Problem 1  Given the matrix

\[ A = \frac{1}{21} \begin{bmatrix} -9 & 32 & -62 \\ -72 & 67 & -34 \\ -18 & 106 & 2 \end{bmatrix} . \]

a) Fill in \( \mu_i, \nu_i, i = 1, 2, 3 \) and \( \sigma_3 \) such that the product

\[ A = \begin{bmatrix} 1/3 & -2/3 & \mu_1 \\ 2/3 & -1/3 & \mu_2 \\ 2/3 & 2/3 & \mu_3 \end{bmatrix} \begin{bmatrix} 7 & \sigma_3 \\ 3 & \nu_1 \\ \nu_2 & \nu_3 \end{bmatrix} \begin{bmatrix} -3/7 & 6/7 & -2/7 \\ 2/7 & 3/7 & 6/7 \end{bmatrix} \]

is a singular value decomposition of \( A \).

b) We define the set of matrices

\[ \mathcal{M} = \{ a_1 b_1^T + a_2 b_2^T, a_1, b_1, a_2, b_2 \in \mathbb{R}^3 \} \]

Determine

\[ \tilde{A} = \arg \min_{B \in \mathcal{M}} \| A - B \|_2 \]

where \( A \) is the matrix defined above.
Problem 2  Let us define the shift matrix \( S \in \mathbb{R}^{n \times n} \) as
\[
S = \begin{bmatrix}
1 & 1 \\
1 & \ddots \\
& & \ddots \\
& & & 1
\end{bmatrix}
\]
the matrix with 1 on the subdiagonal and upper right corner and 0 elsewhere. The effect of applying this matrix to a vector is that all components are shifted one position down and the last component is shifted to the first. Clearly, \( S \) is orthogonal and so \( S^{-1} = S^T \) and solving problems \( Sx = b \) is trivial. Nevertheless, we shall use this linear system as a test case for Krylov subspace methods.

a) Prove that the eigenvalues of \( S \) are the \( n \)th roots of unity, i.e.
\[
\lambda_k = e^{\frac{2i\pi k}{n}}, \quad k = 1, \ldots, n, \quad (i = \sqrt{-1}).
\]

b) Let \( v_1 = e_1 = [1, 0, \ldots, 0]^T \in \mathbb{R}^n \) and for each \( m = 1, \ldots, n \) derive explicitly the matrices \( V_m \) and \( H_m \) from the Arnoldi algorithm, such that the columns of \( V_m \) form an orthonormal basis for \( \mathcal{K}(S, e_1) \).

c) Suppose that we use the GMRES method to solve the linear system \( Sx = b \). We assume that an initial approximation \( x_0 \) has been chosen such that \( r_0 = b - Sx_0 = e_1 \). Compute all approximations \( x_m, m = 1, \ldots, n \). Show how each residual \( r_m \) can be expressed as \( r_m = p_m(S)r_0 \) for some polynomial \( p_m(z) \) of degree at most \( m \), and determine each \( p_m(z) \) for \( m = 1, \ldots, n \). Comment on why the usual convergence analysis presented in the book and lectures fails in this case. Discuss in particular what happens in the very last iteration (\( m = n \)).

d) What happens if we replace GMRES by the full orthogonalization method (FOM).

Problem 3  We now consider the matrix \( A = I + \theta S, |\theta| < 1 \), where \( S \) is the shift matrix defined by (1). You may need the result in the appendix (see below) for this problem.

a) Argue that there exists a diagonal matrix \( \Lambda = \Lambda(\theta) \in \mathbb{C}^{n \times n} \) and a unitary matrix \( X \in \mathbb{C}^{n \times n} \), not depending on \( \theta \), such that \( A = X \Lambda X^H \).

b) We look at solving the equation \( Ax = b \), again by GMRES. Derive an estimate for the convergence of the residual after \( m \) iterations of the form
\[
\|r_m\|_2 \leq \epsilon^{(m)}(\theta) \|r_0\|_2,
\]
that is, determine $\epsilon^{(m)}(\theta)$.

c) Suppose we use a preconditioner, $B^{-1} = I - \theta S$ and consider the system

$$B^{-1}Ax = B^{-1}b$$

Find the corresponding convergence estimate as in (2) obtained by replacing $A$ by $B^{-1}A$.

**Problem 4** Given an arbitrary $2 \times 2$ real symmetric matrix written in the form

$$A = \begin{bmatrix} w + z & \epsilon \\ \epsilon & z \end{bmatrix}.$$

a) Perform the following shifted QR step: $A - zI = QR$, $\bar{A} = RQ + zI$. Show that

$$\bar{A} = \begin{bmatrix} \bar{w} + \bar{z} & \bar{\epsilon} \\ \bar{\epsilon} & \bar{z} \end{bmatrix}, \quad \bar{z} = z - \frac{\epsilon^2 w}{w^2 + \epsilon^2}, \quad \bar{w} = w + 2 \frac{\epsilon^2 w}{w^2 + \epsilon^2}, \quad \bar{\epsilon} = \frac{\epsilon^3}{w^2 + \epsilon^2}.$$

b) What does the result in the previous question tell you about the convergence of the QR-iteration for this type of matrix? What happens to the convergence rate if $w \leq \epsilon$? Draw the Gerschgorin disks for $A$ in the case that $w = \epsilon$ and comment on how this result compare to what you know in general about the convergence of the QR-iteration.

**Appendix.** A version of Zarantonello’s lemma (Saad, Lemma 6.26). Let $C(c, \rho)$ be a circle centered at $c$ with radius $\rho$ where $\rho < |c|$. Then

$$\min_{p \in P_m, \, p(0) = 1} \max_{z \in C(c, \rho)} |p(z)| = \left( \frac{\rho}{|c|} \right)^m$$