IMPERIAL

MATH70024

BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
Summer 2025

This paper is also taken for the relevant examination for the Associateship of the Royal College of Science

Computational Linear Algebra

Date: Friday, May 9, 2025

Time: Start time 14:00 – End time 16:00 (BST)

Time Allowed: 2 hours

This paper has 4 Questions.

Please Answer All Questions in 1 Answer Booklet

This is a closed book examination.

Candidates should start their solutions to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Credit will be given for all questions attempted.

Allow margins for marking.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO DO SO

1. (a) The matrix defining a Householder reflection is given by

$$F = \left(I - 2\frac{vv^*}{v^*v}\right). \tag{1}$$

Show that F is a unitary matrix. (5 marks)

(b) When $v = \pm\|x\|e_1 - x$, where $e_1 = (1, 0, \ldots, 0)^T$, show that $Fx = \pm\|x\|e_1$. (5 marks)

(c) Explain why it is best to choose the sign in v to be $-\operatorname{sign}(x_1)$. (5 marks)

(d) The Householder algorithm for upper triangulation of an $m \times n$ matrix A proceeds as follows. We start with $A^{(0)} = A$, and iterate from $k = 1$ to $k = n - 1$. In iteration k, we have

$$A^{(k+1)} = \begin{pmatrix} I_{k-1} & 0 \\ 0 & I_{m-k+1} - 2\frac{v_k v_k^*}{v_k^* v_k} \end{pmatrix} A^{(k)}, \tag{2}$$

where I_r is an r-dimensional identity matrix, $v_k = \operatorname{sign}(x_1^k)\|x^k\|e_1^k - x^k$, with x^k being the last $m - k + 1$ entries in the kth column of $A^{(k)}$, e_1^k being the $m - k + 1$-dimensional unit vector $e_1^k = (1, 0, \ldots, 0)^T$, and x_1^k indicating the first vector entry of x^k. (In iteration $k = 1$, we just have the bottom-right block of this matrix since $m - k + 1 = m$.)

Show that this is equivalent to the algorithm described in pseudocode below.

- FOR $k = 1$ TO $n - 1$
  - $x^k = A_{k:m,k}$
  - $v_k \leftarrow \text{sign}(x_1^k)\|x\|_2 e_1 + x^k$
  - $v_k \leftarrow v_k / \|v_k\|$
  - $A_{k:m,k:n} \leftarrow A_{k:m,k:n} - 2v_k(v_k^* A_{k:m,k:n})$.
- END FOR

(5 marks)

(Total: 20 marks)

2. (a) Given a problem $f : X \to Y$, we are given

1. A floating point number system,
2. An algorithm for computing f in exact arithmetic,
3. A floating point implementation $\tilde{f}$ for f.

(i) Give the definition of stability of the algorithm $\tilde{f}$.

(4 marks)

(ii) Give the definition of backward stability of the algorithm $\tilde{f}$.

(4 marks)

(b) Given an $m \times n$ matrix A, the Householder algorithm for transformation to Hessenberg form finds an upper Hessenberg matrix H such that that $A = Q^*HQ$ for some unitary matrix Q. What does it mean for this algorithm to be backward stable under perturbations to A? (6 marks)

(c) The Arnoldi iteration starts with an arbitrary vector q_1 with $\|q_1\| = 1$. Then,

- For $k = 2, 3, \ldots$,
  - $q_k \leftarrow Aq_{k-1}$
  - For $j = 1, \ldots, k-1$,
    - $h_{j,k-1} = q_j^* q_k$
    - $q_k \leftarrow q_k - h_{j,k-1}q_j$
  - $h_{k,k-1} = \|q_k\|$,
  - $q_k \leftarrow q_k / h_{k,k-1}$.

Assume that A has $\mathcal{O}(m)$ nonzero entries. Estimate the approximate cost in floating point operations of step k of the iteration if this structure is exploited, and the total approximate cost of the first k iterations. (Just compute the leading order behaviour in k and m, and you do not need to compute a constant of proportionality.) (6 marks)

(Total: 20 marks)

3. Let A be a matrix with splitting $A = M + N$, defining a stationary method

$$Mx^{k+1} = -Nx^k + b, \tag{3}$$

to find approximate solutions of the equation $Ax = b$. We have the result that the method converges if and only if the spectral radius $\rho(C) \leq c < 1$, where $C = -M^{-1}N$.

(a) Show that if c is small, then M is a good preconditioner for A, in other words that GMRES will converge quickly to an approximate solution of $Ax = b$ when using M as a preconditioner. (5 marks)

(b) Let $A = L + D + U$ where L is strictly lower triangular, U is strictly upper triangular and D is diagonal. The (forward) Gauss-Seidel splitting takes $M = L + D$. If A is a symmetric positive-definite matrix, show that Gauss-Seidel always converges.
(Hint: you may use the Theorem saying that $M + M^T - A$ being positive definite implies that $\|I - M^{-1}A\|_A < 1$, where $\|x\|_A^2 = x^T Ax$ and we use the corresponding operator norm.) (5 marks)

(c) If A is symmetric, then a symmetrised version of any stationary method takes the form

$$Mx^{k+1/2} = -Nx^k + b, \quad M^T x^{k+1} = -N^T x^{k+1/2} + b. \tag{4}$$

After eliminating $x^{k+1/2}$, this can be written as

$$x^{k+1} = (I - M_s^{-1}A)x^k + M_s b. \tag{5}$$

Find the specific form of M_s for the symmetrised Gauss-Seidel method. (5 marks)

(d) Show that the symmetrised Gauss-Seidel method always converges. You may use the result that if A is symmetric positive definite, then

$$\rho(I - M^{-1}A) = \|I - M^{-1}A\|_A. \tag{6}$$

(5 marks)

(Total: 20 marks)

4. [Mastery Question.]
Consider the QR eigenvalue algorithm for a real symmetric matrix A, as follows. We start with $A^{(0)} = A$. In iteration k, find $Q^{(k)}$ unitary and $R^{(k)}$ upper triangular, such that $A^{(k)} = Q^{(k)}R^{(k)}$, and set $A^{(k+1)} = R^{(k)}Q^{(k)}$.

(a) Two matrices B and C are defined to be "similar" if there exists an invertible matrix X such that $X^{-1}BX = C$. Show that $A^{(k)}$ is similar to A. (5 marks)

(b) Consider the QR eigenvalue algorithm for a symmetric matrix A. After k iterations, we have transformed A to $A^{(k)}$ where

$$A^{(k)} = \begin{pmatrix} * & * & 0 & \dots & 0 & 0 \\ * & * & * & 0 & \dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots & 0 & * & * & 0 \\ 0 & \vdots & 0 & * & * & \epsilon \\ 0 & \vdots & 0 & 0 & \epsilon & \alpha \end{pmatrix}, \tag{7}$$

with $*$ indicating a nonzero matrix entry.
Show that α is an approximate eigenvalue of $A^{(k)}$ in the sense that

$$\|A^{(k)}v - \alpha v\| \leq \epsilon, \tag{8}$$

for some given eigenvector v, which you should provide. (5 marks)

(c) Hence, show that α is an approximate eigenvalue of A, in the sense that

$$\|Aw - \alpha w\| \leq \epsilon, \tag{9}$$

for some given eigenvector w, which you should provide a formula for in terms of the elements of the QR algorithm described above. (5 marks)

[Note: this question is out of 15 to match the scaling of a 3+1 question exam.]

(Total: 15 marks)

BSc and MSci EXAMINATIONS (MATHEMATICS)

May 2025

This paper is also taken for the relevant examination for the Associateship.

M70024

Computational Linear Algebra (Solutions)

Setter's signature	Checker's signature	Editor's signature
....	....	....

1. (a) $\boxed{\text{seen} \Downarrow}$

$$F^*F = (I - 2\frac{vv^*}{v^*v})^*(I - 2\frac{vv^*}{v^*v}) =(I - 2\frac{vv^*}{v^*v})(I - 2\frac{vv^*}{v^*v}) \quad (1)$$
$$=I - 4\frac{vv^*}{v^*v} + 4\frac{vv^*}{v^*v}\frac{vv^*}{v^*v} \quad (2)$$
$$=I - 4\frac{vv^*}{v^*v} + 4v\underbrace{\frac{v^*v}{(v^*v)^2}}_{=\frac{1}{v^*v}}v^* = I, \quad (3)$$

so $F^* = F^{-1}$, i.e. F is unitary. $\boxed{5, \text{A}}$

(b) $\boxed{\text{seen} \Downarrow}$

$$v^*v =(\pm\|x\|e_1 - x)^*(\pm\|x\|e_1 - x) = \|x\|^2 \mp 2\|x\|x_1 + \|x\|^2 \quad (4)$$
$$=2(\|x\|^2 \mp \|x\|x_1) = -2(\pm\|x\|x_1 - \|x\|^2), \quad (5)$$
$$\text{and } v^*x =(\pm\|x\|e_1 - x)^*x = \pm\|x\|x_1 - \|x\|^2. \quad (6)$$

Then,

$$Fx = x - 2v\frac{v * x}{v^*v},$$
$$= x - 2(\pm\|x\|e_1 - x)\frac{\pm\|x\|x_1 - \|x\|^2}{-2(\pm\|x\|x_1 - \|x\|^2)} = x + \pm\|x\|e_1 - x = \pm\|x\|e_1.$$

$\boxed{5, \text{A}}$

(c) If $x \approx \|x\|e_1$ (i.e. x is nearly colinear with e_1) then $x - \|x\|e_1$ is very close to zero and there is a risk of roundoff error when $\|x\|$ is large. Then, the formula is badly conditioned because small changes in x will lead to large changes in the output relative to the value. Similarly, if $x \approx -\|x\|e_1$ then there are similar issues with $x + \|x\|e_1$. Hence, it is best to choose $v = \pm\|x\|e_1 - x$. $\boxed{\text{seen} \Downarrow}$

$\boxed{5, \text{C}}$

(d) Each iteration in the loop replaces $A(k)$ with $A^{(k+1)}$ by overwriting. Inside the k loop, the first step extracts x^k (the last $m - k - 1$ entries in the kth column of $A^{(k)}$. The second step defines $-v_k$, but the sign doesn't matter because we are going to square it. If we replace v_k with $v_k/\|v_k\|$, then $F_k = I_{m-k} - 2vv^*/(v^*v)$ becomes $I_{m-k} - 2vv^*$. $\boxed{\text{seen} \Downarrow}$

We only need to update the last $m - k + 1$ rows and columns, which transform according to

$$(I_{m-k} - 2vv^*)A^{(k)}_{k:m,k:n} = A^{(k)}_{k:m,k:n} - 2(vv^*)A^{(k)}_{k:m,k:n} = A^{(k)}_{k:m,k:n} - 2v(v^*A^{(k)}_{k:m,k:n}),$$

as required. $\boxed{5, \text{C}}$

2. (a) (i) Stability of $\tilde{f}$ means that for each $x \in X$, there exists $\tilde{x}$ with seen ⇓

$$\frac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = \mathcal{O}(\epsilon), \text{ and } \frac{\|\tilde{x} - x\|}{\|x\|} = \mathcal{O}(\epsilon),$$

where ϵ is the "machine epsilon" for the floating point number system. 4, A

(ii) Backward stability of $\tilde{f}$ means that for each $x \in X$, there exists $\tilde{x}$ with

$$\tilde{f}(x) = f(\tilde{x}), \text{ and } \frac{\|\tilde{x} - x\|}{\|x\|} = \mathcal{O}(\epsilon).$$

4, A

(b) In this case the input to f is A, and the output of f is H, the solution to $Q^*HQ = A$. Let Q^* be the exact unitary matrix formed by

$$Q^* = \prod_{k=2}^{m-1} \begin{pmatrix} I_k & 0 \\ 0 & I_{m-k} - 2\frac{v_k v_k^*}{v_k^* v_k} \end{pmatrix}, \tag{7}$$

where v_k are obtained during the Householder reduction to Hessenberg form approximated in the floating point number system. Then, backwards stability means that there exists $\tilde{A}$ such that $\tilde{A} = Q^*\tilde{H}Q$, where $\tilde{H}$ is obtained by applying the Householder reduction to Hessenberg form, approximated in the floating point number system, such that $\|\tilde{A} - A\|/\|A\| = \mathcal{O}(\epsilon)$. 6, B

(c) The matrix vector product Aq_{k-1} can be evaluated in $\mathcal{O}(m)$ FLOPs, if the sum is only made over the $\mathcal{O}(m)$ nonzero entries in A.

Inside the j loop there are $O(m)$ FLOPs to compute ($O(m)$ multiplications and sums in the inner product, and $O(m)$ multiplications and sums in the update of q). There are $k-1$ of these loops, so this part is $\mathcal{O}(km)$ in total. Then, the norm computation and normalisation are also $\mathcal{O}(m)$. This gives $\mathcal{O}(km)$ for iteration k. Summing from 1 to k is approximately $\mathcal{O}(k^2m)$. 6, B

3. (a) $C = -M^{-1}N = -M^{-1}(A-M) = I - M^{-1}A$. Thus $\rho(I - M^{-1}A) = \rho(C) < c$. If c is small then the eigenvalues of $I - M^{-1}A$ are contained in a disc of radius c around the origin in the complex plane. The eigenvectors of $I - M^{-1}A$ are the same as the eigenvectors of $M^{-1}A$, hence if λ is an eigenvalue of C then $1-\lambda$ is an eigenvalue of $M^{-1}A$. Hence, the eigenvalues of the preconditioned matrix $M^{-1}A$ are contained in a disc of radius c around 1 in the complex plane. If c is small then all of the eigenvalues of $M^{-1}A$ are clustered close together and far from 0, so GMRES will converge quickly. seen ⇓

5, B

(b) For Gauss-Seidel, $M = L + D$, so seen ⇓

$$M^T + M - A = \underbrace{L^T + D^T}_{=U+D} + L + D - L - D - U = 2D. \tag{8}$$

If A is positive definite then so is D. To see this, note that $e_i^T A e_i = A_{ii} > 0$ by positive definiteness. Hence, $M^T + M - A$ is positive definite, so $\|C\| = \|I - M^{-1}A\| < 1$, so Gauss-Seidel always converges.

5, A

(c) unseen ⇓

$$(L+D)x^{k+1/2} = -L^T x^k + b, \quad (L^T + D)x^{k+1} = -Lx^{k+1/2} + b. \tag{9}$$

Eliminating,

$$\begin{aligned}
(L^T + D)x^{k+1} =& -L(L+D)^{-1}(-L^T x^k + b) + b, && (10)\\
=& L(L+D)^{-1}L^T x^k + \left(I - L(L+D)^{-1}\right) b, && (11)\\
=& L(L+D)^{-1}L^T x^k + (L + D - L)(L+D)^{-1} b, && (12)\\
=& L(L+D)^{-1}L^T x^k + D(L+D)^{-1} b, && (13)
\end{aligned}$$

hence,

$$(L+D)D^{-1}(L^T + D)x^{k+1} = (L+D)D^{-1}L(L+D)^{-1}L^T x^k + b, \tag{14}$$

which is in the form

$$M_S x^{k+1} = (M_S - A)x^k + b, \tag{15}$$

with $M_S = (L+D)D^{-1}(L^T + D)$, hence the result. (Note that an alternative form is obtained by using $L^T = U$.)

5, D

(d) We have unseen ⇓

$$\begin{aligned}
(I - (L^T + D)^{-1}A)(I - (L+D)^{-1}A) =& I - (L^T + D)^{-1}A - (L+D)^{-1}A \\
&+ (L^T + D)^{-1}A(L+D)^{-1}A, && (16)\\
=& I - (L^T + D)^{-1}\left(L + D + L^T + D - A\right)(L+D)^{-1}A, && (17)\\
=& I - (L^T + D)^{-1}D(L+D)^{-1}A, && (18)\\
=& I - M_S^{-1}A. && (19)
\end{aligned}$$

From the hint, we know that the $\rho(I - M^{-1}A) = \|I - M^{-1}A\|_A$. Convergence of forward Gauss-Seidel means that $\|I - (L+D)^{-1}A\|_A < 1$, and backward Gauss-Seidel convergence (follows from similar calculation) means that $\|I - (L^T + D)^{-1}A\| < 1$. From the above,

$$\begin{aligned}
\|I - M_S^{-1}A\|_A =& \|(I - (L^T + D)^{-1}A)(I - (L+D)^{-1}A)\|_A, && (20)\\
\leq& \|I - (L^T + D)^{-1}A)\|_A \|I - (L+D)^{-1}A\|_A < 1, && (21)
\end{aligned}$$

so symmetric Gauss-Seidel converges. 5, D

4 (Mastery). (a)

seen ⇓

$$A^{(k)} = R^{(k-1)}Q^{(k-1)} = Q^{(k-1),*}Q^{(k-1)}R^{(k-1)}Q^{(k-1)} \tag{22}$$

$$= Q^{(k-1),*}A^{(k-1)}Q^{(k-1)} = \underbrace{\left(\prod_{j=0}^{k-1} Q^{(j),*}\right)}_{=Q^{-1}} A \underbrace{\prod_{i=0}^{k-1} Q^{(i)}}_{=Q}, \tag{23}$$

by induction, and since $Q^{-1} = Q^*$ because products of unitary matrices are unitary. Hence A is similar to $A^{(k)}$ with Q playing the role of the similarity transformation.

5, M

(b) The basis vector $e_m = (0, 0, \ldots, 1)^T$ is almost an eigenvalue for $A^{(k)}$. To see this, we use the columns space interpretation of matrix vector multiplication, $A^{(k)}e_m = a_m$, the last column of $A^{(k)}$, and we can write $a_m = \alpha e_m + \epsilon e_{m-1}$, and hence

unseen ⇓

$$A^{(k)}e_m - \alpha e_m = \epsilon e_{m-1} \implies \|A^{(k)}e_m - \alpha e_m\| = \epsilon\|e_{m-1}\| = \epsilon. \tag{24}$$

5, M

(c) We have

unseen ⇓

$$\epsilon = \|(A^{(k)} - \alpha I)e_m\| = \|(Q^*(A - \alpha I)Qe_m\| = \|(A - \alpha I)w\|, \tag{25}$$

with $w = Qe_m$, so α is an approximate eigenvalue of A with approximate eigenvector w.

5, M

Review of mark distribution:

Total A marks: 23 of 24 marks

Total B marks: 17 of 15 marks

Total C marks: 10 of 9 marks

Total D marks: 10 of 12 16 marks

Total marks: 75 of 60 marks

Total Mastery marks: 15 of 15 marks

MATH60024 Computational Linear Algebra Markers Comments

Question 1

A large number of candidates struggled with this bookwork question.
In part (a), a number of candidates got confused between the definitions of unitary and Hermitian.
In part (b), a number of candidates made elementary algebra errors, sometimes because of needlessly resorting to writing out components.
In part (c), many candidates made a vague explanation alluding to the geometric motivation from the class notes, but without formulating a convincing logical argument.
Part (d) is testing the ability to translate from a mathematical concept to an efficient implementation. Many candidates gave a very poorly presented answer that did not form a convincing logical argument. A number of candidates missed the important point that the bottom left block of A^k will be zero.

Question 2

Most candidates made a good answer to this question. The main issues were in part c, where a few candidates missed the point that they were being asked to provide the cost after k iterations, not for the whole algorithm (Arnoldi is not usually iterated until $k=n$).

Question 3

There was a spectrum of quality of answers to this question. Several candidates weren't able to recall the connection between the iteration matrix C and the eigenvalues of the preconditioned operator $M^{-1}A$ ($C = I - M^{-1}A$). Many candidates made elementary algebra errors in computing $M^T+M - A = D$. Many candidates did not provide a specific formula for M_s in the case of Gauss-Seidel. No candidates were able to use norm properties to show that convergence of Gauss-Seidel implies convergence of symmetries Gauss-Seidel for SPD matrices.