My Project

Generated by Doxygen 1.9.4

Chapter 1

Design of an adaptive pre-conditioned CG solver

We make a brief note here on rules implemented in the ADP-CG solver.

1.0.0.0.1 Target problem   The solver targets consecutive linear systems $A^k x^{k,j} = b^{k,j}, j = 1, ..., r^k$, where, given a sequence of LHS matrices A^k , we solve a set of linear systems with different RHS. $b^{k,j}$ using pre-conditioned conjugate gradient method. One feature we expect is that spectrums of A^k do not change too aggressively and therefore, it is reasonable to assume that if an expensive but accurate pre-conditioner is computed for A^k , it should also be valid for several future $A^{k+1}, ..., A^{k+t}$. The above scenario is quite common for the normal equation in the interior point method ADA^T and ADP-CG is designed for this purpose.

1.0.0.0.2 Origin   The ADP-CG solver is initially written to accelerate the SDP solver HDSDP [1].

1.0.0.0.3 Adaptive pre-conditioning   One most critical ingredient of ADP-CG is to find **when to update the pre-conditioner**. If the pre-conditioner is too outdated, it is possible that the spectrum after pre-conditioning gets worse. Therefore, we have to decide when to update the pre-conditioner.

In ADP-CG, we use a series of rules to decide whether to update the pre-conditioner.

First we introduce some definitions of statistics that aid our decision

1. We call solution to linear systems with LHS A^k a *round* indexed by k

2. A *solve* is defined by the solution of $A^k x^{k,j} = b^{k,j}$ for some j and round k contains r^k solves

   A solve is either performed by CG or direct solver

3. A *factorization* is defined by the action to factorize A^k for a Cholesky pre-conditioner

4. The (average) solution time of a solve is defined by the (average) time spent in CG (excluding time building pre-conditioner)

5. The (average) factorization time is defined by the (average) time of a factorization

6. A solve is called SUCCESS if the residual norm reaches below tolerance within maximum iteration number

   A solve is called MAXITER if CG exhibits convergence but residual norm fails to reach tolerance within maximum iteration number

   A solve is called FAILED if CG does not exhibit convergence or there is an irreparable error (to be clarified later)

7. Given a pre-conditioner, its *nused* property refers to the number of rounds it has gone through without update

8. The *latesttime* property refers to the average solution time in the lastest round

Based on the above statistics, we now clarify the rules.

1. Update of diagonal pre-conditioner always happens at the beginning of a round and the update of Cholesky pre-conditioner happens **either** at the beginning a round **or** within the first solve in a round

2. In the first solve of each round, if the Cholesky pre-conditioner is not updated at the beginning of the round, there is a chance to regret

   i.e., if the first CG loop is not SUCCESS due to diagonal pre-conditioner or an outdated Cholesky pre-conditioner, it is allowed to perform a make-up Cholesky factorization step and then update the pre-conditioner. Note that by rule (4), the rest of the solves in this round would by done by direct solver and by rule (5), diagonal pre-conditioner will never be used.

3. At the beginning of each round, pre-conditioner is updated if one of the following criteria, checked in order, is satisfied

   - If the system is classified as ill-conditioned or indefinite by some user-defined criterion
   - If the diagonal pre-conditioner is used
   - **If latesttime $>$ 1.5 average solution time**
   - If ADP-CG is asked to perform direct solve
   - **If average solution time $>$ average factorization time**
   - If the *nused* property of the current pre-conditioner exceeds the user-defined threshold

4. If the Cholesky pre-conditioner is updated in round k, then all the solves in round k after the update are solved by the direct solver

5. If CG switches to Cholesky pre-conditioner, it never returns to diagonal pre-conditioner unless requested by the user

6. The following cases result in a FAILED solve:

   - If pre-conditioner build-up fails
   - If direct solve fails
   - If pre-conditioning step fails
   - If NAN appears in the CG solver

   If FAILED occurs, the current solution is not trustworthy and the whole solution procedure stops due to irreparable error

Here are some extra rules tailored for the deteriorating conditioning of normal equations arising from the interior point method

1. If the number of MAXITER exceeds T in round k, then all the solves thereafter are solved by direct solver

   References∗∗

[1] Gao, Wenzhi, Dongdong Ge, and Yinyu Ye. "HDSDP: Software for Semidefinite Programming." *arXiv preprint arXiv:2207.13862* (2022).