MA398 Matrix Analysis and Algorithms: Exercise Sheet 9

1. (SD example) We consider the steepest decent method applied to the following data: For a real $a \gg 1$
$$A = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x_0 = \begin{pmatrix} a \\ 1 \end{pmatrix}$$

   (a) Show that the even iterates are
$$x^{(2m)} = \Big(\frac{a-1}{a+1}\Big)^{2m} x_0, \quad m \in \mathbb{N},$$
and find a formula for the odd iterates.
Compute the $\alpha^{(k)}$ and the residuals $r^{(k)}$, $k \in \mathbb{N}$.

   (b) Show that subsequent search directions are 'almost' parallel in the metric defined by A: Use the formula
$$\cos(\phi) = \frac{\langle x, y\rangle_A}{\|x\|_A \|y\|_A}$$
to measure the angle ϕ between to subsequent search directions and check that ϕ tends to π as $a \to \infty$.

Remark: In contrast, the search directions in the CG method are orthogonal with respect to $\langle\cdot,\cdot\rangle_A$ which results in far better convergence properties.

Answer:

   (a) By induction, and the formula clearly is true in the case $m = 0$.
For any $m \in \mathbb{N}$ we have that
$$r^{(2m)} = b - Ax^{(2m)} = -\Big(\frac{a-1}{a+1}\Big)^{2m} \begin{pmatrix} a \\ a \end{pmatrix}, \quad Ar^{(2m)} = -\Big(\frac{a-1}{a+1}\Big)^{2m} \begin{pmatrix} a \\ a^2 \end{pmatrix}, \tag{1}$$
which yields
$$\alpha^{(2m)} = \frac{\|r^{(2m)}\|_2^2}{\|r^{(2m)}\|_A^2} = \Big(\big(\tfrac{a-1}{a+1}\big)^{4m} 2a^2\Big) \Big/ \Big(\big(\tfrac{a-1}{a+1}\big)^{4m}(a^2 + a^3)\Big) = \frac{2}{1+a}.$$
Since
$$\begin{aligned} a - \frac{2}{a+1}a &= \frac{a^2 + a - 2a}{a+1} = a\frac{a-1}{a+1}, \\ 1 - \frac{2}{a+1}a &= \frac{a+1-2a}{a+1} = -\frac{a-1}{a+1} \end{aligned} \tag{2}$$
we obtain
$$x^{(2m+1)} = x^{(2m)} + \alpha^{(2m)} r^{(2m)} = \Big(\frac{a-1}{a+1}\Big)^{2m}\Big(\begin{pmatrix} a \\ 1 \end{pmatrix} + \frac{2}{1+a}\begin{pmatrix} -a \\ -a \end{pmatrix}\Big) = \Big(\frac{a-1}{a+1}\Big)^{2m+1}\begin{pmatrix} a \\ -1 \end{pmatrix}. \tag{3}$$
In the next step the residual is
$$r^{(2m+1)} = b - Ax^{(2m+1)} = -\Big(\frac{a-1}{a+1}\Big)^{2m+1} \begin{pmatrix} a \\ -a \end{pmatrix}, \quad Ar^{(2m)} = -\Big(\frac{a-1}{a+1}\Big)^{2m+1} \begin{pmatrix} a \\ -a^2 \end{pmatrix}, \tag{4}$$

but the step size is the same again:

$$\alpha^{(2m+1)} = \frac{\|r^{(2m+1)}\|_2^2}{\|r^{(2m+1)}\|_A^2} = \left(\left(\tfrac{a-1}{a+1}\right)^{4m+2} 2a^2\right) \Big/ \left(\left(\tfrac{a-1}{a+1}\right)^{4m+2} (a^2+a^3)\right) = \frac{2}{1+a}.$$

Moreover, by

$$-1 + \frac{2}{a+1} a = \frac{-a-1+2a}{a+1} = \frac{a-1}{a+1}$$

and using (2) we see that

$$x^{(2m+2)} = \left(\frac{a-1}{a+1}\right)^{2m+1} \left(\begin{pmatrix} a \\ -1 \end{pmatrix} + \frac{2}{1+a} \begin{pmatrix} -a \\ a \end{pmatrix} \right) = \left(\frac{a-1}{a+1}\right)^{2m+2} \begin{pmatrix} a \\ 1 \end{pmatrix}$$

which shows that the formula for the even $x^{(k)}$ is true. The odd $x^{(k)}$ fulfil (3). The step sizes are constant, $\alpha^{(k)} = \frac{2}{a+1}$ for all k, and the residuals (= search directions) are (1) for k even and (4) for k odd.

(b) First, using (1) and (4) we have that

$$\langle r^{(2m)}, r^{(2m+1)} \rangle_A = \langle r^{(2m)}, Ar^{(2m+1)} \rangle_2 = \left(\frac{a-1}{a+1}\right)^{4m+1} (a^2 - a^3).$$

Furthermore, as already seen in the computation of the $\alpha^{(k)}$

$$\|r^{(2m)}\|_A = \left(\left(\tfrac{a-1}{a+1}\right)^{4m} (a^2+a^3)\right)^{\frac{1}{2}} = \left(\frac{a-1}{a+1}\right)^{2m} \sqrt{a^2+a^3},$$
$$\|r^{(2m+1)}\|_A = \left(\left(\tfrac{a-1}{a+1}\right)^{4m+2} (a^2+a^3)\right)^{\frac{1}{2}} = \left(\frac{a-1}{a+1}\right)^{2m+1} \sqrt{a^2+a^3}.$$

This yields for the angle ϕ between $r^{(2m)}$ and $r^{(2m+1)}$

$$\cos(\phi) \frac{\langle r^{(2m)}, r^{(2m+1)} \rangle_A}{\|r^{(2m)}\|_A \|r^{(2m+1)}\|_A} = \frac{a^2 - a^3}{a^2 + a^3} = \frac{1-a}{1+a} \to -1 \text{ as } a \to \infty.$$

Therefore $\phi \to \pi$ as $a \to \infty$.

2. Implement the Conjugate Gradient Method for solving linear systems. Given a positive definite matrix A and a vector b, the task is to find a vector x such that $Ax = b$. Use the pseudo code provided in the lecture notes as the base for your implementation.

   The function conjugate_gradient(A, b, x0, eps) takes as input a positive definite matrix A, a vector b, an initial guess for the solution x0, and a tolerance eps. It returns the solution to the linear system $Ax = b$ found by the Conjugate Gradient Method.

   Note: Be sure to implement checks for whether A is symmetric and positive definite. In practice, the Conjugate Gradient Method is usually applied to large, sparse systems, and more efficient implementations would take advantage of this.

3. Implement a Python function that computes the Krylov subspace $\mathcal{K}_k(r^{(0)}, A)$ for a given matrix A and vector $r^{(0)}$, where k is the number of iterations. The function should return a list of vectors spanning the Krylov subspace. You can use the numpy library for computations involving vectors and matrices. Here is a template to start with:

```
import numpy as np

def compute_krylov_subspace(r0, A, k):
    # Your code here
    pass
```

Test your function with a random symmetric positive-definite matrix A and a random vector $r^{(0)}$, and for a few different values of k. Verify that the dimension of the Krylov subspace is as expected. You can use the numpy.linalg.matrix_rank function to compute the dimension of the subspace spanned by a list of vectors.