EXERCISE SET 9 – MATH-250 Advanced Numerical Analysis I

The exercise sheet is divided into two sections: quiz and exercises. The quiz will be discussed in the beginning of the lecture on Thursday, May 8. The exercises marked with (⋆) are graded homework. The exercises marked with **(Python)** are implementation based and can be solved in the Jupyter notebooks which are available on Moodle/Noto. **The deadline for submitting your solutions to the homework is Friday, May 9 at 10h15.**

Quiz

(a) For *any* invertible matrix A, right-hand side $\mathbf{b}$, and starting vector $\mathbf{x}_0$, there is a choice of α such that the Richardson method converges.

☐ True ☐ False

(b) Consider a family of linear systems

$$A_n\mathbf{x} = \mathbf{b}_n, \quad A_n \in \mathbb{R}^{n\times n},$$

such that

- A_n is symmetric positive definite;
- $\kappa_2(A_n) = \|A_n\|_2\|A_n^{-1}\|_2 = O(n^2)$ for $n \to \infty$;
- $\|\mathbf{x}\|_2 = 1$.

Consider fixed accuracy $\varepsilon > 0$. Let k_n denote the minimal number of iterations of the Richardson method (with optimal α, zero starting vector, no preconditioner) needed to attain $\|\mathbf{x}_{k_n} - \mathbf{x}\|_2 \leq \varepsilon$. Then for $n \to \infty$ it holds that

☐ $k_n = O(1)$ ☐ $k_n = O(n)$

☐ $k_n = O(\log n)$ ☐ $k_n = O(n^2)$

(c) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on $\mathbb{R}^n$. If $\mathbf{x}$ is a minimum of f then $\nabla f(\mathbf{x}) = 0$.

☐ True ☐ False

(d) Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable on $\mathbb{R}^n$ and $\mathbf{x}$ such that $\nabla f(\mathbf{x}) \neq 0$. Then for every $\varepsilon > 0$ there is $\mathbf{y}$ with $\|\mathbf{y} - \mathbf{x}\| \leq \varepsilon$ and $f(\mathbf{y}) < f(\mathbf{x})$.

☐ True ☐ False

Exercises

If you have skipped Problem 2 in Exercise Set 8, make sure to catch up on it this week.

Problem 1.

The aim of this exercise is to prove for $A \in \mathbb{R}^{n\times n}$ we have

$$A^k \to 0 \text{ as } k \to \infty \Leftrightarrow \rho(A) < 1,$$

where $\rho(A)$ denotes the spectral radius of A. Let $\|\cdot\|_2$ denote the spectral norm (also called matrix 2-norm).

(a) Show $\rho(A)^k \leq \|A^k\|_2$.
Hint: Use $A\mathbf{x} = \lambda\mathbf{x}$ and submultiplicativity.

(b) Using (a), show

$$A^k \to 0 \text{ as } k \to \infty \Rightarrow \rho(A) < 1$$

(c) Consider the $m \times m$ Jordan block

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & & & 0 \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

Show that

$$(J_m(\lambda)^k)_{ij} = \begin{cases} 0 & \textit{if } i > j \\ \lambda^k & \text{if } i = j \\ \binom{k}{l}\lambda^{k-l} & \text{if } j = i + l \end{cases}$$

where we let $\binom{k}{l} = 0$ if $l > k$.

Hint: First show what happens when $\lambda = 0$. Then use $J_m(\lambda)^k = (\lambda I_m + J_m(0))^k$.

(d) Show that if $|\lambda| < 1$ then

$$J_m(\lambda)^k \to 0 \text{ as } k \to \infty$$

(e) By considering the Jordan canonical form $A = PJP^{-1}$ show

$$A^k \to 0 \text{ as } k \to \infty \Leftarrow \rho(A) < 1$$

Problem 2.

Consider the linear system $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

(a) Determine if Jacobi's method is guaranteed to converge.

(b) Consider the following iterative method

$$L\mathbf{x}^{(k+1)} = L\mathbf{x}^{(k)} + \delta(\mathbf{b} - A\mathbf{x}^{(k)}) \quad k \geq 0 \tag{1}$$

where $\delta > 0$ is a parameter and

$$L := \begin{pmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & -1 & 2 \end{pmatrix}$$

Rewrite the method (1) in the form $\mathbf{x}^{(k+1)} = B^{\delta}\mathbf{x}^{(k)} + \mathbf{z}_{\delta}, k \geq 0$, for a suitable matrix B^{δ} which is to be determined.

(c) Establish for which values of the parameter $\delta > 0$ the method (1) converges.

(d) Let $\delta = \frac{4}{3}$. Considering the results obtained at the point (c), establish whether the method (1) is convergent. If so, which method can be expected to converge faster between method (1) and Jacobi?

Problem 3. Consider the linear system $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

(a) We want to solve the system by the Gauss-Seidel method. Determine the iteration matrix B^{GS}.

(b) What can we say about the convergence of the Gauss-Seidel method?

(c) We now consider the preconditioned Richardson method

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha P^{-1}\mathbf{r}^{(k)} \quad k \geq 0$$

with $P = D$ where D is the diagonal part of A. Verify that if we take $\alpha = 1$ we find the Jacobi method.

(d) Let the starting vector be $\mathbf{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and calculate the first iteration of the preconditioned Richardson method, with $P = D$ being the diagonal part of A, by choosing the optimal parameter α_{opt}.

Problem 4.

The aim of this exercise is to prove that the iterates of the Gauss-Seidel method applied to a strictly diagonally dominant matrix A converge to the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$.

(a) Recall that the error of the Gauss-Seidel iteration can be written as

$$\mathbf{e}_i^{(k+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} \mathbf{e}_j^{(k+1)} - \sum_{j=i+1}^{n} \frac{a_{ij}}{a_{ii}} \mathbf{e}_j^{(k)}$$

where $\mathbf{e}^{(k)} := \mathbf{x}^{(k)} - \mathbf{x}$.Using this, show that there exists an index $p \in \{1, \cdots, n\}$ such that

$$\left(1 - \sum_{j=1}^{p-1} \left|\frac{a_{pj}}{a_{pp}}\right|\right) \|\mathbf{e}^{(k+1)}\|_\infty \leq \left(\sum_{j=p+1}^{n} \left|\frac{a_{pj}}{a_{pp}}\right|\right) \|\mathbf{e}^{(k)}\|_\infty$$

(b) Using that A is strictly diagonally dominant, show that there exists some $\beta \in (0,1)$ such that

$$\|\mathbf{e}^{(k+1)}\|_\infty \leq \beta \|\mathbf{e}^{(k)}\|_\infty$$

and conclude that the Gauss-Seidel method converges.

($\star$) Problem 5.

(a) Consider two symmetric matrices A and P. Show that if P is also positive definite, then $P^{-1}A$ is diagonalisable and all its eigenvalues are real.

(b) **Solving this part is optional and will not be graded.**

Suppose that all eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n > 0$ of A are real and that A satisfies

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \leq \gamma |a_{ii}|, i = 1, 2 \ldots, n, \tag{2}$$

for some $\gamma \in (0,1)$. Using $a_+ = \max_{i=1,2,\ldots,n} |a_{ii}|$ and $a_- = \min_{i=1,2,\ldots,n} |a_{ii}|$ show that

$$\frac{\lambda_1}{\lambda_n} \leq \frac{1+\gamma}{1-\gamma} \cdot \frac{a_+}{a_-}.$$

Hint: You may use Gershgorin's circle theorem.

Gershgorin's Circle Theorem. We define

$$R_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \quad \text{and} \quad B_i = B(a_{ii}, R_i) \subset \mathbb{C},$$

where B_i is the open complex ball with center $a_i i$ and radius R_i. Then, any eigenvalue λ lies within at least one B_i.

(c) Let A be a symmetric and positive definite matrix satisfying (2) with $\gamma = 0.9$. Use (b) to show that the preconditioned Richardson with the diagonal preconditioner $P = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ converges at a rate ≤ 0.9.

(d) Write a Python function `jacobi(A, b, x0, tol, kmax)` that implements the Jacobi method. Choose the right-hand side b as a random vector such that $b_i \sim \mathcal{N}_{0,1}$ for $i = 1, 2, \ldots, n$ follows the standard normal distribution. To this end, define $b =$ `np.random.randn(n)`, or use `numpy.random.randn` if you do not want to use the Jupyter notebooks we provided on Moodle. Run the Jacobi method for

$$A_1 = \begin{pmatrix} 9 & -4 & 0 \\ -4 & 9 & -4 \\ 0 & -4 & 9 \end{pmatrix}$$

and the 100000×100000 matrix A_2 that we provide on Moodle in the file `matrix.npz`. Plot the 2-norm of the residual vector $\|r^{(k)}\|_2 = \|Ax^{(k)} - b\|_2$ for the Jacobi iterate $x^{(k)}$ and increasing numbers of iteration k.

Hint: From SciPy's `sparse` submodule use the function `load_npz` function. In the Jupyter notebook provided on Moodle you can directly call `sps.load_npz`, otherwise you will have to use `scipy.sparse.load_npz`. Look at the function signature of `richardson` we provided in the Jupyter notebook on Moodle as well as the helper functions it contains to handle dense and sparse matrices at the same time.

(e) Write a Python function `richardson(A, b, x0, alpha, P, tol, kmax)` that implements the Richardson method without preconditioning and with diagonal preconditioning (use (c)), respectively. Plot the norms of the residuals $\|r^{(k)}\|_2 = \|Ax^{(k)} - b\|_2$ for the output of both functions for increasing numbers of iteration k. You may choose the Richardson iteration's parameter as $\alpha = \frac{1.9}{\|P^{-1}A\|_\infty}$, where $P = \text{id}$ in case no preconditioning is used.

Hint: Look at the function signature of `richardson` we provided in the Jupyter notebook on Moodle as well as the helper functions it contains to handle dense and sparse matrices at the same time.