Low-Rank Methods for Computations of Posterior Covariance Matrices in Bayesian Inverse Problems

Yue Qiu

We consider the problem of estimating the uncertainty in statistical inverse problems using Bayesian inference. When the probability density of the noise and the prior are Gaussian, the solution of such a statistical inverse problem is also Gaussian. Therefore, the underlying solution is characterized by the mean and covariance matrix of the posterior probability density. However, the covariance matrix of the posterior probability density is full and large. Hence, the computation of such a matrix is impossible for large dimensional parameter spaces. It is shown that for many ill-posed problems, the Hessian matrix of the data misfit part has low numerical rank and it is therefore possible to perform a low-rank approach to approximate the posterior covariance matrix. For such a low-rank approximation, one needs to solve a forward partial differential equation (PDE) and the adjoint PDE in both space and time. This in turn gives $O(n_x n_t)$ complexity for both, computation and storage, where $n_x$ is the dimension of the spatial domain and $n_t$ is the dimension of the time domain. Such computations and storage demand are infeasible for large problems.

To overcome this obstacle, we develop a new approach that utilizes a recently developed low-rank in time algorithm together with the low-rank Hessian method in [1]. This low-rank approach exploits the tensor structure of the underlying problem, which is the crucial part for the low-rank approximation. To perform the low-rank approximation, one needs to solve a forward and adjoint tensor equation. The computed solutions should also be of low-rank tensor format. To solve such a tensor equation, we make use of the “alternative minimal energy (AMEN)” approach developed for the tensor train (TT), where preconditioning is necessary to accelerate the convergence of the “AMEN” solver. We reduce both the computational complexity and storage requirement from $O(n_x n_t)$ to $O(n_x + n_t)$. We use numerical experiments to illustrate the advantages of our approach.

References


1Max Planck Institute for Dynamics of Complex Technical Systems, Germany (qiu@mpi-magdeburg.mpg.de)