DELFT UNIVERSITY OF TECHNOLOGY
Faculty of Electrical Engineering, Mathematics and Computer Science

ANSWERS OF THE TEST SCIENTIFIC COMPUTING (wi4201)
Wednesday January 20 2021, 13:30-16:30

This are short answers, which indicate how the exercises can be answered. In most of the cases more details are needed to give a sufficiently clear answer.

1. (a) Yes. Since Q is an orthogonal matrix we know that $QQ^T = Q^TQ = I$. Suppose $\lambda \in \sigma(A)$.
$$A\mathbf{v} = \lambda\mathbf{v}$$
$$Q^TA\mathbf{v} = \lambda Q^T\mathbf{v}$$
$$Q^TAQQ^T\mathbf{v} = \lambda Q^T\mathbf{v}$$
suppose $\mathbf{w} = Q^T\mathbf{v}$
$$Q^TAQ\mathbf{w} = \lambda\mathbf{w}$$
so $\lambda \in \sigma(Q^TAQ)$.

(b) No. Counterexample:
$$A = \begin{pmatrix} 4 & 1 \\ -1 & 4 \end{pmatrix}.$$
The eigenvalues are $\lambda_1 = 4+i$ and $\lambda_2 = 4-i$ so the eigenvalues are not elements of $\mathbb{R}$.

(c) Yes. Note that D is SPD. Check the properties of the innerproduct.

- Symmetry
$$(\mathbf{x},\mathbf{y})_D = \mathbf{x}^TD\mathbf{y} = (D^T\mathbf{x})^T\mathbf{y} = \mathbf{y}^T(D^T\mathbf{x}) = \mathbf{y}^TD\mathbf{x} = (\mathbf{y},\mathbf{x})_D$$
- Linear
$$(a\mathbf{x}+b\mathbf{y},\mathbf{z})_D = a\mathbf{x}^T\mathbf{z} + b\mathbf{y}^T\mathbf{z} = a(\mathbf{x},\mathbf{z})_D + b(\mathbf{y},\mathbf{z})_D$$
- positivity
$$(\mathbf{x},\mathbf{x})_D = \mathbf{x}^TD\mathbf{x} = \sum_{i=1}^{n} \frac{i}{n}x_i^2$$
This implies that $(\mathbf{x},\mathbf{x})_D > 0$ if $\mathbf{x}$ is not the zero vector and if $(\mathbf{x},\mathbf{x})_D = 0$ $\mathbf{x}$ is the zero vector.

(d) Yes. The Frobenius norm can be viewed as the 2-norm of the vector obtained from all rows of A (see the line below 2.16 in the lecture notes). It is well known that the 2-norm is the same after orthogonal transformations (Proposition 2.5.2) which proves the claim.

(e) No. From the definition it follows that the spectral radius is the in absolute value largest eigenvalues, so it can never be negative.

2. (a) We first start by computing the Gershgorin disks. Note that all disks has center at $\frac{2}{h^2}$. The radius is equal to $\frac{1}{h^2}$ or $\frac{2}{h^2}$. This implies that all eigenvalues should be in the disk with center at $\frac{2}{h^2}$ and radius $\frac{2}{h^2}$. Since the matrix is symmetric that implies that all eigenvalues are real and $0 \leq \lambda \leq \frac{4}{h^2}$. The eigenvalues are given by $\frac{2}{h^2} 2 \sin^2(\frac{\pi h k}{2})$. Since $0 \leq \sin^2(\frac{\pi h k}{2}) \leq 1$ it easily follows that the eigenvalues are contained in the Gershgorin disks.

(b) If $N = 3$ we end up with a 2×2 system. There are only two eigenvectors. Substitute all relevant information in expression (3.52) and check that the innerproduct between both vectors is zero.

(c) There are two correct ways to obtain the result: $\lambda = \frac{4}{h^4} 4 \sin^4(\frac{\pi h k}{2})$. One way is to use the ansatz for the eigenvectors as given in (3.52) and compute the eigenvalues. An easier way is to prove that the discretization matrix for the bi-Harmonic equation as given in (3.69) is the square of the Poisson matrix as given in (3.48). Then the eigenvalues of the bi-Harmonic matrix are the square of the eigenvalues of the Poisson matrix which leads to the correct expression.

(d) The stencil is given by:

$$\frac{1}{h^2} \begin{bmatrix} 0 & -1 & 0 \\ -1 - \frac{h}{2} & 4 + 6h^2 & -1 + \frac{h}{2} \\ 0 & -1 & 0 \end{bmatrix}$$

Central differences are used for all derivatives. To prove that this stencil has second order accuracy, Taylor polynomials have to be used with a remainder term of order h^4 for the second derivatives and of order h^3 for the first derivative.

3. (a) The inverse of Gauss transformation $M_k = I - \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T$ is given by $M_k^{-1} = I + \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T$. To show this claim, we have to check if the following expression is correct $M_k M_k^{-1} = I$. We note that

$$(I - \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T)(I + \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T) =$$

$$I - \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T + \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T + \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T =$$

$$I + \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T = I$$

The final equality follows from the fact that the term $\boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T \boldsymbol{\alpha}^{(k)} \mathbf{e}_k^T$ is equal to the zero matrix. The easiest way to show this is to first do the multiplication $\mathbf{e}_k^T \boldsymbol{\alpha}^{(k)}$. Note that only the $k - th$ component of $\mathbf{e}_k$ is nonzero. However, it is known that the first k components of vector $\boldsymbol{\alpha}^{(k)}$ are equal to zero, so the resulting inner product is also zero.

(b) After the first step of the Gaussian elimination, element $a_{2,2}$ is equal to zero. That means a zero pivot, so the next step can not be done. This problem can be solved by using partial pivotting. Since the pivot is zero one looks in the second column to see if there is a non zero element. That is true for the final row, so one puts that row in the second row and the second row in the third row. Then the matrix is upper triangular and you can find the solution: $\mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

(c) This can be proven by contradiction. Since the matrix is strictly column diagonal dominant we know that $|a_{1,1}| > \sum_{i=2}^{n} |a_{i,1}| \geq 0$. This implies that $a_{1,1} \neq 0$ so the first pivot is non zero and the first Gaussian elimination step can be done.

(d) Let us consider column j of L^{-1}. Suppose there is a k such that $1 \leq k < j$ such that $L^{-1}_{k,j} \neq 0$ and $L^{-1}_{i,j} = 0$, for all $i < k$. Then the result of the multiplication of row k of L with column j of L^{-1} is equal to $L^{-1}_{k,j}$. However, since $LL^{-1} = I$ this element should be equal to zero, so it is impossible to find a k as indicated above.

4. (a) The Jacobi iteration matrix is defined as $B = I - D^{-1}A$ where D is a diagonal matrix and $D_{i,i} = A_{i,i}$. The relevant stencils are:

$$A = \frac{1}{h^2}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}, D^{-1} = h^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ and } I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Substituting these stencils in the definition of B leads to the correct result.

(b) It is important to note that Symmetric Gauss Seidel is the same method as SSOR where $\omega = 1$. Using this equivalence M_{SGS} follows from (5.42). So B_{SGS} is given by:

$$B = I - M_{SGS}^{-1}A = I - (D-E)^{-1}D(D-F)^{-1}(D-E-F)$$

(c) We obtain optimal convergence, if $\rho(B_{Jac(\omega)})$ is minimal. Since $a_{i,i} = 1$ we obtain $B_{Jac(\omega)} = I - \omega A$. This implies that the eigenvectors of A and B are identical, and $\lambda(B_{Jac(\omega)} = 1 - \omega\lambda(A)$. So

$$1 - \omega\lambda_n \leq \lambda(B_{Jac(\omega)} \leq 1 - \omega\lambda_1$$

To minimise the spectral radius of $B_{Jac(\omega)}$ we take ω such that both the smallest and the largest eigenvalue are the same in absolute value. This implies that $1 - \omega\lambda_1 = -1 + \omega\lambda_n$ solving of this equation for ω gives $\omega_{optimal} = \frac{2}{\lambda_1 + \lambda_n}$.

(d) From the theory it follows that $\|\mathbf{e}^k\|_2 \leq (\rho(B_{JAC}))^k \|\mathbf{e}^0\|_2$. So the stopping criterion is satisfied if

$$(1 - \frac{\pi^2}{2}h^2)^k \leq 10^{-4}$$

$$k \log_{10}(1 - \frac{\pi^2}{2} h^2) \le -4$$
$$k \ge \frac{4}{\log_{10}(1 - \frac{\pi^2}{2} h^2)}$$

5. (a) The original question is: given $\mathbf{u}^k = \mathbf{u}^{k-1} + \alpha_k \mathbf{p}^k$ compute α_k such that $\mathbf{e}^k$ is perpendicular to $\mathbf{p}^k$ in the A inner product.

During the exam it seems that it should be "$\mathbf{u}^k$ is perpendicular to $\mathbf{p}^k$ in the A inner product".

Finally it appears that the original question is correct. Both answers will be given and both answers can lead to all points.

Original question:
$$(\mathbf{e}^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{u} - \mathbf{u}^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{u} - \mathbf{u}^{k-1} - \alpha_k \mathbf{p}^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{e}^{k-1} - \alpha_k \mathbf{p}^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{e}^{k-1})^T A \mathbf{p}^k - \alpha_k (\mathbf{p}^k)^T A \mathbf{p}^k = 0$$
Since $A = A^T$ we obtain: $(\mathbf{e}^{k-1})^T A = (A\mathbf{e}^{k-1})^T = (\mathbf{r}^{k-1})^T$. Substituting this leads to:
$$\alpha_k = \frac{(\mathbf{r}^{k-1})^T A \mathbf{p}^k}{(\mathbf{p}^k)^T A \mathbf{p}^k}$$
Adapted question:
$$(u^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{u}^{k-1} + \alpha_k \mathbf{p}^k)^T A \mathbf{p}^k = 0$$
$$(\mathbf{u}^{k-1})^T A \mathbf{p}^k + \alpha_k (\mathbf{p}^k)^T A \mathbf{p}^k = 0$$
This leads to:
$$\alpha_k = \frac{-(\mathbf{u}^{k-1})^T A \mathbf{p}^k}{(\mathbf{p}^k)^T A \mathbf{p}^k}$$

(b) Note that the vectors $\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\}$ form an independent set. To proof this, suppose the vectors are dependent. Let us assume that $\mathbf{v}^1 = \alpha \mathbf{v}^2 + \beta \mathbf{v}^3$. Due to orthogonality we have $(\mathbf{v}^1)^T \mathbf{v}^2 = 0$. Substituting $\mathbf{v}^1 = \alpha \mathbf{v}^2 + \beta \mathbf{v}^3$ shows that $(\mathbf{v}^1)^T \mathbf{v}^2 = (\alpha \mathbf{v}^2 + \beta \mathbf{v}^3)^T \mathbf{v}^2 = \alpha (\mathbf{v}^2)^T \mathbf{v}^2 \neq 0$. This is a contradiction which shows that $\{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^n\}$ is an independent set of vectors, so they span $\mathbb{R}^n$. This implies that if $\mathbf{z}^T \mathbf{v}^i = 0$ for $i = 1, \ldots, n$ then $\mathbf{z} = 0$.

(c) The matrix has only 2 different eigenvalues: 1 and 10. At the top of page 104 of the lecture notes it is shown that the number of iterations of CG before convergence is less than the number of different eigenvalues. Since the matrix has 2 different eigenvalues, CG should converge in 1 or 2 iterations.

(d) See next page.