Andrei Draganescu

A fast multigrid method for inverting linear parabolic problems

The University of Chicago
Department of Mathematics
5734 S University Avenue
Chicago
IL 60637
draga@math.uchicago.edu

This work started as a bid for a cost-efficient method for inverting time-dependent partial differential equations. The original motivation comes from the problem of comparing simulations with experimental results. Given that only partial information is available from experiments, we address the question of finding the best possible match by controlling initial conditions and/or other parameters. This is formulated as an optimization process where the cost functional includes a measure of the distance between simulated and experimental data and regularizing terms. Since in general we may be talking about large problems, for which the forward simulation may take many hours on a powerful machine, it is clear that the attached optimization problem should take no longer than a factor of ten of that time, a demand which is in conflict with the large size of the control space.

As our first model problem we have selected the following inverse parabolic problem: suppose $(x,t) \mapsto U(x,t)$ satisfies a linear parabolic initial value problem with $U(\cdot,0) = v$, and we are given a final state $f \in L^2$ (= the "measurements") at time T. Because f may be only a reconstruction from partial data (e.g. point values, spatial averages), or simply the measurements themselves contain errors, we assume that an initial condition corresponding exactly to the measurements is not unique or doesn't exist. If we have some knowledge about the initial condition leading closely to f, i.e. we have a guess v_0 at the initial value, the question is to find $v \in L^2$ that minimizes $\|U(\cdot,T) - f\|^2_{L^2} + \epsilon\|v - v_0\|^2_{L^2}$. The choice of ϵ in the regularizer depends on the measure of uncertainty in the "measurements" f (an alternative regularizer is $\epsilon|v - v_0|^2_{H^1}$). Being quadratic, the optimization problem converges in one Newton step. The challenge is to invert the Hessian H_ϵ of the cost functional efficiently. If we denote by $\mathcal{S}$ the operator $v \mapsto U(\cdot,T)$ and by $\mathcal{S}^*$ its L^2-adjoint then $H_\epsilon = \epsilon I + \mathcal{S}^*\,\mathcal{S}$. Here, unlike the case of standard multigrid for the Poisson problem, residual computation is expensive: Hessian-vector multiplication (both discrete and continuous) amounts to two forward solves; moreover, the Hessian is a full matrix when using the nodal basis for the finite element space. Direct spectral analysis for the case of the finite difference discretization for the heat equation on a uniform grid shows

that the conjugate gradient method will converge in at most 10 iterations for reasonable choices of all parameters involved, regardless of the resolution. This is due to the exponential decay of the eigenvalues of H_ϵ down to ϵ.

The multigrid algorithm we propose – a mixture between a $\mathcal{V}$ and $\mathcal{W}$-cycle – requires **only one residual computation** at fine (**how** fine depends on ϵ) resolution and two residual computations at coarse resolution in order to resolve the problem to optimal order. The multigrid iteration is based on the design of a multilevel, recursively defined preconditioner for the Hessian of the discrete cost functional. We introduce a scale-free *spectral distance* function on the set of linear operators with positive-definite symmetric part; the spectral distance between the inverse of an operator and its preconditioner gives an estimate of the convergence rate in the simple iteration. Our key result consists in showing that the spectral distance between the inverse of the discrete Hessian and the constructed preconditioner is $O\left((\epsilon\, T)^{-1} h^2\right)$, hence it has the same magnitude as the error between the discrete and the continuous solutions of the minimization problem. Work estimates for the full multigrid algorithm show that in two dimensions the cost is at most $2.5\times$(the work for one residual computation) in addition to the total cost of all exact solves performed at the coarsest level.

This work extends naturally in a few directions. The square-variation regularizer produces a Hessian that is a sum between a smoothing operator – $\mathcal{S}^*\,\mathcal{S}$ – and a roughening one: the Laplacian. For this case we propose a slightly nonsymmetric preconditioner that still approximates the inverse of the discrete Hessian to optimal order in the spectral distance. A second direction refers to the case of semilinear parabolic equations. Adjoint methods allow for exact calculation of gradients and Hessian-vector multiplication at a cost equivalent to two forward computations each. The Hessian has almost the same smoothing properties and in the linear parabolic case, and this is a basic argument in our analysis. Finally, the nonstationary Stokes equations are parabolic, hence we expect our methods to work well for the associated discrete inverse problems, with possible extension to control problems for Navier-Stokes.