MEMORY-EFFICIENT IMPLEMENTATION OF STABLE NONSMOOTH NEWTON’S METHOD: APPLICATION TO CONTROL-STATE CONSTRAINED OPTIMAL CONTROL PROBLEMS

Julia Sternberg, Matthias Gerdts

Department of Mathematics, University of Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany, sternberg@math.uni-hamburg.de, gerdts@math.uni-hamburg.de

Keywords: optimal control, nonsmooth Newton’s method, control-state constraints, stable multiple shooting method, checkpointing

We consider the following optimal control problem (OCP) subject to mixed control-state constraints:

Minimize \( \int_0^1 f_0(x(t), u(t)) dt \)

s.t. \( x'(t) = f(x(t), u(t)) \) a.e. in \([0,1]\),
\( \psi(x(0), x(1)) = 0, \)
\( c(x(t), u(t)) \leq 0 \) a.e. in \([0,1]\),

where \( x \in W^{1,\infty}([0,1], \mathbb{R}^n_x), u \in L^\infty([0,1], \mathbb{R}^n_u) \). Without loss of generality we consider only autonomous problems on the fixed time interval \([0,1]\). The functions \( f_0: \mathbb{R}^n_x \times \mathbb{R}^n_u \to \mathbb{R} \), \( f: \mathbb{R}^n_x \times \mathbb{R}^n_u \to \mathbb{R}^n_x \),
\( \psi: \mathbb{R}^n_x \times \mathbb{R}^n_x \to \mathbb{R}^n_\psi \), \( c: \mathbb{R}^n_x \times \mathbb{R}^n_u \to \mathbb{R}^n_c \),
are supposed to be at least twice continuously differentiable w.r.t. to all arguments.

Several approaches towards the numerical solution of OCP have been investigated in the literature. The so-called direct discretization method is based on a discretization of the infinite dimensional optimal control problem and leads to a finite dimensional nonlinear program. The latter can be solved numerically by suitable programming methods such as, e.g., sequential quadratic programming. The direct discretization method turns out to be very robust in practice. Nevertheless, the computational effort grows at a nonlinear rate with the number of grid points used for discretization.

The so-called indirect method for optimal control problems attempts to satisfy the necessary conditions that are provided by the well-known minimum principle numerically. The exploitation of the minimum principle leads to a nonlinear multi-point boundary value problem that has to be solved. Although the indirect method usually leads to the most accurate solutions, it suffers from the drawback that it requires a good initial guess in order to convergence. One crucial task is to estimate the sequence of active and inactive intervals of the control-state constraint.

In our talk we consider the indirect approach, which avoids the latter drawback, and apply the nonsmooth Newton’s method for its realization. The method is based on a nonsmooth reformulation of the necessary optimality conditions. A brief outline of the essential ideas of the algorithm is as follows. The necessary conditions are stated in terms of a local minimum principle. By use of the Fischer-Burmeister function the local minimum principle is transformed into an equivalent nonlinear and nonsmooth equation in appropriate Banach spaces:

\[ F(z) = 0, \quad F: Z \to Y, \]

where \( Z \) and \( Y \) are appropriate Banach spaces. Application of the globalized nonsmooth Newton’s method generates sequences \( \{z^k\}, \{d^k\} \) and \( \{\alpha_k\} \) related by the iteration

\[ z^{k+1} = z^k + \alpha_k d^k, \quad k = 0, 1, 2, \ldots. \]

Herein, the search direction \( d^k \) is the solution of the linear operator equation \( V_k(d^k) = -F(z^k) \) and the step length \( \alpha_k > 0 \) is determined by a line-search procedure of Armijo’s type for a suitably defined merit function. The linear operator \( V_k \) is chosen from an appropriately defined generalized Jacobian \( \partial_z F(z^k) \) (for details see (1)).
In our talk we describe how the search direction $d^k$ can be computed by the multiple shooting approach. We particularly consider the dichotomy case, so that the linear operator $V_k$ contains both fast growing and fast decaying modes. Utilizing this fact, we can successfully solve the linear system $V_k(d^k) = -F(z^k)$ combining compactification and decoupling, where the decoupling corresponds to the splitting of growing and decaying modes, which results in a stable version of the nonsmooth Newton’s method.

Each iteration of the Newton’s method contains three alternative sweeps through a time horizon, so that the information evaluated within each sweep is required to integrate the others two. To reduce the huge memory requirement, resulting by the straightforwardly storing all information evaluated during each sweep, we apply checkpointing techniques. As developed in (2; 3), checkpointing means that not all intermediate states are saved but only a small subset of them is stored as checkpoints. Because of the triple sweep within each Newton iteration, we are faced here with a nested checkpointing, where checkpoints from various sweeps must be kept simultaneously. In our talk we describe some heuristics to construct appropriate nested reversal schedules.

Finally we present some numerical examples.

REFERENCES

