Optimality conditions and optimization algorithms in infinite-dimensional spaces

Victor Tolstykh
https://tolstykh.com

Donetsk State University, Donetsk, Russia

2. Finite-dimensional spaces

Smooth objective function $J(u)$, control $u \in E^n$.
Classical necessary condition of optimality (NCO):

$$\|\nabla J(u^k)\|_{E^n} \xrightarrow{k\to\infty} 0. \tag{1}$$

It can be implemented using gradient methods, e.g.:

$$u_i^{k+1} = u_i^k - b^k \nabla_i J^k, \quad i = 1, ..., n, \; k = 0, 1, ...$$

In this case, the convergence of the control vectors u^k to optimum u_* in norm implies convergence in components (wonderfully!):

$$\|u^k - u_*\|_{E^n} \xrightarrow{k\to\infty} 0 \quad \Longrightarrow \quad \left| u_i^k - u_{*i} \right|_{\forall i=1...n} \xrightarrow{k\to\infty} 0.$$

3. Infinite-dimensional spaces

Fréchet differentiable functional $J(u)$, control $u(\tau) \in L_2(S)$.
Classical NCO:

$$\|u^k(\tau) - u_*(\tau)\|_{L_2(S)} \xrightarrow{k\to\infty} 0. \tag{2}$$

It is implemented by gradient methods, e.g.:

$$u^{k+1}(\tau) = u^k(\tau) - b^k \nabla J(u^k; \tau), \quad \tau \in S, \; k = 0, 1, \ldots$$

However, convergence in norm in $L_2(S)$ (integral over the root-mean-square) does not imply uniform (pointwise) convergence of control functions (bad!):

$$\|u^k - u_*\|_{L_2(S)} \xrightarrow{k\to\infty} 0 \quad \not\Rightarrow \quad \left|u^k(\tau) - u_*(\tau)\right| \xrightarrow{k\to\infty} 0 \;\forall \tau \in S.$$

4. Infinite-dimensional optimality conditions

If S contains a zero-measure subset where pointwise convergence is impossible, then restrict S to S_Δ, where pointwise convergence is possible in principle [*Controlability conditions*].

Theorem 1 (Strong NCO).
If the strictly convex quadratic functional $J(u)$, $u \in L_2(S)$, has a minimum at u_*, and the sequence of controls u^k satisfies $\left|u_* - u^k\right| \xrightarrow{k\to\infty} 0$ uniformly on S_Δ, then:

$$\left|\nabla J(u^k;\tau)\right| \xrightarrow{k\to\infty} 0 \text{ uniformly on } S_\Delta \subseteq S. \qquad (3)$$

Theorem 2 — traditional (Weak NCO)
$\|\nabla J(u^k;\tau)\|_{L_2(S)} \xrightarrow{k\to\infty} 0.$

Does not depend on S_Δ and does not satisfy (3).

5. Infinite-dimensional optimization algorithms

Gradient-based method with adjustable descent direction (**ADDMg**) for convex functionals:

$$u^{k+1} = u^k - b^k \alpha \nabla J^k \text{ uniformly on } S_\Delta, \quad k = 0, 1 \dots \tag{4}$$

The parameter $\alpha^k(\tau) \in C^1_+(S_\Delta)$ controls the direction of descent to ensure uniform convergence, as required by the strict NCO conditions of Theorem 1. It is necessary for the first step:

$$u^0(\tau) \xrightarrow{\text{uniform step}} u^1(\tau), \quad \tau \in S_\Delta,$$
$$\nabla J(u^0;\tau) \xrightarrow{\text{uniform change}} \nabla J(u^1;\tau), \quad \tau \in S_\Delta.$$

I.e. $\forall \tau \in S_\Delta$ the step $\left|u^1 - u^0\right|$ should be appreciable and should lead to appreciable changes $\left|\nabla J(u^0) - \nabla J(u^1)\right|$.

6. Practical recommendations for choosing α

1. Test: ADDMg should start with the parameter $\alpha = 1$, i.e., as a gradient method. If convergence is bad, then:
2. Do the first step "at 45°": $u^1(\tau) = u^0(\tau) \pm \delta$. We get

$$\alpha(\tau) = \frac{\delta u^0(\tau)}{|\nabla J(u^0;\tau)|}, \quad \operatorname{sgn} J(u^0;\tau) = const, \ \tau \in S_\Delta,$$

   where $\delta > 0$ shifts the function $u^0(\tau)$ "at 45°".
3. If step 2 is not effective enough, you should either change u^0 or set a trial function $u^1(\tau)$ for

$$\alpha(\tau) = \left| \frac{u^1(\tau) - u^0(\tau)}{\nabla J(u^0;\tau)} \right|, \quad \operatorname{sgn} J(u^0;\tau) = const, \ \tau \in S_\Delta,$$

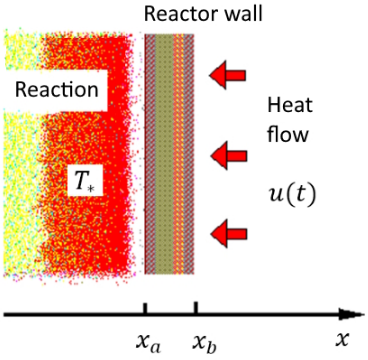
   which satisfy to Strong NCO.

7. Illustration of infinite-dimensional optimization

One-dimensional
non-stationary
heat transfer process:

$$C\rho\frac{\partial T}{\partial t} - \lambda\frac{\partial^2 T}{\partial x^2} = 0,$$

$T(x,t)$ is the temperature in the reactor wall, C, ρ, and λ are the heat capacity, density, and thermal conductivity.
The boundary conditions are:



$$\lambda\frac{\partial T}{\partial x} = q \text{ on } \Gamma_a = x_a \times (t_0, t_1), \quad \lambda\frac{\partial T}{\partial x} = u \text{ on } \Gamma_b = x_b \times (t_0, t_1),$$

$$T = T_0, \quad \text{on } [x_a, x_b] \times t_0.$$

8. Objective functional and gradient

Find the heat flux density $u(t) \in L_2(S)$ that minimizes the quadratic objective functional

$$J(u) = \int_{t_0}^{t_1} \big(T(x_a, t) - T_*(t)\big)^2 dt,$$

where $T_*(t)$ is the nominal reaction temperature.
The gradient

$$\nabla J(u;t) = -f \quad \text{на } S_\Delta.$$

Adjoint problem:

$$-C\rho \frac{\partial f}{\partial t} - \lambda \frac{\partial^2 f}{\partial x^2} = 0,$$

$$\lambda \frac{\partial f}{\partial x} = 2\big(T\big|_{\Gamma_a} - T_*\big) \ \text{ on } \Gamma_a, \quad \lambda \frac{\partial f}{\partial x} = 0 \ \text{ on } \Gamma_b,$$

$$f = 0 \ \text{ on } [x_a, x_b] \times t_1.$$

9. Preparation of test calculations

Space-time grid 10×100 cells. Problem solving time — 150 s. Selected time step $\Delta t = 1.5$ s, $S_\Delta = x_b \times [1.5, 148.5]$ s. The test control was set

$$u_*(t) = 150 + 50 \sin \frac{2\pi t}{t_1 - t_0},$$

the temperature T_* was calculated, and the inverse problem of recovering the control $u_*(t)$ that minimizes the objective functional $J(u)$ was solved with precision

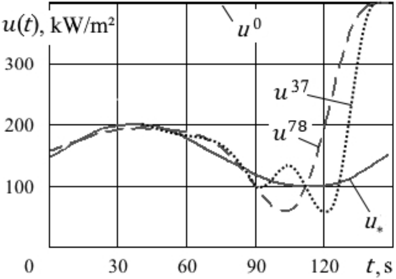
$$\frac{\|u^k - u^{k-1}\|}{\|u^{k-1}\|} \leq 10^{-6}.$$

Initial guess

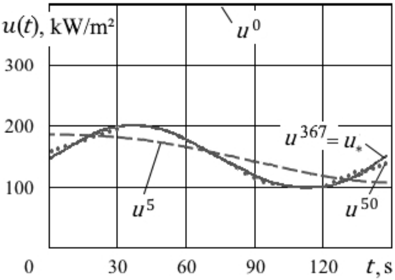
$$u^0(t) = 400.$$

10. Results: Traditional methods

Steepest descent method — SDM (dotted line)
and L-BFGS (dots)



11. Results: Adjustable descent direction method — ADDMg



Parameter for adjusting the descent direction

$$\alpha(t) = \frac{0.2u^0}{|\nabla J(u^0;t)|}, \quad \delta = 0.2.$$

The convergence is uniform and ended at $k = 367$. The value of $J(u)$ decreased by 5 orders of magnitude better than SDM.

Publications

- Tolstykh V.K. Collinear Gradients Method for Minimizing Smooth Functions. Oper. Res. Forum. 2023. Vol. 4, no 20.
- Tolstykh V.K. Practical optimization and identification of distributed systems. Nauka, Moscow, 2026. [In Russian]
- Tolstykh V.K. On the gradient in optimization problems of nonstationary systems with distributed control. Numerical Methods and Programming. 2025. Vol. 26, no 3. Pp. 229–244.
- Tolstykh V.K. Algorithms for optimizing systems with multiple extremum functionals. J. of Calcul. Mathem. and Mathem. Physics. 2024. Vol. 64, no 3. Pp. 415–423.
- Tolstykh V.K. Controllability of distributed parameter systems. J. of Calcul. Mathem. and Mathem. Physics. 2024. Vol. 64, no 6. Pp. 959–972.