1. The Conjugate Gradient Method and its convergence (15 points)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and $b \in \mathbb{R}^n$. Write

$$\mathcal{K}^k(A,b) = \operatorname{span}\{b, Ab, \ldots, A^{k-1}b\}$$

for the k-th Krylov subspace for A with start vector b.

Assume that the only A-invariant subspace that contains b is $\mathbb{R}^n$.

The *Conjugate Gradient Method* defines x_k as the element from $\mathcal{K}^k(A,b)$ with the property that

$$r_k = b - Ax_k \perp \mathcal{K}^k(A,b). \tag{1}$$

(a=3) Show that such x_k exists and is unique for all $k \leq n$.

Define the A-inner product as $\langle x,y\rangle_A = x^\top Ay$ for all $x, y \in \mathbb{R}^n$ and set $\|x\|_A = \langle x,x\rangle_A^{\frac{1}{2}}$.

(b=4) Prove that for all $y \in \mathcal{K}^k(A,b)$,

$$\|x - x_k\|_A \leq \|x - y\|_A. \tag{2}$$

Now, write $\mathbb{R}[X]_{\leq k}$ for the real vector space of polynomials of degree at most k.

(c=4) Prove that for each $y \in \mathcal{K}^k(A,b)$ there exists a $p \in \mathbb{R}[X]_{\leq k}$ with $p(0) = 1$ such that

$$x - y = p(A)x. \tag{3}$$

(d=4) Finally, combine (b) and (c) to show that

$$\frac{\|x - x_k\|_A}{\|x\|_A} \leq \min_{p\in\mathbb{R}[X]_{\leq k},\ p(0)=1} \ \max_{\lambda\in\sigma(A)} |p(\lambda)|. \tag{4}$$

This result enables the use of polynomial approximation theory to further bound the relative errors.

2. Lanczos Biorthogonalization and Biconjugate Gradients (10 points)

Let $A \in \mathbb{R}^{n\times n}$ and $b, c \in \mathbb{R}^n$. Assume that $\dim \mathcal{K}^k(A,b) = k = \dim \mathcal{K}^k(A^\top, c)$ for all $k \leq n$.

The Lanczos biorthogonalization method aims to compute bases $\beta = \{v_1, \ldots, v_k\}$ for $\mathcal{K}^k(A,b)$ and $\gamma = \{w_1, \ldots, w_k\}$ for $\mathcal{K}^k(A^\top, c)$ with the property that

$$v_i^\top w_j = 0 \quad \text{if } i \neq j, \quad \text{and} \quad v_i^\top w_i \neq 0 \tag{5}$$

for all $i \in \{1, \ldots, k\}$. However, this is not always possible.

(a=3) Show by a simple example that the given assumptions do not guarantee existence of β and γ.

Assume from now on that such bi-orthogonal bases β and γ do exist for each k. As usual, write V_k for the matrix with columns $v_1, \ldots, v_k$ and W_k for the matrix with columns $w_1, \ldots, w_k$.

(b=4) Prove that there exist upper Hessenberg matrices $H_{k+1,k}$ and $G_{k+1,k}$ such that for all $k < n$,

$$AV_k = V_{k+1}H_{k+1,k} \quad \text{and} \quad A^\top W_k = W_{k+1}G_{k+1,k}. \tag{6}$$

The *Biconjugate Gradient Method* defines x_k as the element from $\mathcal{K}^k(A,b)$ with the property that

$$r_k = b - Ax_k \perp \mathcal{K}^k(A^\top, c). \tag{7}$$

(c=4) Show that x_k is the solution of a system with matrix $W_k^\top AV_k$, and that $W_k^\top AV_k$ is tridiagonal.

3. The Arnoldi Method and Implicit Restart (15 points)

Let $A \in \mathbb{R}^{n\times n}$ be invertible. Applying an arbitrarily given number $k < n$ steps of the Arnoldi method with start vector v with $\|v\| = 1$ *usually* results in $V_{k+1} \in \mathbb{R}^{n\times(k+1)}$ and $H_{k+1,k} \in \mathbb{R}^{(k+1)\times k}$ such that

$$AV_k = V_{k+1}H_{k+1,k}, \tag{8}$$

where $V_{k+1}^\top V_{k+1} = I$ and $H_{k+1,k}$ is upper Hessenberg.

(a=3) Give an example of such A, v, k so that the Arnoldi method *does not* result in the above.

Suppose now that the Arnoldi method *does not* break down.

Each eigenvalue μ of $H_{k,k}$ (the top k rows of $H_{k+1,k}$) is a *Ritz value* and each corresponding eigenvector y with $\|y\| = 1$ yields a *Ritz vector* $z = V_k y$. The pair (μ, z) is a *Ritz pair* with *residual* $r = Az - \mu z$.

(b=4) Show that each residual r for a Ritz pair is a multiple of one and the same vector, and that

$$\|r\| \leq |h|. \tag{9}$$

Here, h is the entry of $H_{k+1,k}$ at position $(k+1, k)$.

Now, let $H_{k+1,k} =: Q_{k+1,k}R_{k,k}$ be a thin QR-decomposition, i.e., one in which $R_{k,k}$ is square. Set

$$W_k := V_{k+1}Q_{k+1,k} \tag{10}$$

and note that this implies that $W_k R_{k,k} = V_{k+1}H_{k+1,k}$.

(c=4) Prove that the first column of W_k equals $\hat{v} = \pm Av/\|Av\|$.

Let $Q_{k,k-1}$ be the top left $k \times (k-1)$ part of $Q_{k+1,k}$ and W_{k-1} the first $k-1$ columns of W_k.

(d=4) Finally, show that

$$AW_{k-1} = W_k R_{k,k} Q_{k,k-1}, \tag{11}$$

and that $\hat{H}_{k-1,k} := R_{k,k}Q_{k,k-1}$ is upper Hessenberg.

Thus (11) is an Arnoldi decomposition for A with start vector $\hat{v}$.