Application of gradient methods to solving ill-posed problems of mathematical physics

Nikita V. Pletnev
Moscow Institute of Physics and Technology

ASCOMP 2024

Ill-posed and inverse problems

- Various physical meaning and applications. Different type: differential or integral equations.
- A small change in the initial data leads to a significant change in the solution.
- Often ill-posed problems are inverse to well-posed and relatively easy-to-solve problems.
- The main idea is to replace condition f by condition q on another boundary.
- So we seek q making the well-posed problem equivalent to ill-posed.

Operator of the problem. Reduction to optimization

- Let $q : [0, 1] \to \mathbb{R}$ be an element of functional Hilbert space H with usual scalar product.
- We define the operator $A : H \to H$, which associates the corresponding f to a known q.
- Calculation of Aq is a well-posed (direct) problem.
- Each ill-posed Cauchy problem under consideration is reduced to the operator equation $Aq = f$.
- $Aq^* = f \Rightarrow q^* = \arg\min_{q \in H} J(q)$. If $J(q^*) = 0$, q^* is the solution. Otherwise, solution does not exist.
- $J(q) = \frac{1}{2}||Aq - f||^2$ is convex and smooth functional. Its gradient can be calculated by general formula: $\nabla J(q) = A^*(Aq - f)$, A^* is a conjugate operator, similar (sometimes equal) to A, can be found by the method of Lagrange's multipliers.

Optimization methods: classic

- Gradient descent: $q_{n+1} = q_n - \alpha_n \nabla J(q_n)$. $\quad J(q_n) = O(1/n)$.
  Step: constant; $\alpha_n = \frac{||\nabla J(q_n)||^2}{||A_0 \nabla J(q_n)||^2}$ (fastest descent); $\alpha_n = \frac{2J(q_n)}{||\nabla J(q_n)||^2}$ (nearest descent).
- Accelerated methods: $q_{n+1} = q_n - \alpha_n \nabla J(q_n) + \beta_n(q_n - q_{n-1})$. $\quad J(q_n) = O(1/n^2)$.
  Conjugate Gradient Descent ($J(q_{n+1})$ is minimized on every step); Heavy Ball; STM.

New method: Conjugate Gradient Descent with minimization of distance to the exact solution

$$q_{n+1} = q_n - \alpha_n \nabla J(q_n) + \beta_n(q_n - q_{n-1}); \quad (\alpha_n, \beta_n) = \arg\min_{\alpha_n, \beta_n} ||q_{n+1} - q^*||^2.$$

It can be rewritten:

$$s_n = -\nabla J(q_n) + \frac{\langle \nabla J(q_n), s_{n-1} \rangle}{||s_{n-1}||^2} s_{n-1}; \quad q_{n+1} = q_n + \frac{2J(q_n)}{||s_n||^2} s_n.$$

- The nearby steps produced by this method are orthogonal: $\langle s_n, s_{n-1} \rangle = 0$.
- q_n converges to q^*, and the distance decreases monotonically.
- The functional converges to 0, but not monotonically.
- The method is appropriate only for quadratic functions, but it's OK: $J(q)$ is quadratic.
- $||q_{n+1} - q^*||^2 \le \left(1 - \frac{\mu}{L \sin^2 \varphi_n}\right) \cdot ||q_n - q^*||^2$, where φ_n is angle between $-\nabla J(q_n)$ and $q_n - q_{n-1}$.
- The method can be improved by using more steps in definition of s_n. The modification using *all* previous steps is the best first order method, but its computational complexity is $O(n^2)$, not $O(n)$.

Results are submitted at arXiv.org: «On the modification of the conjugate gradient method with minimization of the distance to the exact solution when choosing the step length».

Retrospective Cauchy problem for heat equation: parabolic type

$$\begin{cases} u_t - \kappa^2(x,t)\Delta_x u = 0, & (x,t) \in \Omega = \Pi \times (0,1) \\ u|_{x\in\partial\Pi} = 0 \text{ or } \frac{\partial u}{\partial n}|_{x\in\partial\Pi} = 0, & t \in [0,1] \\ u|_{t=1} = f(x), & x \in \Pi \end{cases}$$

$\kappa(x,t)$, $f(x) = (Aq^*)(x)$ are known continuous functions.

$(Aq)(x) = u(x,1)$, where $q(x) \in C^2(\Pi)$, $u(x,t)$ is solution of

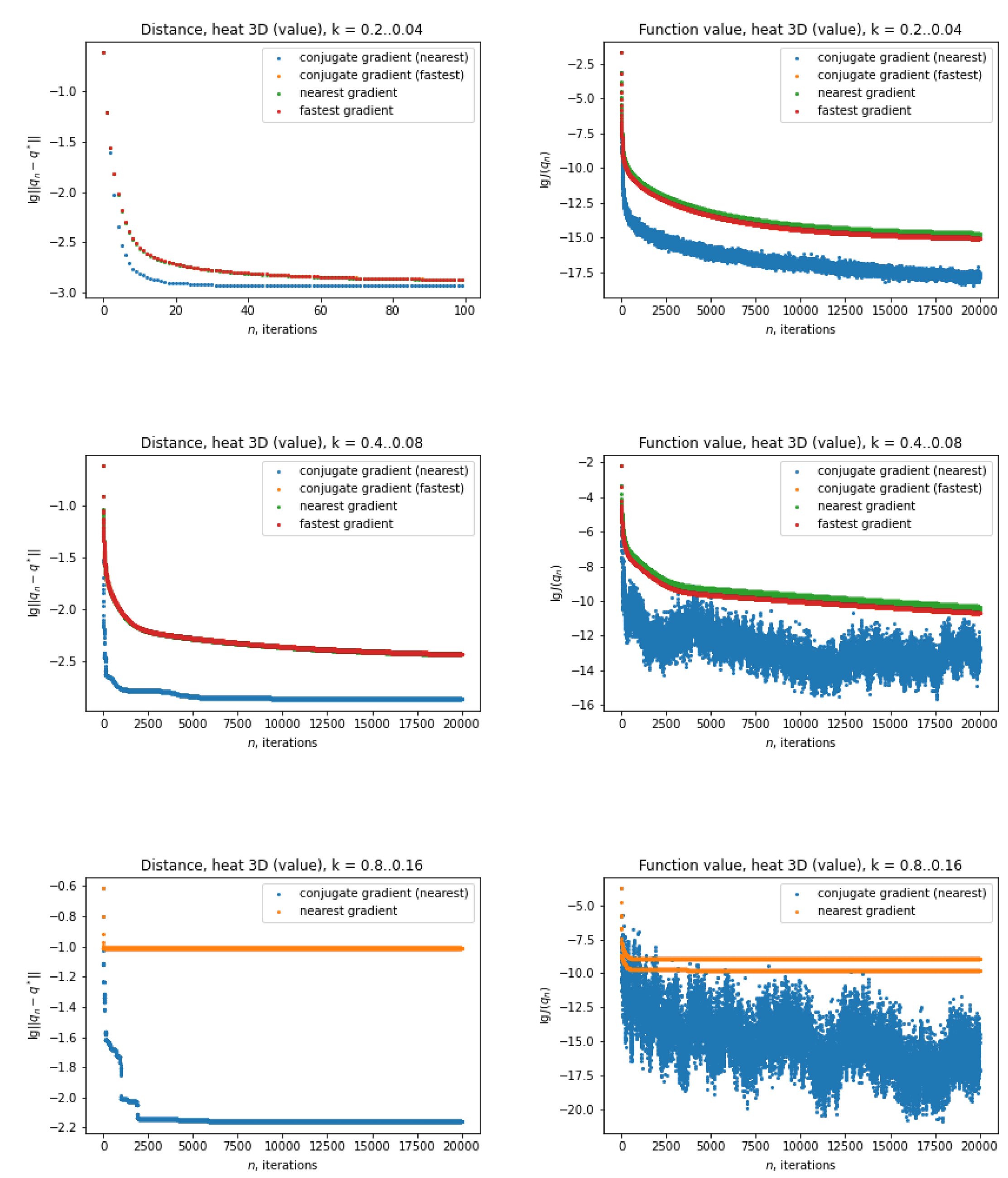
$$\begin{cases} u_t - \kappa^2(x,t)\Delta_x u = 0, & (x,t) \in \Omega = \Pi \times (0,1) \\ u|_{x\in\partial\Pi} = 0 \text{ or } \frac{\partial u}{\partial n}|_{x\in\partial\Pi} = 0, & t \in [0,1] \\ u|_{t=0} = q(x), & x \in \Pi \end{cases}$$

$$\kappa(x,t) \equiv const \Leftrightarrow A^* = A.$$

Experiments: boundary conditions on the value

$f(x) = (Aq^*)(x)$, where $q^*(x) = \sin 2\pi x_1 \sin^2 2\pi x_2 \sin^3 2\pi x_3$,

$$\kappa(x,t) = \begin{cases} \kappa_{max}, & 0.4 < x_i < 0.6, \ i = 1,2,3 \\ \frac{\kappa_{max}}{5}, & \text{otherwise} \end{cases}$$



Achieved distance to minimum point (boundary conditions on value), $\times 10^{-3}$. Initial distance: 0.242.

κ_{max}	grad J	grad ρ	conj J	conj ρ
0.2	1.1757	1.1757	1.1757	1.1756
0.4	3.67	3.66	3.67	1.36
0.6	46.8	40.0	46.8	3.17
0.8	-	96.7	-	6.95
1.0	-	-	-	21.1
1.2	-	-	-	47.6

Fredholm integral equation of the first kind

$$\int_0^1 K(x,s)q(s)ds = f(x), \quad x \in [0,1]$$

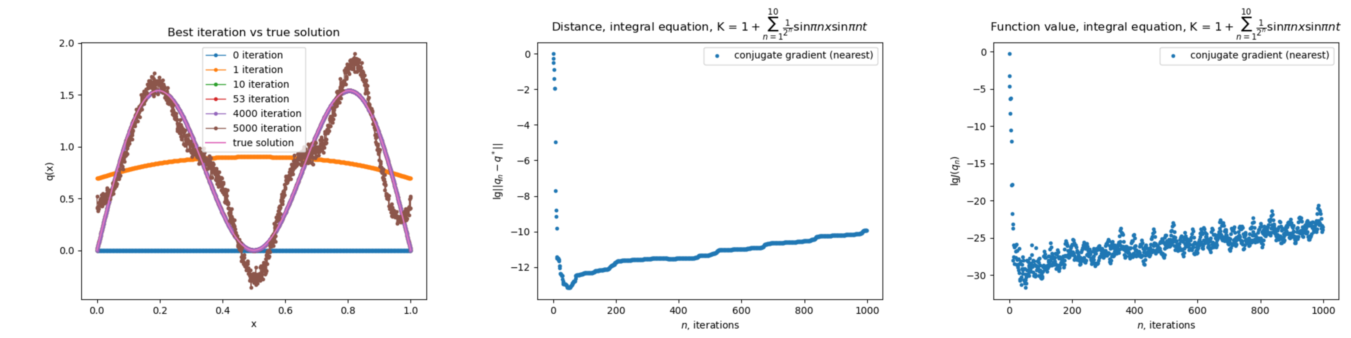
$K(x,s)$, $f(x) = (Aq^*)(x)$ are known continuous functions.

$$(Aq)(x) = \int_0^1 K(x,s)q(s)ds, \quad x \in [0,1];$$
$$K(x,s) = K(s,x) \Leftrightarrow A^* = A.$$

Experiments

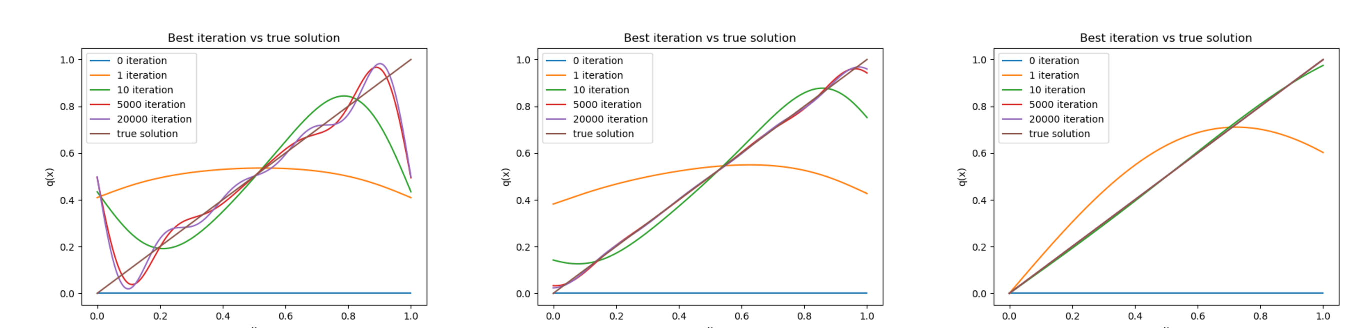
Symmetric kernel: conjugate gradient method with minimization of distance to the exact solution

$$K(x,s) = 1 + \sum_{n=1}^{10} \frac{1}{2^n} \sin \pi n x \sin \pi n s; \ q^*(x) = \sin \pi x + \sin 3\pi x.$$



After quickly reaching the optimum (distance to the exact solution is $6.789 \cdot 10^{-14}$, functional value is $2.126 \cdot 10^{-32}$) on 53-th iteration the quality deteriorates.

Reconstruction of $q^*(x) = x$ for symmetric kernels: conjugate gradient method with minimization of functional



$K(x,s)$	$\|\|q - q^*\|\|_2$	$J(q)$
$1 + \sum_{n=1}^{10} \frac{1}{2^n} \sin \pi n x \sin \pi n s$	0.104	$2.39 \cdot 10^{-10}$
$1 + \sum_{n=1}^{10} \frac{1}{2^n} \cos \pi n (x - s)$	$6.79 \cdot 10^{-3}$	$8.08 \cdot 10^{-11}$
$\sin \pi x s$	$1.52 \cdot 10^{-4}$	$3.54 \cdot 10^{-17}$

Conclusions

- Conjugate gradient descent with minimization of the distance to the exact solution is quite simple and effective to solve inverse and ill-posed problems.
- Retrospective Cauchy problem for heat equation is as worse as κ is larger. New method can be used for $\kappa \lesssim 0.8$, unlike old methods (only for $\kappa \lesssim 0.5$).
- Integral equation is solved very good, but using of new method requires the stopping rule.