Solution 9 – MATH-250 Advanced Numerical Analysis I

($\star$) 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

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

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 (1) 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 $P = \mathrm{id}$ 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.

Solution.

(a) By assumption, P is symmetric and positive definite. Therefore, $P^{-1/2}$ exists and is invertible. Thus it holds that

$$P^{1/2}(P^{-1}A)P^{-1/2} = P^{-1/2}AP^{-1/2}, \text{ and}$$
$$\left(P^{-1/2}AP^{-1/2}\right)^\top = \left(P^\top\right)^{-1/2}A^\top\left(P^\top\right)^{-1/2} = P^{-1/2}AP^{-1/2}.$$

We have shown that $P^{-1/2}AP^{-1/2}$ is symmetric, wherefore all its eigenvalues are real and it is diagonalisable. Lastly, $P^{-1}A$ is similar to $P^{-1/2}AP^{-1/2}$, meaning that its eigenvalues are real and it too is diagonalisable.

(b) We directly use the notation as given in the task. By the Gershgorin Circle Theorem we know that for every eigenvalue λ_k there exists an open ball B_i such that $\lambda_k \in B_i$.

We show that $a_{ii} > 0$: By the inclusion $\lambda_k \in B_i$ it is evident that

$$|a_{ii} - \lambda_k| < R_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \leq \gamma|a_{ii}| < |a_{ii}|.$$

By assumption we know $\lambda_k > 0$, and hence $a_{ii} > 0$.

We further manipulate $|a_{ii} - \lambda_k| \leq \gamma|a_{ii}|$ and use $a_{ii} > 0$ to see that

$$(1-\gamma)|a_{ii}| = a_{ii} - \gamma|a_{ii}| \leq \lambda_k \leq a_{ii} + \gamma|a_{ii}| = (1+\gamma)|a_{ii}|.$$

This holds irrespective of the eigenvalue λ_k we selected initially, meaning that for all $k = 1, 2, \ldots, n$ we have

$$(1-\gamma) \min_{i=1,2,\ldots,n} |a_{ii}| \leq \lambda_k \leq (1+\gamma) \max_{i=1,2,\ldots,n} |a_{ii}|.$$

Dividing this inequality for λ_1 by that for λ_n we finally get our desired result

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

(c) The matrix P is invertible because none of the main diagonal entries are zero, implying by (b) that none of the eigenvalues may be zero. The matrix A is symmetric and positive definite, whence all its diagonal entries are positive, which also holds true for

P. By applying (a) we see that all eigenvalues of $P^{-1}A$ are real and positive. Next, we can write every ij-th entry of $P^{-1}A$ because

$$(P^{-1}A)_{ij} = \frac{a_{ij}}{a_{ii}}.$$

This way of writing implies that $P^{-1}A$ satisfies (1). Labelling the eigenvalues of $P^{-1}A$ by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ we can also apply (b) yielding

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

Seeing that the rate of convergence is linked to the spectral radius of the iteration matrix $P^{-1}A$ we argue with Theorem 5.4 and see

$$\rho_{\text{opt}} = \frac{\frac{\lambda_1}{\lambda_n} - 1}{\frac{\lambda_1}{\lambda_n} + 1} \leq \frac{\frac{1+\gamma}{1-\gamma} - 1}{\frac{1+\gamma}{1-\gamma} + 1} = \gamma = 0.9.$$

(d, e) You can find the solution in the Jupyter notebook provided on Moodle.