R. Hiptmair
S. Hosseini
C. Pagliantini

Homework Problem Sheet 12

Introduction. This problem sheet is devoted to the Conjugate Gradient method.

Problem 12.1 CG Iteration Error

Consider a s.p.d. matrix $\mathbf{A} \in \mathbb{R}^{n,n}$, $n \in \mathbb{N}$, with spectrum $\sigma(\mathbf{A})$ that satisfies

$$\sigma(\mathbf{A}) \subset [1,2] \cup \{\xi\} , \quad \xi \gg 2 .$$

(12.1a) Give an estimate for the reduction of the energy norm of the iteration error achieved by the CG algorithm for solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ for any $\mathbf{b} \in \mathbb{R}^n$.

HINT: Take the cue from the analysis developed in Section 6.2.2 in NMI_6 using in particular Proposition 3.26.

Use the error decay formula for the conjugate gradient subspace correction scheme

$$\left\|\mathbf{e}^{(k)}\right\|_A \leq \inf\{\max_{\lambda\in\sigma(\mathbf{A})} |p(\lambda)|;\ p \in \mathbb{P}_k,\ p(0) = 1\}\left\|\mathbf{e}^{(0)}\right\|_A .$$

In this formula make a smart choice of the polynomial p.

(12.1b) Compare the result obtained in subproblem (12.1a) with the estimate obtained from a standard convergence theorem for the conjugate gradient method that relies on the spectral condition number $\kappa(\mathbf{A})$ only.

Problem 12.2 Conjugate Gradient Method: Number of Iterations

This exercise is devoted to the computation of the number of CG iterations required to reduce the initial error by a given factor.

Consider the matrix $\mathbf{A} = \left(4^{-|i-j|}\right)_{i,j=1}^{n} \in \mathbb{R}^{n,n}$, $n \in \mathbb{N}$.

(12.2a) Write an efficient MATLAB function `y = multA(x)` that realizes the multiplication of $\mathbf{A}$ with an arbitrary vector $\mathbf{x}$.

(12.2b) What is the asymptotic complexity of your implementation of `multA` in subproblem (12.2a) in terms of the problem size parameter n?

(12.2c) The Gershgorin circle theorem states that for any symmetric matrix $\mathbf{A} \in \mathbb{R}^{n,n}$, the spectrum $\sigma(\mathbf{A})$ of $\mathbf{A}$ satisfies

$$\sigma(\mathbf{A}) \subset \{\lambda \in \mathbb{R} : \exists i \in \{1,\ldots,n\} :\ |a_{ii} - \lambda| \leq \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|\} .$$

Use this result to prove that $\mathbf{A}$ is positive definite by deriving bounds for the largest and smallest eigenvalue of $\mathbf{A}$.

(12.2d) A linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ can be approximately solved by means of the CG method. Estimate the number of CG steps required to reduce the energy norm of the initial error by a factor of 10^8?

(12.2e) Let $\mathbf{x}^*$ be the solution of the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. Assume we approximately solve the system with the Conjugate Gradient algorithm with initial guess $\mathbf{x}^{(0)} = 0$. How many CG steps are needed to achieve

$$\frac{\left\|\mathbf{x}^{(k)} - \mathbf{x}^*\right\|_2}{\left\|\mathbf{x}^*\right\|_2} \leq 10^{-8} \ ?$$

HINT: You may use the fact that, since $\mathbf{A}$ is s.p.d, $\|\mathbf{A}\|_2$ agrees with the maximal eigenvalue of $\mathbf{A}$. Combine this with the estimate

$$\left|\mathbf{x}^\top \mathbf{M}\mathbf{x}\right| \leq \|\mathbf{M}\|_2 \|\mathbf{x}\|_2^2 \qquad \forall \mathbf{M} \in \mathbb{R}^{n,n},\ \mathbf{x} \in \mathbb{R}^n$$

to establish a relationship between the energy norm and the Euclidean norm.

Problem 12.3 Newton meets CG

Computing the Newton correction entails solving a linear system of equations. In light of the fact that in the context of Newton's method we always deal with approximations, it may not be necessary to solve this linear system to very high accuracy. Thus, we may use an iterative solver, for instance a Krylov subspace solver.

In this problem we study a particular non-linear system $F(\mathbf{x}) = 0$ arising from a minimization problem for the convex smooth functional $\Phi : \mathbb{R}^n \mapsto \mathbb{R}$, given by

$$\Phi(\mathbf{x}) := \cosh\left(\mathbf{x}^T \mathbf{A}\mathbf{x}\right) - \mathbf{b}^T\mathbf{x} \quad \text{with} \quad \mathbf{A} = \begin{pmatrix} 4 & 1 & & \\ 1 & 4 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 4 \end{pmatrix} \in \mathbb{R}^{n,n}, \ \ \mathbf{b} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n \ .$$

A necessary condition for the existence of a minimizer $\mathbf{x}^*$ of Φ is $\operatorname{grad}\Phi(\mathbf{x}^*) = 0$. Thanks to the positive definiteness of $\mathbf{A}$ and the convexity of $\cosh$, the condition is also sufficient. The minimization problem is therefore equivalent to a non-linear system of equations.

(12.3a) Display in a surface plot the function $\Phi(\mathbf{x})$ for $n = 2$ and $\mathbf{x} \in [-0.4, 0.4] \times [-0.4, 0.4]$.

(12.3b) Show that $\mathbf{A}$ is s.p.d.

(12.3c) Compute $\mathbf{F}(\mathbf{x}) := \operatorname{grad}\Phi(\mathbf{x})$ using the chain rule for differentiation in multiple dimensions.

(12.3d) Compute the Hessian $\mathbf{H}\Phi$ of Φ, i.e, the matrix of the second derivatives. It coincides with the Jacobian $D\mathbf{F}$ of $\mathbf{F}$.

HINT: If you do not feel comfortable with the product rule for multidimensional differentiation, consider the components of $\mathbf{F}$ and write all the expressions using sums.

(12.3e) Show that $D\mathbf{F}(\mathbf{x})$ is s.p.d. for any $\mathbf{x} \in \mathbb{R}^n,\ \mathbf{x} \neq 0$.

(12.3f) The original matrix $\mathbf{A}$ is sparse. Conversely, we find that $D\mathbf{F}(\mathbf{x})$ is dense. However, it arises from a special modification of a sparse matrix. What kind of modification?

(12.3g) Determine the computational effort needed to compute $D\mathbf{F}(\mathbf{x}) \cdot \mathbf{y}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

(12.3h) We have transformed the minimization problem

$$\mathbf{x} = \operatorname*{arg\,min}_{\mathbf{y} \in \mathbb{R}^n} \Phi(\mathbf{y})$$

in the non-linear root-finding problem

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}.$$

Write the formula for the Newton iteration corresponding to this system.

(12.3i) Write a MATLAB function `minPhi(x0)` that solves the non-linear system of equations $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ by means of Newton's method. The argument `x0` is supposed to pass the initial guess. Here the Newton correction may be computed using the MATLAB \-solver.

(12.3j) Use your `minPhi`-code to find $\mathbf{x}^* = \operatorname*{argmin}_{\mathbf{x} \in \mathbb{R}^n} F(\mathbf{x})$ up to a relative error of 10^{-6}. What is the problem when you start with $\mathbf{x}^{(0)} = \mathbf{0}$?

(12.3k) Now implement `minPhi` in such a way that the Newton corrections are (approximately) determined by means of conjugate gradient iterations. Use the MATLAB `pcg`-function.

Compare the convergence of this new "inexact" Newton iteration for different CG-tolerances `tol` with that of the "exact" Newton iteration from subproblem (12.3j) for $100 \leq n \leq 300$.

Published on May 20, 2014.

To be submitted on May 27/28, 2014.

MATLAB: Submit all files in the online system. Include the files that generate the plots. Label all your plots. Include commands to run your functions. Comment on your results.

Written exercises: Hand-in the solutions during the exercise class or in the labeled boxes in HG G 53.x.

References

[NMI] Lecture Notes for the course "Numerical Analysis I".

Last modified on May 20, 2014