IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

EXAMINATIONS 2021

MATHEMATICS FOR SIGNALS AND SYSTEMS SOLUTIONS

SOLUTIONS

1. (a) i. The matrix $\mathbf{A}$ has two non-zero singular values. We first find the eigenvalues of $\mathbf{A}\mathbf{A}^T$:

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 16 & 0 \\ 0 & 4 \end{bmatrix}.$$

This means $\lambda_1 = 16$ and $\lambda_2 = 4$. Therefore, $\sigma_1 = 4$ and $\sigma_2 = 2$,

ii. Here, $\mathbf{v}_1$ and $\mathbf{v}_2$ are related to the two non-zero singular values, therefore they are given by:

$$\mathbf{v}_1^T = \frac{1}{\sigma_1}\mathbf{u}_1^T\mathbf{A} = \frac{1}{2}\begin{bmatrix} 2 & 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}$$

and

$$\mathbf{v}_2^T = \frac{1}{\sigma_2}\mathbf{u}_2^T\mathbf{A} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & -0.5 & -0.5 \end{bmatrix}.$$

The remaining two singular vectors are found by imposing the $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ form an orthonormal basis. Therefore, $\mathbf{v}_3 = \begin{bmatrix} 0.5 & -0.5 & 0.5 & -0.5 \end{bmatrix}^T$ and $\mathbf{v}_4 = \begin{bmatrix} -0.5 & 0.5 & 0.5 & -0.5 \end{bmatrix}^T$.

iii. The pseudo inverse of $\mathbf{A}$ is given by:

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^T,$$

where

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix},$$

$$\mathbf{\Sigma}^+ = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{bmatrix}.$$

Therefore

$$\mathbf{A}^+ = \begin{bmatrix} 1/8 & 1/4 \\ 1/8 & 1/4 \\ 1/8 & -1/4 \\ 1/8 & -1/4 \end{bmatrix}$$

and $\mathbf{x}_2 = \mathbf{A}^+\mathbf{b} = [1.5, 1.5, 3.5, 3.5]^T$. We first note that $\mathbf{A}$ is full row rank, therefore the problem $\mathbf{A}\mathbf{x} = \mathbf{b}$ has an infinite number of *exact* solutions and we have found two such solutions. The pseudo-inverse normally provides the minimum norm least square solution, but given that $\mathbf{A}$ is full row rank, it is in this case just providing the *exact* minimum norm solution and indeed $\|\mathbf{x}_2\| < \|\mathbf{x}_1\|$.

(b) Let $\mathbf{B} = \mathbf{A}\mathbf{A}^H$, we have that $\mathbf{B}$ is positive definite if and only if $\mathbf{x}^H\mathbf{B}\mathbf{x} > 0$ for any vector $\mathbf{x} \neq 0$. This is satisfied since for any $\mathbf{x} \neq 0$ we can write $\mathbf{x}^H\mathbf{B}\mathbf{x} = \mathbf{x}^H\mathbf{A}\mathbf{A}^H\mathbf{x} = \|\mathbf{A}^H\mathbf{x}\|^2 > 0$. The last inequality is due to the fact that, given that $\mathbf{A}$ is nonsingular, then $\mathbf{A}^H\mathbf{x} \neq 0$ when $\mathbf{x} \neq 0$.

(c) The matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ -2 & -8 \\ 3 & 12 \end{bmatrix}$$

has clearly rank one and a basis for the range space is $\mathbf{u}_1 = [1, -2, 3]^T$. The null space has dimension $2 - \text{rank}(\mathbf{A}) = 1$ and a possible basis for the null space is $\mathbf{n}_1 = [-4, 1]^T$; $\text{rank}(\mathbf{A}^T) = 1$ and $\text{range}(\mathbf{A}^T) = \text{span}([1, 4]^T)$. Since the null space of $\mathbf{A}^T$ has dimension two, the easiest way to find its basis is by finding two linearly independent vectors orthogonal to $\mathbf{u}_1$. This yields $\text{null}(\mathbf{A}^T) = \text{span}([-1, 1, 1]^T, [5, 1, -1]^T)$.

(d) i. The transition matrix is

$$\mathbf{P} = \begin{bmatrix} 0 & 0.6 & 0.8 \\ 0.5 & 0 & 0.2 \\ 0.5 & 0.4 & 0 \end{bmatrix}.$$

ii. Since the cat has moved many times, we can assume that we have now reached the steady state, so the location where the cat currently is is independent of the initial state. The eigenvector related to $\lambda_1 = 1$ indicates the steady state and so will indicate with which probability the cat is in any of the three locations. We then pick the location with highest probability. So we first find $\mathbf{u}_1$ such that $\mathbf{P}\mathbf{u}_1 = \mathbf{u}_1$. This yields $\mathbf{u}_1 = [0.7064, 0.4607, 0.5374]^T$. We therefore conclude that location 1 (under the bed) is the most likely one. We also note that to complete the proof we need to compute λ_2 and λ_3 and show that $|\lambda_i| < 1$ for $i = 2, 3$. This guarantees that the process converges to a steady state. Given that in this case $\text{trace}(\mathbf{P}) = 0$ and $\det(\mathbf{P}) = 0.22$, we conclude that this is the case. In fact, if we compute the eigenvalues directly we obtain $\lambda_2 = -0.6734$ and $\lambda_3 = -0.3268$.

(e) i. The rows of $\mathbf{P}^T$ add to one, this means that when we multiply one row of $\mathbf{P}^T$ with $\mathbf{1} = [1, 1, ..., 1]^T$, we obtain one. This means that $\mathbf{P}^T\mathbf{1} = \mathbf{1}$, which shows that $\mathbf{1}$ is the eigenvector related to the eigenvalue $\gamma = 1$.

ii. Because $\lambda_1 = 1$ and all other eigenvalues have $|\lambda| < 1$, the matrix

$$\mathbf{P}^n = \mathbf{X}\mathbf{\Lambda}^n\mathbf{X}^{-1} \text{ approaches } \mathbf{P}^\infty = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots \\ & & \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & . \end{bmatrix} \mathbf{X}^{-1}.$$

We now observe that $\mathbf{P}^T = (\mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1})^T = (\mathbf{X}^{-1})^T\mathbf{\Lambda}\mathbf{X}^T$. This means that the rows of $\mathbf{X}^{-1}$ are the eigenvectors of $\mathbf{P}^T$. Therefore the first row of $\mathbf{X}^{-1}$ is $\mathbf{1}^T$. This means that

$$\mathbf{P}^\infty = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots \\ & & \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & . \end{bmatrix} \begin{bmatrix} \mathbf{1}^T \\ . \\ . \\ . \end{bmatrix} = \mathbf{u}_1\mathbf{1}^T.$$

2. (a) i. Since h_n has only two non-zero entries, the convolution formula is in this case $y_n = \sum_{k=0}^{1} h_k x_{n-k}$. Since $x_n = 0$ for $n \neq 0, 1$ we have that $y_n = 0$ for $n \neq 0, 1, 2$ and that we can express the non-zero entries of the filtered sequence as follows: $\mathbf{y} = \mathbf{H}\mathbf{x}$ where the matrix $\mathbf{H}$ is 'tall' with size 3×2 and is given by

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

ii. The two columns are clearly linearly independent, so the matrix is full column rank and the null space is trivial. A possible basis for the range space is given by the two columns of $\mathbf{H}$.

iii. To find the least-equare solution we need to solve the normal equation:

$$\mathbf{H}^T\mathbf{H}\mathbf{x} = \mathbf{H}^T\mathbf{y}.$$

Since $\mathbf{H}$ is full column rank, $\mathbf{H}^T\mathbf{H}$ is invertible and we have:

$$\mathbf{x} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}.$$

Clearly

$$(\mathbf{H}^T\mathbf{H}) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

and

$$\mathbf{x} = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \mathbf{y} = \begin{bmatrix} 4/3 \\ 1/3 \end{bmatrix}.$$

iv. We are trying to solve a least-square problem with a linear constraint. Specifically, we want to minimise $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$ subject to $\mathbf{A}\mathbf{x} = 1$ with $\mathbf{A} = [1, 1]$. This optimisation can be solved using Lagrangian multipliers leading to the new function:

$$L(\mathbf{x}, \lambda) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \lambda(\mathbf{A}\mathbf{x} - 1).$$

Taking the derivative and equating to zero yields the system:

$$\begin{bmatrix} 2\mathbf{H}^T\mathbf{H} & \mathbf{A}^T \\ \mathbf{A} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} 2\mathbf{H}^T\mathbf{y} \\ 1 \end{bmatrix}$$

and the solution is $\mathbf{x} = [1, 0]^T$.

(b) i. Due to periodicity assumption we have that $\mathbf{H}$ is circulant and is given by :

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}.$$

The matrix is now square and of size 4×4.

ii. By inspection we see that the rank of the matrix is 3 since the last column can be obtained with a proper linear combination of the other 3 columns. The consequence is that now the null space is not trivial and has dimension $N-3 = 1$. The first 3 columns of $\mathbf{H}$ are a possible basis of the range space of $\mathbf{H}$. We then need to find a vector $\mathbf{n}$ such that $\mathbf{H}\mathbf{n} = \mathbf{0}$. One such vector is $\mathbf{n} = [1, -1, 1, -1]^T$ and the null space of $\mathbf{H}$ is given by $\text{span}\{\mathbf{n}\}$.

3. (a) i. Using the definition of Frobenius norm, we have that $\|\mathbf{A}\|_F^2 = \sum_{i,j} a_{i,j}^2$. The trace of a matrix is computed as the sum of its diagonal elements. By expanding the elements of $\mathbf{A^T A}$, it becomes evident that the sum of its diagonal elements equals $\|\mathbf{A}\|_F^2$, since the entry (k,k) of $\mathbf{A^T A}$ is $\sum_i a_{i,k}^2$.

ii. The diagonal elements of $\mathbf{A}^T\mathbf{B}$ are of the form $\sum_i a_{i,k}b_{i,k}$ and the same applies to $\mathbf{B}^T\mathbf{A}$ so the two matrices have the same trace.

iii. Based on these preliminary observations we have that

$$\|\mathbf{B}-\mathbf{SA}\|_F^2 = \text{trace}((\mathbf{B}-\mathbf{SA})^T(\mathbf{B}-\mathbf{SA})) = \text{trace}(\mathbf{A^T S^T S A})+\text{trace}(\mathbf{B^T B})-2\text{trace}(\mathbf{A^T S^T B}),$$

where we have used the fact, proved in (ii), that $\text{trace}(\mathbf{A^T S^T B}) = \text{trace}(\mathbf{B^T S A})$. We also have that $\text{trace}(\mathbf{A^T S^T S A}) = \text{trace}(\mathbf{A^T A})$, where in the last equality we have used the fact that $\mathbf{S^T S} = \mathbf{I}$ since $\mathbf{S}$ is an orthogonal matrix. We therefore conclude that:

$$\|\mathbf{B}-\mathbf{SA}\|_F^2 = \text{trace}(\mathbf{A^T A}) + \text{trace}(\mathbf{B^T B}) - 2\text{trace}(\mathbf{A^T S^T B}).$$

This means that finding $\mathbf{S}$ that minimises $\|\mathbf{B}-\mathbf{SA}\|_F^2$ is equivalent to maximizing $\text{trace}(\mathbf{A^T S^T B})$ since this is the only term that depends on $\mathbf{S}$.

iv. By using the cyclic property of the trace: $\text{trace}(\mathbf{ABC}) = \text{trace}(\mathbf{CAB}) = \text{trace}(\mathbf{BCA})$, we obtain

$$\text{trace}(\mathbf{A^T S^T B}) = \text{trace}(\mathbf{B A^T S^T}) = \text{trace}(\mathbf{U\Sigma V^T S^T}) = \text{trace}(\mathbf{V^T S^T U \Sigma}) = \text{trace}(\mathbf{Z\Sigma}),$$

where $\mathbf{Z} = \mathbf{V^T S^T U}$.

v. We note that $\mathbf{U}$, $\mathbf{S}$ and $\mathbf{V}$ are orthogonal matrices. Moreover, $\mathbf{Z}^T = \mathbf{U^T S V}$ and $\mathbf{Z}^T\mathbf{Z} = \mathbf{U^T S V V^T S^T U} = \mathbf{I}$. Therefore $\mathbf{Z}$ is orthogonal.

vi. From the fact that $\mathbf{Z} = \mathbf{I}$, we conclude that $\mathbf{S} = \mathbf{UV^T}$. What we have just solved is known as *the orthogonal Procrustes problem.* This problem appears in many engineering applications including computer graphics, robotics and navigation as well as in sensor networks and graph theory.

(b) i. Matrix $\mathbf{A}$ can be defined as the coordinates of the three points:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Similarly, matrix $\mathbf{B}$ can be defined as:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

ii. Matrix $\mathbf{M}$ is given by:

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}.$$

Moreover, we have that:

$$\mathbf{M^T M} = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix},$$

and:

$$\mathbf{M M^T} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}.$$

We can decompose $\mathbf{M^T M} = \mathbf{W D W^T}$, where:

$$\mathbf{W} = \begin{bmatrix} -0.9239 & -0.3827 \\ -0.3827 & 0.9239 \end{bmatrix},$$

and:

$$\mathbf{D} = \begin{bmatrix} 5.8284 & 0 \\ 0 & 0.1716 \end{bmatrix}.$$

Similarly $\mathbf{M M^T} = \mathbf{V D V^T}$, where:

$$\mathbf{V} = \begin{bmatrix} -0.3827 & -0.9239 \\ -0.9239 & 0.3827 \end{bmatrix}.$$

iii. Finally:

$$\mathbf{S} = \mathbf{W V^T} = \frac{1}{2}\begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}$$

and clearly $\mathbf{S}\mathbf{S}^T = \mathbf{I}$.