A sharp regularization error estimate for bang-bang solutions for an iterative Bregman regularization method for optimal control problems

Frank Pörner

In the present work, we present numerical results for an iterative method for solving an optimal control problem with inequality constraints. The method is based on generalized Bregman distances. Under a combination of a source condition and a regularity condition on the active sets convergence results are presented. Furthermore we show by numerical examples that the provided a-priori estimate is sharp in the bang-bang case.

1 Introduction

In this article we consider optimization problems of the following form

\[
\text{Minimize } \frac{1}{2} \| Su - z \|_Y^2 \quad \text{such that } \quad u_a \leq u \leq u_b \quad \text{a.e. in } \Omega \quad (P)
\]

which can be interpreted both as an optimal control problem or as an inverse problem. Here \( \Omega \subseteq \mathbb{R}^n \), \( n \geq 1 \) is a bounded, measurable set, \( Y \) a Hilbert space, \( z \in Y \) a given function. The operator \( S : L^2(\Omega) \to Y \) is linear and continuous. Here, the interesting situation is, when \( z \) cannot be reached due to the presence of the control constraints (non-attainability). The set of admissible functions is abbreviated by \( U_{ad} := \{ u \in L^2(\Omega) : u_a \leq u \leq u_b \} \). We are interested in an iterative method to solve \((P)\) based on generalized Bregman distances. In [1] the algorithm was analysed under a suitable regularity assumption. Here we recall the most important results, followed by numerical results.

*Department of Mathematics, University of Würzburg, Emil-Fischer-Str. 40, 97074 Würzburg, Germany, E-mail: frank.poerner@mathematik.uni-wuerzburg.de
2 Bregman iteration

The Bregman distance associated with the regularization functional $J : u \mapsto \frac{1}{2}\|u\|_{L^2(\Omega)}^2 + I_{U_{\text{ad}}}(u)$ is defined as $D^\lambda(u, v) := J(u) - J(v) - (u - v, \lambda)$ where $\lambda \in \partial J(v)$. In the following $(\alpha_k)_k$ denotes a positive, bounded sequence of real numbers. The algorithm is given by:

1. Let $u_0 = P_{U_{\text{ad}}}(0) \in U_{\text{ad}}$, $\lambda_0 = 0 \in \partial J(u_0)$ and $k = 1$.
2. Solve for $u_k$: Minimize $\frac{1}{2}\|Su - z\|^2_Y + \alpha_k D^{\lambda_{k-1}}(u, u_{k-1})$.
3. Set $\lambda_k := \sum_{i=1}^k \frac{1}{\alpha_i} S^*(z - Su_i) \in \partial J(u_k)$.
4. Set $k := k + 1$, go back to 1.

The algorithm is well-defined due to the convexity of $D^\lambda(\cdot, v)$ with respect to the first argument (see [1] and the references therein).

3 A-priori error estimates

Let $u^\dagger$ be a solution of $(P)$ and $p^1 = S^*(Su - z)$ be the adjoint state, then $(p^1, u - u^\dagger) \geq 0$, $\forall u \in U_{\text{ad}}$ is satisfied. To derive our error estimates furthermore assume that there exists a set $I \subset \Omega$, $w \in Y$ and $\kappa, c > 0$ such that $I \supset \{x \in \Omega : p^1(x) = 0\}$ holds. In addition assume that $\chi_I u^\dagger = \chi_I P_{U_{\text{ad}}}(S^*w)$ and $S^*w \in L^\infty(\Omega)$ holds. On the set $A := \Omega \setminus I$ we assume that the following structural assumptions $\{x \in A : 0 < |p^1(x)| < \varepsilon\} \leq c \varepsilon^\kappa \forall \varepsilon > 0$ holds.

Under this regularity assumption strong convergence of the iterates $(u_k)_k$ can be established together with the a-priori error estimate

$$\|u^\dagger - u_k\|^2_{L^2(\Omega)} \leq O\left(\gamma_k^{-1} + \gamma_k^{-1} \sum_{j=1}^k \alpha_j^{-1} \gamma_j^{-\kappa}\right),$$

with the abbreviation $\gamma_k := \sum_{j=1}^k \alpha_j^{-1}$. For details - both for the regularity assumption and the convergence - we refer to [1]. For the special choice of a constant sequence $\alpha_k = \alpha > 0$ and $\kappa < 1$ the a-priori estimate reduces to $\|u^\dagger - u_k\|^2 = O(k^{-\kappa})$ and to $\|u^\dagger - u_k\|^2 = O\left(k^{-1} \log(k)\right)$ for $\kappa = 1$.

4 Numerical examples

In this section we present numerical results. The implementation is done in FEniCS [2] with a semi-smooth Newton solver (see [3]). We use constant $\alpha_k = \alpha$ and compute the numerical approximation

$$\kappa_k := \frac{1}{\log(2)} \log\left(\frac{\|u_k/2 - u^\dagger\|^2_{L^2(\Omega)}}{\|u_k - u\|^2_{L^2(\Omega)}}\right).$$

2
for bang-bang test examples (\(A = \Omega\)). Here our operator \(y = Su\) is chosen to be the solution of the equation \(-\Delta y = u\) in \(\Omega\) and \(y = 0\) on \(\partial\Omega\). First we compute 1D examples with \(\kappa = 1\), \(\kappa = \frac{1}{2}\), and \(\kappa = \frac{1}{3}\) for different mesh sizes \(h\). The results are listed in Table 1 and 3 respectively. For the details of the construction of bang-bang examples with given adjoint state \(p^*\) we refer to [4, Chapter 2.9]. To obtain \(\kappa = 1\) we use \(p^*(x) = \sin(\pi x)\) on \(\Omega = [-1, 1]\). The other examples can be constructed using polynomials and limiting the slope near the zeros. For \(\kappa = \frac{1}{3}\) we use \(p^*(x) = x(1-x)(3x-1)^3\) on \(\Omega = [0, 1]\).

Second we present a 2D bang-bang example, namely \(p^*(x, y) = \sin(2\pi x)\sin(2\pi y)\) on \(\Omega = [0, 1]^2\). Numerical estimates indicate \(\kappa = 1\), which is supported by our numerical results. Note that if the grid is too coarse the discretization error is dominating the regularization error, leading to unreliable results for \(\kappa_k\). In all cases we obtain \(\kappa_k \approx \kappa\) for \(k\) large and \(h\) small enough, indicating that our a-priori error estimate is sharp for the bang-bang case.

| \(h\) | \(10^{-3}\) | \(10^{-4}\) | \(10^{-5}\) | \(10^{-6}\) | \(\kappa_k\) | \(\kappa_k\) | \(\kappa_k\) | \(\kappa_k\) |
|---|---|---|---|---|---|---|---|---|
| 4  | 0.646 | 0.602 | 0.601 | 0.601 | 0.522 | 0.520 | 0.520 | 0.520 |
| 8  | 0.839 | 0.752 | 0.750 | 0.750 | 0.648 | 0.644 | 0.643 | 0.643 |
| 16 | 1.027 | 0.860 | 0.856 | 0.857 | 0.641 | 0.635 | 0.634 | 0.634 |
| 32 | 1.211 | 0.927 | 0.922 | 0.923 | 0.646 | 0.636 | 0.635 | 0.635 |
| 64 | 1.229 | 0.960 | 0.958 | 0.960 | 0.639 | 0.624 | 0.622 | 0.622 |
| 128| -0.001| 0.945 | 0.975 | 0.979 | 0.625 | 0.605 | 0.602 | 0.602 |
| 256| -0.004| 0.786 | 0.980 | 0.989 | 0.609 | 0.585 | 0.581 | 0.581 |
| 512| -0.020| 0.271 | 0.972 | 0.991 | 0.591 | 0.567 | 0.562 | 0.562 |
| 1024| -0.081| -0.054| 0.938 | 0.978 | 0.571 | 0.553 | 0.547 | 0.546 |
| 2048| -0.217| -0.149| 0.826 | 0.919 | 0.545 | 0.542 | 0.534 | 0.534 |

Table 1: 1D example 1 (\(\kappa = 1\)).

| \(h\) | \(10^{-3}\) | \(10^{-4}\) | \(10^{-5}\) | \(10^{-6}\) | \(\kappa_k\) | \(\kappa_k\) | \(\kappa_k\) | \(\kappa_k\) |
|---|---|---|---|---|---|---|---|---|
| 4  | 0.286 | 0.286 | 0.286 | 0.286 | 0.509 | 0.472 | 0.458 | 0.456 |
| 8  | 0.312 | 0.312 | 0.312 | 0.312 | 0.676 | 0.622 | 0.595 | 0.592 |
| 16 | 0.325 | 0.327 | 0.328 | 0.328 | 0.789 | 0.759 | 0.711 | 0.705 |
| 32 | 0.329 | 0.337 | 0.338 | 0.338 | 0.720 | 0.885 | 0.803 | 0.791 |
| 64 | 0.321 | 0.339 | 0.340 | 0.341 | 0.411 | 1.000 | 0.884 | 0.863 |
| 128| 0.301 | 0.338 | 0.340 | 0.340 | 0.216 | 1.027 | 0.968 | 0.935 |
| 256| 0.272 | 0.335 | 0.338 | 0.339 | 0.145 | 0.855 | 1.039 | 1.012 |
| 512| 0.236 | 0.332 | 0.337 | 0.338 | 0.166 | 0.556 | 1.011 | 1.039 |
| 1024| 0.193| 0.328 | 0.336 | 0.337 | 0.129 | 0.295 | 0.805 | 0.936 |
| 2048| 0.132| 0.322 | 0.335 | 0.336 | -0.045| 0.126 | 0.545 | 0.693 |

Table 3: 1D example 3 (\(\kappa = \frac{1}{3}\)).

| \(h\) | \(10^{-3}\) | \(10^{-4}\) | \(10^{-5}\) | \(10^{-6}\) | \(DOF\) | \(10^4\) | \(10^5\) | \(10^6\) | \(2 \cdot 10^6\) |
|---|---|---|---|---|---|---|---|---|---|
| 4  | 0.286 | 0.286 | 0.286 | 0.286 | 4 | 0.509 | 0.472 | 0.458 | 0.456 |
| 8  | 0.312 | 0.312 | 0.312 | 0.312 | 8 | 0.676 | 0.622 | 0.595 | 0.592 |
| 16 | 0.325 | 0.327 | 0.328 | 0.328 | 16 | 0.789 | 0.759 | 0.711 | 0.705 |
| 32 | 0.329 | 0.337 | 0.338 | 0.338 | 32 | 0.720 | 0.885 | 0.803 | 0.791 |
| 64 | 0.321 | 0.339 | 0.340 | 0.341 | 64 | 0.411 | 1.000 | 0.884 | 0.863 |
| 128| 0.301 | 0.338 | 0.340 | 0.340 | 128| 0.216 | 1.027 | 0.968 | 0.935 |
| 256| 0.272 | 0.335 | 0.338 | 0.339 | 256| 0.145 | 0.855 | 1.039 | 1.012 |
| 512| 0.236 | 0.332 | 0.337 | 0.338 | 512| 0.166 | 0.556 | 1.011 | 1.039 |
| 1024| 0.193| 0.328 | 0.336 | 0.337 | 1024| 0.129 | 0.295 | 0.805 | 0.936 |
| 2048| 0.132| 0.322 | 0.335 | 0.336 | 2048| -0.045| 0.126 | 0.545 | 0.693 |

Table 4: 2D example (\(\kappa = 1\)).
References

[1] D. Wachsmuth and F. Pörner, arXiv:1603.05792 (2016)

[2] FEniCS, FEniCS project, http://www.fenicsproject.org/, 2016

[3] S. Beuchler, C. Pechstein and D. Wachsmuth, Comput Optim Appl. 51, 883-908 (2010)

[4] F. Tröltzsch, Optimal Control of Partial Differential Equations (American Mathematical Society, Providence, RI, 2010)