NUMERICAL SOLUTION OF OPTIMAL CONTROL PROBLEMS WITH CONVEX CONTROL CONSTRAINTS

[1]D. Wachsmuth
[1]*Institut für Mathematatik, TU Berlin, Str. des 17. Juni 136, D-1062 Berlin, Germany,*
wachsmut@math.tu-berlin.de

Abstract We study optimal control problems with vector-valued controls. In the article, we propose a solution strategy to solve optimal control problems with pointwise convex control constraints. It involves a SQP-like step with an imbedded active-set algorithm.

keywords: Optimal control, convex control constraints, set-valued mappings.

1. Introduction

In this article, we want to investigate solution strategies to solve optimal control problems with pointwise convex control constraints. In flow control, any control regardless of distributed or boundary control is a vector-valued function. That means, there are many possibilities to formulate control constraints. Here, we will study a general concept of such control constraints. As an example of an optimal control problem with vector-valued controls we chose the following problem of optimal distributed control of the instationary Navier-Stokes equations in two dimensions. To be more specific, we want to minimize the following quadratic objective functional:

$$J(y,u) = \frac{\alpha_T}{2}\int_\Omega |y-y_T|^2 dx + \frac{\alpha_Q}{2}\int_Q |y-y_Q|^2 dxdt + \frac{\gamma}{2}\int_Q |u|^2 dxdt \quad (1)$$

subject to the instationary Navier-Stokes equations

$$\begin{aligned} y_t - \nu\Delta y + (y\cdot\nabla)y + \nabla p &= u && \text{in } Q,\\ \operatorname{div} y &= 0 && \text{in } Q,\\ y(0) &= y_0 && \text{in } \Omega, \end{aligned} \quad (2)$$

and to the control constraints $u \in U_{ad}$ with U_{ad} defined by

$$U_{ad} = \{u \in L^2(Q)^2 : \; u(x,t) \in U(x,t) \text{ a.e. on } Q\}. \quad (3)$$

Please use the following format when citing this chapter:

Wachsmuth, D., 2006, in IFIP International Federation for Information Processing, Volume 202, Systems, Control, Modeling and Optimization, eds. Ceragioli, F., Dontchev, A., Furuta, H., Marti, K., Pandolfi, L., (Boston: Springer), pp. 319-327.

Here, Ω is a bounded domain in R^2, Q denotes the time-space cylinder $Q := \Omega \times (0, T)$. The set of admissible controls is generated by the set-valued mapping U, $U : Q \mapsto 2^{R^2}$. The conditions imposed on the various ingredients of the optimal control problem are specified in Sections 2.1 and 2.2, see assumptions (A) and (AU).

Several choices of the control constraint U are discussed in [13]. Let us only mention that the case of box-constraints is a special case of our general convex constraint. Optimal control problems with such control constraints are rarely investigated in literature. We refer to Bonnans [2], Dunn [3], Páles and Zeidan [7], and the authors article [14].

The plan of the article is as follows. In Section 2 we collect results concerning the solvability of the state equation and the optimal control problem. Necessary optimality conditions are stated in Section 3. Section 4 is devoted to the derivation of a solution strategy. There, a new active-set algorithm is described. Numerical experiments that confirm the efficiency of the proposed algorithm are presented in Section 5.

2. Statement of the optimal control problem

We define the spaces of solenoidal functions $H := \{v \in L^2(\Omega)^2 : \operatorname{div} v = 0\}$, $V := \{v \in H_0^1(\Omega)^2 : \operatorname{div} v = 0\}$. Further, we will work with the standard spaces of abstract functions from $[0, T]$ to a real Banach space X, $L^p(0, T; X)$, endowed with their natural norms.

2.1 The state equation

Before we start with the discussion of the state equation, we specify the requirements for the various ingredients describing the optimal control problem. In the sequel, we assume that the following conditions are satisfied:

$$
(\mathrm{A}) \begin{cases} 1 \;\; \Omega \subset R^2 \text{ is domain with Lipschitz boundary } \Gamma := \partial\Omega, \\ 2 \;\; y_0,\, y_T \in H,\, y_Q \in L^2(Q)^2, \\ 3 \;\; \alpha_T,\, \alpha_Q \geq 0,\, \gamma,\, \nu > 0. \end{cases}
$$

We introduce a linear operator $A : L^2(0, T; V) \mapsto L^2(0, T; V')$ by

$$
\int_0^T \langle (Ay)(t),\, v(t) \rangle_{V',V} dt := \int_0^T (y(t), v(t))_V dt,
$$

and a nonlinear operator B by

$$
\int_0^T \langle (B(y))(t),\, v(t) \rangle_{V',V}\, dt := \int_0^T \int_\Omega \sum_{i,j=1}^2 y_i(t) \frac{\partial y_j(t)}{\partial x_i}\, v_j(t) dx dt.
$$

Now, we define the notation of weak solutions for the instationary Navier-Stokes equations (2) in the Hilbert space setting. A function $y \in L^2(0,T;V)$ with $y_t \in L^2(0,T;V')$ is called weak solution of (2) if it satisfies

$$\begin{aligned} y_t + \nu A y + B(y) &= f \text{ in } L^2(0,T;V'), \\ y(0) &= y_0. \end{aligned} \tag{4}$$

Results concerning the solvability of (4) are standard, cf. [8] for proofs and further details.

2.2 Set-valued functions

Before we begin with the formulation of the optimal control problem with inclusion constraints, we will specify the notation and assumptions for the admissible set $U(\cdot)$. It is itself a mapping from the control domain Q to the set of subsets of R^2, a so-called *set-valued mapping.*

The controls are taken from the space $L^2(Q)^2$, so we have to impose at least some measurability conditions on the mapping U. In the sequel, we will work with measurable set-valued mappings, we refer to the textbook by Aubin and Frankowska [1]. Once and for all, we specify the requirements for the function U, which defines the control constraints.

(AU)
The set-valued function $U : Q \to R^2$ satisfies:

1. *U is a measurable set-valued function.*
2. *The images of U are non-empty, closed, and convex a.e. on Q. That is, the sets $U(x,t)$ are non-empty, closed and convex for almost all $(x,t) \in Q$.*
3. *There exists a function $f_U \in L^2(Q)^2$ with $f_U(x,t) \in U(x,t)$ a.e. on Q.*

The assumption (AU) is as general as possible. In the case that the set-valued function U is a constant function, i.e. $U(x,t) \equiv U_0$, we can give a simpler characterization: it suffices that U_0 is non-empty, closed, and convex. For further details, we refer to [14].

3. First-order necessary conditions

We will repeat the exact statement of the necessary optimality conditions for convenience of the reader.

THEOREM 1 (NECESSARY CONDITION) *Let $\bar{u}$ be locally optimal in $L^2(Q)^2$ with associated state $\bar{y} = y(\bar{u})$. Then there exists a unique Lagrange multiplier $\bar{\lambda} \in L^2(0,T;V)$ with $\bar{\lambda}_t \in L^{4/3}(0,T;V)$, which is the weak solution of the* *adjoint equation*

$$\begin{aligned} -\bar\lambda_t + \nu A\bar\lambda + B'(\bar y)^*\bar\lambda &= \alpha_Q(\bar y - y_Q) \\ \bar\lambda(T) &= \alpha_T(\bar y(T) - y_T). \end{aligned} \tag{5}$$

Moreover, it holds

$$(\gamma\bar u + \bar\lambda, u - \bar u)_Q \geq 0 \quad \forall u \in U_{ad}. \tag{6}$$

We can reformulate the variational inequality (6) as the projection representation of the optimal control

$$\bar u(x,t) = \mathrm{Proj}_{U(x,t)}\left(-\frac{1}{\gamma}\bar\lambda(x,t)\right) \qquad \text{a.e. on } Q. \tag{7}$$

Secondly, another formulation of the variational inequality uses the normal cone $N_{U_{ad}}(\bar u)$. Then inequality (6) can be written equivalently as

$$\nu\bar u + \bar\lambda + N_{U_{ad}}(\bar u) \ni 0. \tag{8}$$

It will allow us to write the optimality system as a generalized equation.

4. Solution strategy

Now, we are going to describe our strategy to solve optimal control problems with pointwise convex control constraints. The starting point of our considerations is Newtons method applied to generalized equations.

4.1 Generalized Newton's method

In the sequel, we want to apply Newton's method to the optimality system. This system can be written as a generalized equation. Let us define a function F by

$$F(y,u,\lambda) = \begin{pmatrix} y_t + \nu Ay + B(y) \\ y(0) \\ -\lambda_t + \nu A\lambda + B'(y)^*\lambda \\ \lambda(T) \\ \gamma u + \lambda \end{pmatrix} - \begin{pmatrix} u \\ y_0 \\ \alpha_Q(y - y_Q) \\ \alpha_T(y(T) - y_T) \\ 0 \end{pmatrix}. \tag{9}$$

Then a triple $(\bar y, \bar u, \bar\lambda)$ fulfills the necessary optimality conditions consisting of the state equation (4), the adjoint equation (5), and the variational inequality (6) if and only if it fulfills the generalized equation

$$F(\bar y, \bar u, \bar\lambda) + (0,0,0,0,N_{U_{ad}}(\bar u))^T \ni 0. \tag{10}$$

Now, we will apply the generalized Newton method to equation (10). If the control constraints are box constraints then this method is equivalent to the SQP-method, see e.g. [9]. Given iterates (y_n, u_n, λ_n) we have to solve the linearized generalized equation: 0 should belong to the set

$$F(y_n, u_n, \lambda_n) + F'(y_n, u_n, \lambda_n)[(y - y_n, u - u_n, \lambda - \lambda_n)] + (0, 0, 0, 0, N_{U_{ad}}(\bar{u}))^T \tag{11}$$

in every step. It turns out that this equation is the optimality system of the linear-quadratic optimal control problem.

$$\min J_n(y, u) = \frac{\alpha_T}{2}|y(T) - y_d|_H^2 + \frac{\alpha_Q}{2}\|y - y_Q\|_2^2 - B(y - y_n)\lambda_n \tag{12}$$

subject to the linearized state equation

$$\begin{aligned} y_t + \nu A y + B'(y_n)y &= u + B(y_n) \\ y(0) &= y_0 \end{aligned} \tag{13}$$

and the control constraint

$$u \in U_{ad}. \tag{14}$$

Let us emphasize the following observation. In the generalized equation (10) only $N(x)$ represents the control constraint. And it is the only term that was *not* linearized in (11). Consequently, the subproblem (12)–(14) is subject to the same control constraint $u(x, t) \in U(x, t)$ as the original non-linear problem. That means, the control constraint is not linearized, even if it is written as an inequality like $u_1(x, t)^2 + u_2(x, t)^2 \leq \rho(x, t)^2$. This is a difference to the standard SQP-method for optimization problems with nonlinear inequalities. An inequality $g(x) \leq 0$ would be linearized to $g(x_n) + g'(x_n)(x - x_n) \leq 0$.

4.2 Active-set strategy

As for the box-constrained case we will use an active set algorithm. It is very similar to the well-known primal-dual active-set strategy. The algorithm we propose here tries to solve the projection representation of optimal controls given by (7).

The algorithm to solve the subproblem (12)–(14) works as follows. We denote the control iterates of step k by u^k. The state y^k is the solution of (13) with right-hand side u^k, and λ^k is the solution of the adjoint equation associated to (13).

Algorithm. Take a starting guess u^0 with associated state y^0 and adjoint λ^0. Set $k = 0$.

1 Given u^k, y^k, λ^k. Determine the active set $A^{k+1} = A^{k+1}(\lambda^k)$ by

$$A^{k+1} := \left\{(x, t) : -\frac{1}{\gamma}\lambda^k(x, t) \notin U(x, t)\right\}.$$

2 Minimize the functional J_n given by (12) subject to the linearized state equation (13) and to the control constraints

$$\left.\begin{array}{lcl} u|_{A^{k+1}} & = & \mathrm{Proj}_{U(x,t)}\left(-\frac{1}{\gamma}\lambda^k(x,t)\right) \\ u|_{Q\setminus A^{k+1}} & & \text{free.} \end{array}\right\}$$

Denote the solution by $(\tilde{u},\, \tilde{y},\, \tilde{\lambda})$.

3 Project $u^{k+1} := \mathrm{Proj}_{U(x,t)}\left(-\frac{1}{\gamma}\tilde{\lambda}(x,t)\right)$, compute y^{k+1} and λ^{k+1}.

4 If $\left\|u^{k+1} - \mathrm{Proj}_{U_{ad}}\left(-\frac{1}{\gamma}\lambda^{k+1}\right)\right\|_2 < \epsilon \to$ ready,
else set $k := k+1$ and go back to 1.

In the first step of the algorithm, the active set is determined. The control constraint at a particular point (x,t) is considered active if $-1/\gamma\,\lambda^k(x,t)$ does not belong to $U(x,t)$. In the second step, a linear-quadratic optimization problem is solved. It involves no inequality constraints, since the control is fixed on the active set and the control is free on the inactive set. After that, the optimal adjoint state of that problem is projected on the admissible set to get the new control. The algorithm stops if the residual in the projection representation is small enough.

Let us compare the behaviour of our algorithm and the primal-dual method in the detection of the active sets for the simple box constraint $|u_i| \le 1$, $i = 1,2$. Let us assume that $-1/\gamma\,\lambda^k(x_0,t_0) = (2,0)$, i.e. $-1/\gamma\,\lambda^k(x_0,t_0)$ is not admissible. Then in our method the point (x_0,t_0) will belong to the active set A^{k+1} in the next step. And at this point the value $(1,0)$ is prescribed for u^{k+1}: $u^{k+1}(x_0,t_0) = (1,0)$. In the primal-dual method, the point (x_0,t_0) will be added to the active set A_1^{k+1} associated to the inequality $|u_1| \le 1$. It will not belong to the active set A_2^{k+1} for the second inequality $|u_2| \le 1$ since this inequality was not violated. And for the inner problem we will get the constraints $u_1^{k+1}(x_0,t_0) = 1$ and $u_2^{k+1}(x_0,t_0)$ =free, that is the control is allowed to vary in tangential directions on the right side of the box!

5. Numerical results

Now, let us report about the performance of the proposed algorithm. To study the convergence speed We present results for an optimal control problem, where the solution is known.

5.1 Problem setting

The computational domain was chosen to be the unit square $\Omega = (0,1)^2$ with final time $T = 1$. We want to minimize the functional

$$\frac{1}{2}\int_0^1\int_\Omega |y(x,t) - y_d(x,t)|^2 dxdt + \frac{\gamma}{2}\int_0^1\int_\Omega |u(x,t)|^2 dxdt$$

subject to the instationary Navier-Stokes equations on $\Omega\times(0,1)$ with distributed control

$$\begin{aligned} y_t - \nu\Delta y + (y\cdot\nabla)y + \nabla p &= u + f && \text{in } Q,\\ \operatorname{div} y &= 0 && \text{in } Q,\\ y(0) &= y_0 && \text{in } \Omega. \end{aligned}$$

and subject to some control constraints to be specified later on. Let us construct a triple of state, control and adjoint, that satisfies the first-order optimality system. We chose as state and adjoint state

$$\bar{y}(x,t) = e^{-\nu t}\begin{pmatrix} \sin^2(\pi x_1)\sin(\pi x_2)\cos(\pi x_2) \\ -\sin^2(\pi x_2)\sin(\pi x_1)\cos(\pi x_1) \end{pmatrix},$$

$$\bar{\lambda}(x,t) = \left(e^{-\nu t} - e^{-\nu}\right)\begin{pmatrix} \sin^2(\pi x_1)\sin(\pi x_2)\cos(\pi x_2) \\ -\sin^2(\pi x_2)\sin(\pi x_1)\cos(\pi x_1) \end{pmatrix}.$$

Regardless of the choice of U_{ad}, the control is computed using the projection formula as

$$\bar{u} = \mathrm{Proj}_{U_{ad}}\left(-\frac{1}{\gamma}\bar{\lambda}(x,t)\right).$$

The quantities f, y_0, y_d are now chosen in such a way that $\bar{y}$ and $\bar{\lambda}$ are the solutions of the state and adjoint equations, respectively.

5.2 Discretization

The continuous problem was discretized using Taylor-Hood finite elements with different mesh sizes. The grid consists of 4096 triangles with 8321 velocity and 2113 pressure nodes yielding a mesh size $h = 0.03125$. Further, we use the semi-implicit Euler scheme for time integration with a equidistant time discretization with step length $\tau = 0.000625$. The control is approximated by piecewise continuous functions in time and space.

We used the SQP-method without any globalization to solve the problem with box constraints. The constrained SQP-subproblems were solved by the primal-dual active-set method. This method is known to converge locally with super-linear convergence rate [4, 11] if a sufficient optimality condition holds. Under some strong assumptions it converges even globally [6].

To solve the optimal control problem with convex constraints, we employ the solution strategy proposed in the previous section: generalized Newton method with active-set strategy as inner loop. The quadratic subproblems of the active-set method were again solved by the CG algorithm.

5.3 Results

Now, let us report about the results for convex control constraints compared to the box-constraints considered above. The parameters of the example are set to $\gamma = 0.01$ and $\nu = 0.1$. We computed solutions for the box constraint

$$|u_i(x,t)| \leq 1.0$$

and for the convex constraint in polar coordinates

$$0 \leq u_r(x,t) \leq \psi(u_\phi(x,t)).$$

The function $\psi(\phi)$ is given as the spline interpolation with cubic splines and periodic boundary conditons of the function values in Table 1. It is chosen such that the admissible set is comparable to the box constrained case. The projection on that set was computed by Newtons method.

ϕ	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3}{4}\pi$	π	$\frac{5}{4}\pi$	$\frac{3}{2}\pi$	$\frac{7}{4}\pi$
$\psi(\phi)$	1.0	1.0	0.64	0.44	0.4	0.44	0.64	1.0

Table 1. Convex constraint

Let us compare briefly the convergence of the active set algorithms: the primal-dual active set method for the box constrained problem and the active set method proposed above to solve the convex constrained problem. In all our computations both methods showed a similar behaviour, which can be seen in Table 2. Although they solved optimal control problems with different control constraints, they needed almost the same number of outer (SQP/generalized Newton) and inner (active-set) iterations. Also the residual $u - \text{Proj}(-1/\gamma\, \lambda)$ depicted by 'Res' decreased in the same way. So, we can say that our algorithm is as efficient as the primal dual active set method applied to box constraints.

References

[1] J.-P. Aubin and H. Frankowska. *Set-valued analysis.* Birkhäuser, 1990.

[2] J. F. Bonnans. Second-order analysis for constrained optimal control problems of semilinear elliptic equations. *Appl. Math. Optim.*, 38:303–325, 1998.

[3] J. C. Dunn. Second-order optimality conditions in sets of L^∞ functions with range in a polyhedron. *SIAM J. Control Optim.*, 33(5):1603–1635, 1995.

[4] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13:865–888, 2003.

Box constraints			Convex constraints		
SQP		PD-AS	gNewton		AS
It	$\|u^n - u^{n-1}\|_\infty$	Res	It	$\|u^n - u^{n-1}\|_\infty$	Res
1		0.1564	1		$0.9912 \cdot 10^{-1}$
		$0.4823 \cdot 10^{-1}$			$0.2327 \cdot 10^{-1}$
		$0.1128 \cdot 10^{-1}$			$0.3475 \cdot 10^{-2}$
2		$0.1128 \cdot 10^{-1}$	2		$0.3498 \cdot 10^{-2}$
	0.1221	$0.4971 \cdot 10^{-3}$		0.1311	$0.1976 \cdot 10^{-3}$
3		$0.6967 \cdot 10^{-3}$	3		$0.2530 \cdot 10^{-3}$
	$0.1148 \cdot 10^{-1}$	$0.1561 \cdot 10^{-5}$			$0.3741 \cdot 10^{-4}$
				$0.1186 \cdot 10^{-1}$	$0.7047 \cdot 10^{-5}$
4		$0.1891 \cdot 10^{-5}$	4		$0.7035 \cdot 10^{-5}$
	$0.4105 \cdot 10^{-6}$	$0.4643 \cdot 10^{-6}$			$0.1291 \cdot 10^{-5}$
				$0.2804 \cdot 10^{-3}$	$0.2308 \cdot 10^{-6}$

Table 2. Comparison of the algorithms

[5] M. Hinze. *Optimal and instantaneous control of the instationary Navier-Stokes equations.* Habilitation, TU Berlin, 2002.

[6] K. Kunisch and A. Rösch. Primal-dual active set strategy for a general class of constrained optimal control problems. *SIAM J. Optim.*, 13:321–334, 2002.

[7] Zs. Páles and V. Zeidan. Optimum problems with measurable set-valued constraints. *SIAM J. Optim.*, 11:426–443, 2000.

[8] R. Temam. *Navier-Stokes equations.* North Holland, Amsterdam, 1979.

[9] F. Tröltzsch. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations. *SIAM J. Control Optim.*, 38:294–312, 1999.

[10] F. Tröltzsch and D. Wachsmuth. Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations. *ESAIM: COCV*, 12:93–119, 2006.

[11] M. Ulbrich. Constrained optimal control of Navier-Stokes flow by semismooth Newton methods. *Systems & Control Letters*, 48:297–311, 2003.

[12] D. Wachsmuth. Regularity and stability of optimal controls of instationary Navier-Stokes equations. *Control and Cybernetics*, 34:387–410, 2005.

[13] D. Wachsmuth. Optimal control problems with convex control constraints. Preprint 35-2005, Institut für Mathematik, TU Berlin, submitted, 2005.

[14] D. Wachsmuth. Sufficient second-order optimality conditions for convex control constraints. *J. Math. Anal. App.*, 2006. To appear.