Infinite-Dimensional Semi-Smooth Newton and Augmented Lagrangian Methods for Friction and Contact Problems in Elasticity*, (2004). PhD thesis, University of Graz.

[5] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan Vol. 29, No. 4, (1977).

[6] K. Kunisch, D. Wachsmuth, *Sufficient Optimality Conditions and Semi-Smooth Newton Methods for Optimal Control of Stationary Variational Inequalities*, ESIAM, Control Optim. Calc. Var. 18(2) (2012).

[7] F. Mignot, *Contrôle dans les Inéquations Variationelles Elliptiques*, J. Func. Anal. 22, (1975).