Numerical Linear Algebra Assignment 12

Exercise 1. (10 points)

Let $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{b} \in \mathbb{R}^m$, and $\phi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} - \mathbf{x}^\top \mathbf{b}$. Compute the gradient of $\phi(\mathbf{x})$.

Exercise 2. (TreBau Exercise 38.5, 10 points)

We have described CG as an iterative minimization of the function $\varphi(\mathbf{x})$ of (38.7). Another way to minimize the same function—far slower, in general—is by the method of *steepest descent*.

(a) Derive the formula $\nabla\varphi(\mathbf{x}) = -\mathbf{r}$ for the gradient of $\varphi(\mathbf{x})$. Thus the steepest descent iteration corresponds to the choice $\mathbf{p}_k = \mathbf{r}_k$ instead of $\mathbf{p}_k = \mathbf{r}_k + \beta_k \mathbf{p}_{k-1}$ in Algorithm 38.1.

(b) Determine the formula for the optimal step length α_k of the steepest descent iteration.

(c) Write down the full steepest descent iteration. There are three operations inside the main loop.

Exercise 3. (Shufang Xu, 10 points)

Suppose that the steepest descent iteration finds the exact solution of a symmetric positive definite linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ within finite steps. Prove that the last search direction is an eigenvector of $\mathbf{A}$.

Exercise 4. (10 points)

Prove the following: if the vector $\mathbf{b}$ is orthogonal to n linearly independent eigenvectors of the HPD matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$, then the CG iteration for $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $\mathbf{x}_0 = \mathbf{0}$ converges in at most $m - n$ steps.

Exercise 5. (10 points)

Assume the HPD matrix $\mathbf{A}$ has only n distinct eigenvalues $\{\lambda_i\}_{i=1}^n$ and the vector $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ is non-degenerate (i.e., the projections onto each of the n eigenspaces, $\mathrm{null}(\mathbf{A} - \lambda_i \mathbf{I})$, are nonzero). Prove the following: CG converges at step n and the polynomial $p(\lambda)$ in $\mathbf{x}_n - \mathbf{x}_0 = p(\mathbf{A})\mathbf{r}_0$ satisfies $p(\lambda_i) = 1/\lambda_i$ and $p(\mathbf{A}) = \mathbf{A}^{-1}$.

Exercise 6. (10 points)

Assume that $\mathbf{A}$ and $\mathbf{M}$ are both HPD. Prove that PCG for $\mathbf{A}\mathbf{M}^{-1}\mathbf{z} = \mathbf{b}$, $\mathbf{x} = \mathbf{M}^{-1}\mathbf{z}$ has the convergence estimate

$$\frac{\|\boldsymbol{\varepsilon}_j\|_{\mathbf{A}}}{\|\boldsymbol{\varepsilon}_0\|_{\mathbf{A}}} \le 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^j.$$

Here $\boldsymbol{\varepsilon}_j = \mathbf{A}^{-1}\mathbf{b} - \mathbf{x}_j$ and $\kappa = \lambda_{\max}(\mathbf{A}\mathbf{M}^{-1})/\lambda_{\min}(\mathbf{A}\mathbf{M}^{-1})$.

Exercise 7. (Programming, 10 points)

TreBau Exercise 38.6. Also plot the curve: $(\kappa-1)^n/(\kappa+1)^n$.

Exercise 8. (Programming, 10 points)

Write matlab code to plot Figure 40.1 of TreBau's book.

Additional Exercise 1.

How to recover the tridiagonal matrix $\mathbf{T}_j$ generated in the Lanczos process via α_i, β_i in CG, and $\|\mathbf{r}_i\|_2$, $i = 1, 2, \ldots, j$.