Delft University of Technology
Faculty of Electrical Engineering, Mathematics, and Computer Science
Section Signal Processing Systems

Partial exam EE2S31 SIGNAL PROCESSING
Part 1: 22 May 2024 (13:30-15:30)

Closed book; two sides of one A4 with handwritten notes permitted. No other tools except a basic pocket calculator permitted. Note the attached tables!

This exam consists of four questions (34 points). Answer in English. Make clear in your answer how you reach the final result; the road to the answer is very important.

Question 1 (10 points)

Given the joint probability density function of two random variables X and Y:

$$f_{X,Y}(x,y) = \begin{cases} c & 0 \le x \le 3,\ 0 \le y \le \frac{1}{2}x \\ 0 & \text{otherwise.} \end{cases}$$

(a) Calculate the value of the constant c.

(b) Show that $f_X(x)$ and $f_Y(y)$ are given by

$$f_X(x) = \begin{cases} \frac{1}{2}cx & 0 \le x \le 3 \\ 0 & \text{otherwise} \end{cases} \qquad f_Y(y) = \begin{cases} 3c - 2cy & 0 \le y \le \frac{3}{2} \\ 0 & \text{otherwise} \end{cases}$$

and argue whether the random variables X and Y are independent or not.

(c) Determine the blind MMSE estimator $\hat{x}_1$, which is based only on prior information.

(d) Suppose that we have additional knowledge that $X \in A$ with $A = \{X > 2\}$.
Determine the MMSE estimator $\hat{x}_2$ that takes this information into account.

(e) Instead, suppose that we observe a realization y of the random variable Y.
Determine the MMSE estimator $\hat{x}_3(y)$ that uses this information.

Solution

2p (a) It is recommended to first make a plot showing the (triangular) region over which $f(x,y)$ is nonzero.

$$\begin{aligned} \iint f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y &= \int_0^{3/2}\int_{2y}^{3} c\,\mathrm{d}x\mathrm{d}y \\ &= \int_0^{3/2} [cx]_{2y}^{3}\,\mathrm{d}y \\ &= \int_0^{3/2} c(3-2y)\,\mathrm{d}y \\ &= c\left[3y - y^2\right]_0^{3/2} \\ &= c(\frac{9}{2} - \frac{9}{4}) = \frac{9}{4}c \quad = 1 \end{aligned}$$

$$\begin{aligned} \iint f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x &= \int_0^{3}\int_0^{\frac{1}{2}x} c\,\mathrm{d}y\mathrm{d}x \\ &= \int_0^{3} [cy]_0^{\frac{1}{2}x}\,\mathrm{d}x \\ &= \int_0^{3} \frac{1}{2}cx\,\mathrm{d}x \\ &= \frac{1}{2}c\left[\frac{1}{2}x^2\right]_0^3 \\ &= \frac{9}{4}c \quad = 1 \end{aligned}$$

Hence, $c = \frac{4}{9}$.

2p (b) For $0 \le x \le 3$, resp. $0 \le y \le \frac{3}{2}$,

$$\begin{aligned} f_X(x) &= \int_0^{x/2} c\,\mathrm{d}y = \frac{1}{2}cx \\ f_Y(y) &= \int_{2y}^{3} c\,\mathrm{d}x = 3c - 2cy\,. \end{aligned}$$

X and Y are not independent as $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$.

2p (c) The blind MMSE estimator is given by

$$\hat{x}_1 = \mathrm{E}[X] = \int x\, f_X(x)\,\mathrm{d}x = \int_0^3 \frac{cx^2}{2}\mathrm{d}x = \frac{cx^3}{6} = \frac{27c}{6} = \frac{9c}{2}$$

2p (d) The MMSE estimator is $\hat{x}_2 = \mathrm{E}[X|A]$. Since A does not depend on Y,

$$\begin{aligned} \mathrm{P}[X > 2] &= \int_{x>2} f_X(x)\,\mathrm{d}x = \int_2^3 \frac{1}{2}cx\,\mathrm{d}x = \left[\frac{1}{4}cx^2\right]_2^3 = \frac{5c}{4} \\ f_{X|X>2}(x) &= \begin{cases} \dfrac{f_X(x)}{\mathrm{P}[X>2]} & x > 2 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \frac{4}{10}x & x > 2 \\ 0 & \text{otherwise} \end{cases} \\ \hat{x}_2 &= \mathrm{E}[X|X>2] = \int x\, f_{X|X>2}(x)\,\mathrm{d}x = \int_2^3 \frac{4x^2}{10}\mathrm{d}x = \left[\frac{4}{30}x^3\right]_2^3 = \frac{38}{15} \end{aligned}$$

2p (e) The MMSE estimator is $\hat{x}_3(y) = \mathrm{E}[X|Y=y]$.

$$\begin{aligned} f_{X|Y}(x|y) &= \begin{cases} \dfrac{f_{X,Y}(x,y)}{f_Y(y)}, & 2y \le x \le 3 \\ 0\,, & \text{otherwise} \end{cases} = \begin{cases} \dfrac{1}{3-2y} & 2y \le x \le 3 \\ 0 & \text{otherwise} \end{cases} \\ \hat{x}_3(y) &= \mathrm{E}[X|Y=y] = \int x\, f_{X|Y}(x|y)\,\mathrm{d}x = \int_{2y}^3 \frac{x}{3-2y}\mathrm{d}x = \left[\frac{x^2}{6-4y}\right]_{2y}^3 = \frac{9}{6-4y} - \frac{4y^2}{6-4y} \end{aligned}$$

Question 2 (8 points)

Last month, newspapers reported that King's Day (27 April) is apparently 1°C colder than Queen's Day (30 April). This was based on observations since 1949.

Denote the temperature on 27 April by the random variable T_K and on 30 April by T_Q. Long-term KNMI statistics suggest the following:

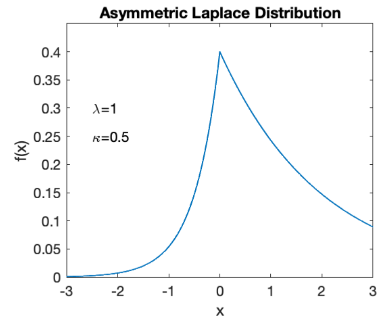
$$\begin{aligned} &\mathrm{E}[T_K] = 14.2, \qquad \mathrm{std}[T_K] = 2.0 \\ &\mathrm{E}[T_Q] = 14.7 \end{aligned}$$

Let Y be the average of the temperatures measured on King's Day over the last 75 years.

(a) Using Chebyshev's inequality, find an upper bound for the probability that Y is at least 1°C colder than $\mathrm{E}[T_Q]$.

(b) Estimate the probability that Y is at least 1°C colder than $\mathrm{E}[T_Q]$ using the central limit theorem.

The statistics on temperature suggest an asymmetric distribution with outliers: hot extremes are more likely than cold extremes. Let's consider an *Asymmetric Laplace Distribution* (ALD) with parameters λ and κ. The Moment Generating Function (MGF) for this distribution is given by



$$\phi(s) = \frac{1}{(1+s\frac{\kappa}{\lambda})(1-s\frac{1}{\lambda\kappa})}, \qquad \text{ROC: } -\frac{\lambda}{\kappa} < s < \lambda\kappa$$

Consider $T \sim \text{ALD}(\lambda, \kappa)$.

(c) Show that T can be obtained as the difference of two independent exponentially distributed random variables, $T = X_1 - X_2$, where $X_1 \sim \text{Exp}(\lambda\kappa)$ and $X_2 \sim \text{Exp}(\frac{\lambda}{\kappa})$. *See table.*

(d) Determine $\text{E}[T]$ and $\text{var}[T]$.

Solution

2p (a) $\text{E}[Y] = 14.2, \ \text{var}[Y] = \frac{2.0^2}{75} = 0.053$

$$\text{P}[Y \le 13.7] = \text{P}[Y - 14.2 \le -0.5] \le \text{P}[|Y - 14.2| \ge 0.5] \le \frac{\text{var}[Y]}{0.5^2} = \frac{0.053}{0.25} = 0.21$$

2p (b) $\text{std}[Y] = \sqrt{0.053} = 0.231$.

$$\text{P}[Y \le 13.7] = \text{P}\left[\frac{Y - \text{E}[Y]}{\text{std}[Y]} < \frac{13.7 - \text{E}[Y]}{\text{std}[Y]}\right] = \text{P}[Z < -2.17] = 1 - \Phi(2.17) = 1 - 0.985 = 0.0150$$

2p (c) The MGF of X_1 is (see table)

$$\phi_{X_1}(s) = \frac{\lambda\kappa}{\lambda\kappa - s}, \qquad \text{ROC: } s < \lambda\kappa$$

Likewise, the MGF of X_2 is (see table)

$$\phi_{X_2}(s) = \frac{\lambda/\kappa}{\lambda/\kappa - s}, \qquad \text{ROC: } s < \lambda/\kappa$$

The scaling formula for a MGF shows that the MGF of $-X_2$ is

$$\phi_{-X_2}(s) = \phi_{X_2}(-s) = \frac{\lambda/\kappa}{\lambda/\kappa + s}, \qquad \text{ROC: } s > -\lambda/\kappa$$

Finally, since for the sum of two independent RVs the MGFs multiply, we obtain

$$\phi_T(s) = \frac{\lambda\kappa}{\lambda\kappa - s}\frac{\lambda/\kappa}{\lambda/\kappa + s} = \cdots = \frac{1}{(1+s\frac{\kappa}{\lambda})(1-s\frac{1}{\kappa\lambda})}$$

The ROC of the product is the intersection of both ROCs.

2p (d) Two different approaches:

Using $T = X_1 - X_2$ and the table on the exponential distribution:

$$\begin{aligned} \text{E}[T] &= \text{E}[X_1] - \text{E}[X_2] = \frac{1}{\lambda\kappa} - \frac{\kappa}{\lambda} = \frac{1-\kappa^2}{\lambda\kappa} \\ \text{var}[T] &= \text{var}[X_1] + \text{var}[X_2] = \frac{1}{\lambda^2\kappa^2} + \frac{\kappa^2}{\lambda^2} = \frac{1+\kappa^4}{\lambda^2\kappa^2} \quad \text{(used independence of } X_1, X_2) \end{aligned}$$

Alternatively, use the MGF (more work!):

$$
\begin{aligned}
\frac{\mathrm{d}\phi}{\mathrm{d}s} &= \frac{1/(\lambda\kappa)}{(1-s/(\lambda\kappa))^2(1+s\kappa/\lambda)} - \frac{\kappa/\lambda}{(1-s/(\lambda\kappa))(1+s\kappa/\lambda)^2} \\
\mathrm{E}[T] &= \left.\frac{\mathrm{d}\phi}{\mathrm{d}s}\right|_{s=0} = \frac{1}{\lambda\kappa} - \frac{\kappa}{\lambda} = \frac{\kappa^2-1}{\lambda\kappa} \\
\frac{\mathrm{d}^2\phi}{\mathrm{d}s^2} &= \cdots \quad \text{(4 terms)} \\
\mathrm{E}[T^2] &= \left.\frac{\mathrm{d}^2\phi}{\mathrm{d}s^2}\right|_{s=0} = \cdots = \frac{2\kappa^4 - 2\kappa^2 + 2}{\lambda^2\kappa^2} \\
\mathrm{var}[T] &= \mathrm{E}[T^2] - (\mathrm{E}[T])^2 = \cdots = \frac{1+\kappa^4}{\lambda^2\kappa^2}
\end{aligned}
$$

(The details are straightforward but a bit too painful to write out.)

Question 3 (6 points)

An audio transmitter emits a sine wave $x(t)$ at a frequency between 20 kHz and 23 kHz, described by the equation

$$x(t) = \sin(2\pi F_T t),$$

where F_T is between 20 kHz and 23 kHz. The receiver aims to estimate the frequency of the transmitted signal.

(a) What is the frequency band of interest for the receiver?

(b) Which sampling frequency should the receiver choose if it follows the Nyquist rate?

Suppose the receiver samples the signal at 16 kHz. Let the discrete time Fourier transform (DTFT) of the signal sampled at 16 kHz be $X(f)$, where f is the normalized frequency.

(c) Sketch the spectrum $|X(f)|$ of the received signal for the normalized frequency range $-0.5 \le f \le 0.5$ for two cases: when the signal frequency is 20 kHz and when it is 23 kHz.

(d) How does the receiver estimate the transmit frequency using the DTFT $X(f)$? Justify your estimation technique.

Hint: The Fourier transform of $\sin(2\pi F_0 t)$ *is* $X(F) = \frac{j}{2}\delta(F - F_0) - \frac{j}{2}\delta(F + F_0)$.

Solution

1p (a) The highest frequency is $F_H = 23$ kHz, and the lowest is $F_L = 20$ kHz. Therefore, the bandwidth is $B = 3$ kHz, and the center frequency is $F_c = 21.5$ kHz.

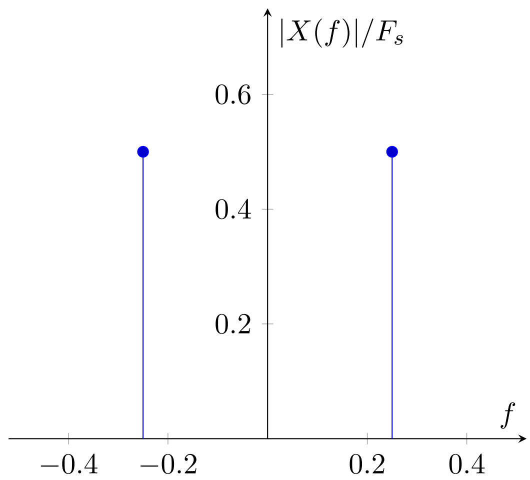
1p (b) The Nyquist rate is $2F_H = 46$ kHz.

2p (c) When the transmit frequency is F_T, the DTFT is

$$|X(f)| = \frac{F_s}{2} \sum_{k=-\infty}^{\infty} \delta(f - \frac{F_T}{F_s} + k) + \delta(f + \frac{F_T}{F_s} + k),$$

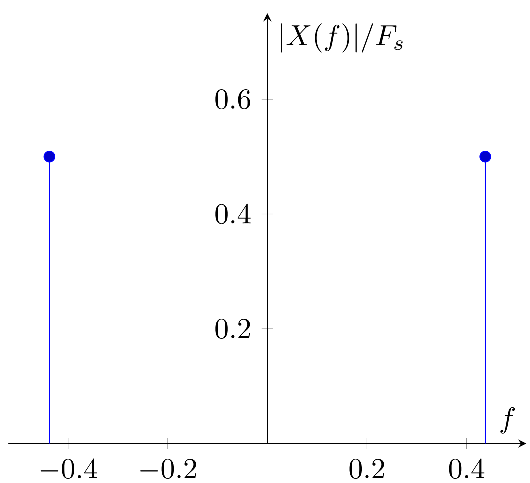
where $F_s = 16$ kHz. When the transmit frequency is $F_T = 20$ kHz, the DTFT is

$$\frac{|X(f)|}{F_s} = \frac{1}{2} \sum_{k=-\infty}^{k=\infty} \delta(f - \frac{1}{4} + k) + \delta(f + \frac{1}{4} + k).$$



When the transmit frequency is 23 kHz, the DTFT is

$$\frac{|X(f)|}{F_s} = \frac{1}{2}\sum_{k=-\infty}^{k=\infty} \delta(f - \frac{7}{16} + k) + \delta(f + \frac{7}{16} + k).$$



2p (d) Since F_T is between 20 kHz and 23 kHz, we derive

$$\frac{5}{4} \leq \frac{F_T}{F_s} \leq \frac{23}{16} \implies \frac{1}{4} \leq \frac{F_T}{F_s} - 1 \leq \frac{7}{16}.$$

Therefore, a peak in the DTFT occurs at $\pm\left(\frac{F_T}{F_s} - 1\right)$. Hence, the estimated frequency is $16(|f_P| + 1)$ kHz, where f_P corresponds to the peak in the DTFT where $-0.5 \leq f_P \leq 0.5$.

Question 4 (10 points)

Consider a discrete-time signal $x[n] = [x[0], x[1], x[2], \ldots, x[N-1]]$. Let the DTFT of $x[n]$ be $X(f)$ and its N-point discrete Fourier transform (DFT) be $X[k]$.

Consider another signal $\tilde{x}[n] = [x[N-1], x[0], x[1], \ldots, x[N-2]]$. Let $\tilde{X}[k]$ denote its DFT.

(a) Prove that $\tilde{X}[k] = e^{-j2\pi k/N} X[k]$, for $k = 0, 1, \ldots, N-1$.

Let $y[n]$ be obtained by zero padding $x[n]$ to make its length $2N$, i.e.,

$$y[n] = [x[0], x[1], \ldots, x[N-1], \underbrace{0, 0, \ldots, 0}_{N \text{ zeros}}].$$

Let the DTFT and DFT of $y[n]$ be $Y(f)$ and $Y[k]$, respectively.

(b) Derive a relation between $Y(f)$ and $X(f)$.

(c) What is the relation between the DFTs $X[k]$ and $Y[k]$? Explain your answer.

Consider a filter $h[n] = [h[0], h[1], h[2]]$ and let $N = 32$. Suppose $z[n] = x[n] * h[n]$, where $*$ represents linear convolution.

(d) Explain how to obtain $z[n]$ using the circular convolution.

(e) Explain how to obtain $z[n]$ using the overlap-save method with an 8-point DFT. How many 8-point DFTs and IDFTs are required?

Solution

2p (a) The DFT is given by

$$\begin{aligned}
e^{-j2\pi k/N} X[k] &= \sum_{n=0}^{N-2} x[n] e^{-j2\pi k(n+1)/N} + x[N-1] \\
&= \sum_{n=1}^{N-1} x[n-1] e^{-j2\pi kn/N} + \tilde{x}[0] \\
&= \sum_{n=1}^{N-1} \tilde{x}[n] e^{-j2\pi kn/N} + \tilde{x}[0] e^{-j2\pi k0/N} \\
&= \sum_{n=0}^{N-1} \tilde{x}[n] e^{-j2\pi kn/N} = \tilde{X}[k].
\end{aligned}$$

2p (b) The DTFT $Y[k]$ is

$$Y(f) = \sum_{n=0}^{2N-1} y[n] e^{-j2\pi fn} = \sum_{n=0}^{N-1} x[n] e^{-j2\pi fn} = X(f).$$

2p (c) For $k = 0, 1, \ldots, N-1$,

$$X[k] = X(k/N) = Y(k/N) = Y(2k/2N) = Y[2k].$$

2p (d) First zero pad $x[n]$ and $h[n]$ to get $x_0[n]$ and $h_0[n]$, respectively, such that they have the same length as $z[n]$, i.e., $N + 3 - 1 = 34$. Next, take the DFT of both $x_0[n]$ and $h_0[n]$, resulting in $X_0[k]$ and $H_0[k]$, respectively, each of length 34. Then, $Z[k] = X[k]H[k]$, for all k. Finally, the IDFT of $Z[k]$ provides $z[n]$.

2p (e) We start by adding the $M - 1 = 2$ zeros to the input and then dividing it into blocks of size $L = 8 - (M-1) = 6$. The number of blocks is $\lfloor 34/6 \rfloor = 6$. The first interval $x_1[n]$ is $[0, 0, x[0], \ldots, x[5]]$. For $r = 2, 3, \ldots, 5$, the $x_r[n]$ interval is given by $[x[6r-8], x[6r-7], \ldots, x[6r-1]]$. The last interval is $x_6[n]$ interval is given by $[x[28], x[29], \ldots, x[31], 0, 0, 0, 0]$.

We zero pad $h[n]$ to obtain $h_c[n]$ such that it has length 8. Then, we compute the circular convolution of $x_r[n]$ and $h_c[n]$ to get $y_r[n]$. Finally, to match $z[n]$ and $z_r[n]$, we discard the first $M - 1 = 2$ samples of $z_r[n]$.

The above method requires computing the DFT of $h_c[n]$, and each interval needs one DFTs and one IDFT, requiring a total of 7 DFTs and 6 IDFTs.