problem sheet 13

discussion: week of Monday, 23.1.2020

13.1. Let $\mathbf{A} \in \mathbb{R}^{N\times N}$ be positive definite (but not necessarily be symmetric). The *minimal residual method* defines $\mathbf{x}_{\ell+1}$ from $\mathbf{x}_\ell$ by minimizing the function $\phi(\mathbf{x}) := \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ on the line $\{\mathbf{x}_\ell + t\mathbf{r}_\ell \,|\, t \in \mathbb{R}\}$ with $\mathbf{r}_\ell = \mathbf{A}\mathbf{x}_\ell - \mathbf{b}$.

a) Formulate the method, i.e., compute α_ℓ such that α_ℓ is the minimizer of

$$t \mapsto \|\mathbf{A}(\mathbf{x}_\ell + t\mathbf{r}_\ell)\|_2^2.$$

b) Define

$$\mu := \frac{1}{2}\lambda_{min}(\mathbf{A} + \mathbf{A}^T), \qquad \sigma := \|\mathbf{A}\|_2,$$

where $\lambda_{min}(\mathbf{A} + \mathbf{A}^T)$ is the smallest eigenvalue of $\mathbf{A} + \mathbf{A}^T$. Show: $\mu \leq \sigma$.

c) Shat that the *minimal residual method* converges by showing

$$\|\mathbf{r}_{\ell+1}\|_2 \leq \left(1 - \frac{\mu^2}{\sigma^2}\right)\|\mathbf{r}_\ell\|_2^2$$

13.2. *Steepest descent* could be employed for SPD matrices $\mathbf{A}$. Define for the exact solution $\mathbf{x}^*$ of $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ the function $\phi(\mathbf{x}) := \frac{1}{2}(\mathbf{A}\mathbf{x}, \mathbf{x}) - (\mathbf{b}, \mathbf{x})$.

a) Show: $\phi(\mathbf{x}) - \phi(\mathbf{x}^*) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}^*\|_{\mathbf{A}}^2$.

b) Show that the direction of steepest descent at $\mathbf{x}$, which is given by $-\nabla\phi(\mathbf{x})$, is

$$\nabla\Phi(\mathbf{x}) = -\mathbf{r}(\mathbf{x}) := \mathbf{A}\mathbf{x} - \mathbf{b}.$$

c) *Steepest descent* computes from an approximation $\mathbf{x}_\ell$ the new approximation $\mathbf{x}_{\ell+1} = \mathbf{x}_\ell + \alpha_\ell\mathbf{r}_\ell$ (with $\mathbf{r}_\ell = \mathbf{r}(\mathbf{x}_\ell)$), where α_ℓ is such that $\mathbf{x}_{\ell+1}$ minimizes the function $\mathbf{x} \mapsto \phi(\mathbf{x})$ on the line $\{\mathbf{x}_\ell + t\mathbf{r}_\ell \,|\, t \in \mathbb{R}\}$. compute α_ℓ. Formulate the *steepest descent* algorithm.

d) Show that the search directions $\mathbf{r}_{\ell+1}$ and $\mathbf{r}_\ell$ are orthogonal.
Remark: 1. The convergence of the *steepest descent* method is

$$\|\mathbf{x}_{\ell+1} - \mathbf{x}^*\|_{\mathbf{A}} \leq \frac{\kappa(\mathbf{A}) - 1}{\kappa(\mathbf{A}) + 1}\|\mathbf{x}_\ell - \mathbf{x}^*\|_{\mathbf{A}}, \qquad \kappa(\mathbf{A}) = \frac{\lambda_{max}(\mathbf{A})}{\lambda_{min}(\mathbf{A})}$$

2. Also CG can be viewed as a descent method. Then, the search direction $\mathbf{d}_\ell$ is *not* $\mathbf{r}_\ell$. Rather, CG determines search directions $\mathbf{d}_0, \ldots, \mathbf{d}_\ell$ that are (pairwise) $\mathbf{A}$-orthogonal. The convergence of CG is

$$\|\mathbf{x}_\ell - \mathbf{x}^*\|_{\mathbf{A}} \leq 2\left(\frac{\sqrt{\kappa(\mathbf{A})} - 1}{\sqrt{\kappa(\mathbf{A})} + 1}\right)^\ell \|\mathbf{x}_0 - \mathbf{x}^*\|_{\mathbf{A}},$$

which is much better than *steepest descent* for the typical case of large $\kappa(\mathbf{A})$.

13.3. Program for SPD-matrices $\mathbf{A} \in \mathbb{R}^{N\times N}$ the *steepest descent* method of Problem 13.2. The routine should return two vectors: a vector with the ℓ^2-norms of the residuals, $\|\mathbf{b} - \mathbf{A}\mathbf{x}_\ell\|_2$, and a vector with the errors $\|\mathbf{x}^* - \mathbf{x}_\ell\|_{\mathbf{A}}$. (The exact solution $\mathbf{x}^*$ can be computed with `matlab`-command `\`.) Compare this method with the CG-method (see homepage) by plotting semilogarithmically (`semilogy`) the error and the norm of the residual versus the iteration number $\ell \in \{1, \ldots, N\}$. Use the matrices $\mathbf{A} = \texttt{gallery('poisson',n)}$ for $n = 10$, 20, 40 and $\mathbf{b} = \texttt{ones}(n * n, 1)$. Which method is better? Explain.

Remark: The matrices `gallery('poisson',n)` are $n^2 \times n^2$-matrices that correspond to a discretization of the differential operator $-\Delta$ on a regular grid with n points in each direction.