SSP1 Homeworks

John Doe, *Fellow, OSA,* and Jane Doe, *Life Fellow, IEEE*

Abstract

The abstract goes here.

Index Terms

IEEE, IEEEtran, journal, LaTeX, paper, template.

I. Probability and statistics

A. Question 1

Calculate expected value μ and variance σ^2 of the random variable x which follows the probability distribution

$$f(x) = \begin{cases} b\exp\left(-\frac{x}{a}\right) & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $a > 0$ is known. Determine the constant b first.

Answer: When integrated over the probability density function, the result must be 1. We get that

$$\int_0^\infty b\exp\left(-\frac{x}{a}\right)dx = b\int_0^\infty \exp\left(-\frac{x}{a}\right)dx = ba$$

so that $b = 1/a$. Therefore, the distribution is

$$f(x) = \begin{cases} \frac{1}{a}\exp\left(-\frac{x}{a}\right) & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The mean is by definition

$$\mu = \int_0^\infty \frac{x}{a}\exp\left(-\frac{x}{a}\right)dx = a$$

The variance is by definition

$$\sigma^2 = E\left[x^2\right] - E[x]^2 = \int_0^\infty \frac{x^2}{a}\exp\left(-\frac{x}{a}\right)dx - a^2$$
$$= 2a^2 - a^2 = a^2$$

B. Question 3

Assume that random variable x has the distribution

$$f(x) = \begin{cases} \frac{1}{2}\exp(x) & x < 0 \\ \frac{1}{2}\exp(-x) & x \geq 0 \end{cases}$$

Find mean and variance of x.

M. Shell was with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA e-mail: (see http://www.michaelshell.org/contact.html).

J. Doe and J. Doe are with Anonymous University.

Manuscript received April 19, 2005; revised August 26, 2015.

Answer: The mean is obtained with

$$E[x] = \int_0^\infty \frac{x}{2}\exp(-x) + \int_{-\infty}^0 \frac{x}{2}\exp(x) = 0$$

and the variance is

$$\mathrm{Var}[x] = E\left[x^2\right] - 0^2 = \int_0^\infty \frac{x^2}{2}\exp(-x) + \int_{-\infty}^0 \frac{x^2}{2}\exp(x) = 2$$

C. Question 4

We use estimator

$$\widehat{\theta} = \frac{1}{2N}\sum_{i=1}^{N}\left(z[i]^2 + 2\right)$$

to estimate unknown parameter θ. In addition we know that $E\left[z[i]^2\right] = 2(\theta - 1)$. Prove that $E\left[\hat{\theta}\right] = \theta$.

Answer: By the linearity of expectation we get

$$\begin{aligned} E\left[\widehat{\theta}\right] &= \tfrac{1}{2N}\sum_{i=1}^{N}\left(E\left[z[i]^2\right] + 2\right) = \tfrac{1}{2N}\sum_{i=1}^{N}\left(2(\theta - 1) + 2\right) \\ &= \tfrac{1}{2N}\sum_{i=1}^{N}(2\theta - 2 + 2) = \tfrac{1}{2N}\sum_{i=1}^{N}(2\theta) = \tfrac{2\theta N}{2N} = \theta \end{aligned}$$

D. Question 5

A store has on average $E[X] = 100$ customers per day with variance $\mathrm{Var}[X] = 225$. By using Chebyshev's inequality find upper bound on the number of customers X being more than 120 or less than 80.

Answer: The Chebyshev's inequality is

$$P(|X - E[X]| \geq b) \leq \frac{\mathrm{Var}[X]}{b^2}$$

In this case, direct application yields,

$$P(|X - 100| \geq 20) \leq \frac{225}{400} = \frac{9}{16}$$

E. Question 6

Assume that

$$X \sim Geometric(p)$$

where Geometric distribution refers to the probability distribution of the number X of Bernoulli trials necessary to get one success $(1, 2, 3, \cdots)$. Use Markov inequality to find upper bound for

$$P(X \geq a)$$

for a positive integer a. Also find the exact value for $P(X \geq a)$. The probability mass function (PMF) for Geometric distribution is

$$P(X = k) = (1 - p)^{k-1}p$$

for $k = 1, 2, 3, \cdots$.

Answer: The mean of geometric distribution is

$$E[X] = \frac{1}{p}$$

The generic Markov inequality for any non-negative random variable X is

$$P(X \geq a) \leq \frac{E[X]}{a}$$

where $a > 0$. When applied to geometric distribution, we directly get that

$$P(X \geq a) \leq \frac{1}{pa}$$

To get the exact value we use symbolic summation (symsym) in MATLAB:

```
syms p k a positive
assume(a,{'positive','integer'});
assume(k,{'positive','integer'});
symsum((1-p)^(k-1)*p,k,a,Inf)
```

and get that

$$P(X \geq a) = (1-p)^{a-1}$$

F. *Question 7*

Assume that X is continuous random variable with probability density function (PDF)

$$f(x) = \begin{cases} x^2 + \frac{2}{3} & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- Find $E[X^n]$, for $n = 1, 2, 3, \cdots$
- Find variance of X

Answer: The raw moments are

$$E[X^n] = \int_0^1 x^n \left(x^2 + \frac{2}{3}\right) dx = \int_0^1 x^{n+2} dx + \frac{2}{3}\int_0^1 x^n dx = \frac{1}{n+3} + \left(\frac{2}{3}\right)\frac{1}{n+1}$$

for $n = 1, 2, \cdots$.

The variance can be obtained with the raw moments:

$$\begin{aligned} &Var[X] = E[X^2] - (E[X])^2 = \tfrac{1}{5} + \tfrac{2}{9} - \left(\tfrac{1}{4} + \tfrac{2}{6}\right)^2 \\ &= \tfrac{1}{5} + \tfrac{2}{9} - \left(\tfrac{1}{4} + \tfrac{2}{6}\right)^2 = \tfrac{59}{720} \end{aligned}$$

G. *Question 8*

Assume that X is continuous random variable with probability density function (PDF)

$$f(x) = \begin{cases} \frac{5}{32}x^4 & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Now let $Y = X^2$.

- Find Cumulative Distribution Function (CDF) of Y
- Find the PDF of Y
- Find $E[Y]$

Answer: When $0 \le y \le 4$, we get

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P\left(0 \le X \le \sqrt{y}\right)$$
$$= \int_0^{\sqrt{y}} \tfrac{5}{32} x^4 dx = \tfrac{1}{32} y^{5/2}$$

The complete answer for all cases is:

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{1}{32} y^{5/2} & 0 \le y \le 4 \\ 1 & y > 4 \end{cases}$$

Now the PDF is obtained as the derivative of the CDF

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{5}{64} y^{3/2} & 0 \le y \le 4 \\ 0 & y > 4 \end{cases}$$

Finally the mean is

$$E[Y] = \frac{5}{64} \int_0^4 y \cdot y^{3/2} dy = \frac{5}{64} \int_0^4 y^{5/2} dy = \frac{5}{64} \frac{2}{7} 4^{7/2} = \frac{5}{64} \frac{2}{7} 128 = \frac{20}{7} \approx 2.9$$

H. *Question 9*

Assume that Y is random variable and that its probability density function (PDF) is

$$f_Y(y) = \begin{cases} 1 + y & -1 \le y \le 0 \\ y & 0 < y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

- Find $P\left(|Y| < \frac{1}{2}\right)$
- Find $P\left(Y > 0 | Y < \frac{1}{2}\right)$
- Find $E[Y]$

Answer:

$$P\left(|Y| < \frac{1}{2}\right) = \int_0^{1/2} y dy + \int_{-1/2}^{0} (1 + y)\, dy = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$$

By the definition of conditional probability:

$$P\left(Y > 0 | Y < \frac{1}{2}\right) = \frac{P\left(0 < Y < \frac{1}{2}\right)}{P\left(Y < \frac{1}{2}\right)} = \frac{\int_0^{1/2} y dy}{\frac{1}{2} + \int_0^{1/2} y dy} = \frac{1/8}{5/8} = \frac{1}{5}$$

The expected value is

$$E[Y] = \int_0^1 y^2 dy + \int_{-1}^0 y(1 + y)\, dy = \int_{-1}^1 y^2 dy + \int_{-1}^0 y dy = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$$

I. *Question 10*

Assume that X is random variable that is uniformly distributed between 0 and 1. Let $Y = \exp(-X)$.

- What is the cumulative distribution function of Y?
- What is the probability density function of Y?
- Find $E[Y]$.

Remember to handle all possible values as input to PDF and CDF.

Answer: First, we notice that range of possible values for Y is $[1/e, 1]$. The CDF is $F_Y(y) = P(Y \leq y)$. We can easily see that when $y > 1$, the CDF is one and when $y < 1/e$, CDF is zero (since those values are impossible). Now let us focus on the case where y is in the range $[1/e, 1]$:

$$F_Y(y) = P(Y \leq y) = P(\exp(-X) \leq y) = P(-X \leq \ln(y)) = P(X \geq -\ln(y))$$
$$= 1 - P(X < -\ln(y)) = 1 - (-\ln(y)) = 1 + \ln(y)$$

where we have used the fact that $-\ln(y)$ is between 0 and 1. Therefore, the overall CDF is:

$$F_Y(y) = \begin{cases} 0 & y < \frac{1}{e} \\ 1 + \ln(y) & \frac{1}{e} \leq y \leq 1 \\ 1 & y > 1 \end{cases}$$

The PDF is simply derivative of the CDF:

$$f_Y(y) = \begin{cases} 0 & y < \frac{1}{e} \\ \frac{1}{y} & \frac{1}{e} \leq y \leq 1 \\ 0 & y > 1 \end{cases}$$

And the expected value is

$$E[Y] = \int_{1/e}^{1} y\frac{1}{y}dy = \int_{1/e}^{1} 1dy = 1 - \frac{1}{e}$$

J. *Question 11*

John went fishing. He knows that it takes 1 hour to get a fish (on average).

- Find upper bound for probability that it takes more than 4 hours to catch a fish.
- John has only two hours available for fishing. Find upper bound on the probability that he will not catch fish.

Answer: We directly use the Markov inequality. Let X denote the time it takes to catch a fish. Clearly X is non-negative. Now,

$$P(X \geq 4) \leq \frac{E[X]}{4} = \frac{1}{4}$$

For the second question,

$$P(\text{nofish}) = P(X > 2) = P(X \geq 2)$$

and again using the Markov inequality we get

$$P(X \geq 2) \leq \frac{E[X]}{2} = \frac{1}{2}$$

II. Multivariate Distributions

A. Question 1

Random variables x and y have the probability distribution

$$f(x,y) = \begin{cases} 8xy & 0 \le x \le 1, 0 \le y \le x \\ 0 & \text{otherwise} \end{cases}$$

- Find marginal distribution of x
- Find marginal distribution of y
- Find conditional distribution of x
- Find conditional distribution of y

Answer: The marginal distribution of x is

$$f(x) = \int_0^x 8xydy = 8x \int_0^x ydy = 8x\frac{x^2}{2} = 4x^3$$

The marginal distribution of y is obtained by noticing that for given y, the valid range for x is between y and 1. Then we simply integrate and get

$$f(y) = \int_y^1 8xydx = 8y \int_y^1 xdx = 8y\left(\frac{1}{2} - \frac{y^2}{2}\right) = 4y\left(1 - y^2\right)$$

To find the conditional distribution of x we use

$$f_{x|y}(x|y) = \frac{f(x,y)}{f(y)} = \frac{8xy}{4y(1-y^2)} = \frac{2x}{1-y^2}$$

where $0 \le y \le 1, y \le x \le 1$. Similarly, the conditional distribution of y is

$$f_{y|x}(y|x) = \frac{f(x,y)}{f(x)} = \frac{8xy}{4x^3} = \frac{2y}{x^2}$$

when $0 \le x \le 1, 0 \le y \le x$.

B. Question 2

Random variables x and y have the probability distribution

$$f(x,y) = \begin{cases} \frac{3}{2} & x^2 \le y \le 1, 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find conditional distribution of y.

Answer: First, we need to get marginal distribution of x. It is

$$f(x) = \int_{x^2}^1 \frac{3}{2}dy = \frac{3}{2}\left(1 - x^2\right)$$

where $0 < x < 1$. Now directly the conditional distribution is

$$f_{y|x}(y|x) = \frac{f(x,y)}{f(x)} = \frac{1}{1-x^2}$$

where $0 < x < 1, x^2 \le y \le 1$.

C. Question 3

Let

$$x(n) = sin(\omega_0 n + \phi_0),$$

be a real-valued discrete-time random process so that n is an integer. The ϕ_0 is a random variable, uniformly distributed in the range [-π, π] and all samples share the same ϕ_0. ω_0 is constant.

Show that $x(n)$ is zero mean and show that the autocovariance sequence of $x(n)$ is

$$r_x(k) = \frac{1}{2}\cos(\omega_0 k).$$

Answer: First, we show that $x(n)$ has mean zero:

$$E\left[x(n)\right] = \frac{1}{2\pi}\int_{-\pi}^{\pi} sin(\omega_0 n + \phi_0) d\phi_0 = 0$$

To find autocovariance we first note the trigonometric identity

$$\sin\left(\theta\right)\sin\left(\varphi\right) = \frac{\cos\left(\theta - \varphi\right) - \cos\left(\theta + \varphi\right)}{2}$$

Now we find autocovariance or equivalently autocorrelation (since mean is zero) with

$$\begin{aligned}
&r_x\left(k\right) = E\left[x(n+k)x\left(n\right)\right] = E\left[\sin(\omega_0\left(n+k\right) + \phi_0)\sin(\omega_0 n + \phi_0)\right] \\
&= \tfrac{1}{2}E\left[\cos\left(\omega_0 k\right)\right] - \tfrac{1}{2}E\left[\cos\left(\omega_0\left(2n+k\right) + 2\phi_0\right)\right] \\
&= \tfrac{1}{2}\cos\left(\omega_0 k\right) - \tfrac{1}{2}E\left[\cos\left(\omega_0\left(2n+k\right) + 2\phi_0\right)\right] \\
&= \tfrac{1}{2}\cos\left(\omega_0 k\right) - \tfrac{1}{4\pi}\int_{-\pi}^{\pi}\cos\left(\omega_0\left(2n+k\right) + 2\phi_0\right)d\phi_0 \\
&= \tfrac{1}{2}\cos\left(\omega_0 k\right)
\end{aligned}$$

which does not depend on n.

D. Question 4

Let $X \sim Binomial(n, p)$ where n is the number of trials and p is the success probability. Find $E[X]$ and Var$[X]$. Hint: decompose the binomial random variable into sum of independent Bernoulli random variables.

Answer:

$$X = X_1 + X_2 + \cdots + X_n$$

where each X_i is a Bernoulli random variable with success probability p. Now, to get mean we use the linearity of expectation (would work even if the variables would not be independent!)

$$E\left[X\right] = E\left[X_1\right] + E\left[X_2\right] + \cdots + E\left[X_n\right] = nE\left[X_i\right]$$

and since the X_is are independent we can also sum the variances (note that generally we cannot do this!):

$$Var\left[X\right] = Var\left[X_1\right] + Var\left[X_2\right] + \cdots + Var\left[X_n\right] = nVar\left[X_i\right]$$

Now all that is left is to find the mean and variance for one Bernoulli distributed random variable. Since each Bernoulli random variable is 1 with probability p,

$$E\left[X_i\right] = p$$

and the mean of the Binomial random variable is

$$E\left[X\right] = np$$

The variance of each Bernoulli random variable is

$$Var\left[X_i\right] = E\left[X_i^2\right] - E[X_i]^2 = p - p^2 = p\left(1-p\right)$$

and so the variance of the Binomial random variable is

$$Var\left[X\right] = np\left(1-p\right)$$

E. *Question 5*

Assume that random variables x and y have joint probability density function (PDF)

$$f_{xy}(x,y) = \begin{cases} x^2 + \frac{1}{3}y & -1 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the conditional PDF of x given y. Also determine if x and y are independent.

Answer: First we need to find marginal PDF of y

$$f_y(y) = \int_{-1}^{1} \left(x^2 + \frac{1}{3}y\right) dx = \frac{2}{3}(y+1)$$

Now, we directly get that

$$f_{x|y}(x|y) = \frac{f_{xy}(x,y)}{f_y(y)} = \frac{x^2 + \frac{1}{3}y}{\frac{2}{3}(y+1)}$$

where $-1 \le x \le 1, 0 \le y \le 1$ (otherwise conditional PDF is zero).

To check for independence let us find the marginal PDF of x

$$f_x(x) = \int_{0}^{1} \left(x^2 + \frac{1}{3}y\right) dy = x^2 + \frac{1}{6}$$

Now let us check if the joint PDF is product of the marginal PDFs

$$f_x(x)\, f_y(y) = \frac{y}{9} + \frac{2x^2y}{3} + \frac{2x^2}{3} + \frac{1}{9} \neq f_{xy}(x,y)$$

we can see that x and y are not independent.

F. *Question 6*

Assume two random variables x and y that have joint probability density function (PDF)

$$f_{xy}(x,y) = \begin{cases} \frac{1}{2}(3x+y) & 0 \le x, y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Define random vector

$$\mathbf{U} = \begin{bmatrix} X \\ Y \end{bmatrix}$$

Find the mean and covariance matrix of $\mathbf{U}$.

Answer: First, let us find the marginal PDFs

$$f_x(x) = \begin{cases} \int_0^1 \frac{1}{2}(3x+y)\, dy = \frac{3x}{2} + \frac{1}{4} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_y(y) = \begin{cases} \int_0^1 \frac{1}{2}(3x+y)\, dx = \frac{y}{2} + \frac{3}{4} & 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Now, the mean vector is

$$E[\mathbf{U}] = \begin{bmatrix} \int_0^1 x f_x(x)\, dx \\ \int_0^1 y f_y(y)\, dy \end{bmatrix} = \begin{bmatrix} \frac{5}{8} \\ \frac{13}{24} \end{bmatrix}$$

The covariance matrix is

$$\mathbf{C} = \begin{bmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(Y,X) & \mathrm{Var}(Y) \end{bmatrix}$$

where

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X) = E[XY] - E[X]E[Y]$$

We find that

$$E\left[X^2\right] = \int_0^1 x^2 f_x(x)\,dx = \frac{11}{24}$$

and

$$E\left[Y^2\right] = \int_0^1 y^2 f_y(y)\,dy = \frac{3}{8}$$

Therefore, the variances are

$$\mathrm{Var}(X) = E\left[X^2\right] - E[X]^2 = \frac{13}{192}$$

and

$$\mathrm{Var}(Y) = E\left[Y^2\right] - E[Y]^2 = \frac{47}{576}$$

The expected value of XY is

$$E[XY] = \int_0^1 \int_0^1 \frac{xy}{2}(3x+y)\,dxdy = \frac{1}{3}$$

Therefore, the covariance is

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = -\frac{1}{192}$$

Finally, we get the covariance matrix:

$$\mathbf{C} = \begin{bmatrix} \frac{13}{192} & -\frac{1}{192} \\ -\frac{1}{192} & \frac{47}{576} \end{bmatrix}$$

G. Question 7

Assume that X_1 and X_2 are random variables with joint probability density function (PDF)

$$f(x_1,x_2) = \exp(-x_1 - x_2),\, 0 < x_1 < \infty, 0 < x_2 < \infty$$

Let us consider transformation of random variables:

$$Y_1 = X_1 - X_2$$

and

$$Y_2 = X_1 + X_2$$

Find the joint PDF of Y_1 and Y_2.

Answer: We immediately notice that

$$X_1 = \frac{Y_1 + Y_2}{2}$$

and

$$X_2 = \frac{Y_2 - Y_1}{2}$$

Now, we know from the theory of transformation of random variables

$$f\left(y_1, y_2\right) = \exp\left(-\left(\frac{Y_1 + Y_2}{2}\right) - \left(\frac{Y_2 - Y_1}{2}\right)\right)|J| = \exp\left(-Y_2\right)|J|$$

where the Jacobian is

$$J = \det\begin{bmatrix} \frac{\partial X_1}{\partial Y_1} & \frac{\partial X_1}{\partial Y_2} \\ \frac{\partial X_2}{\partial Y_1} & \frac{\partial X_2}{\partial Y_2} \end{bmatrix} = \det\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} = \frac{1}{2}$$

Finally, we get

$$f\left(y_1, y_2\right) = \frac{\exp\left(-y_2\right)}{2}, 0 < y_1 + y_2 < \infty, 0 < y_2 - y_1 < \infty$$

If we let y_2 vary from zero to infinity, we get the limits

$$f\left(y_1, y_2\right) = \frac{\exp\left(-y_2\right)}{2}, 0 < y_2 < \infty, -y_2 < y_1 < y_2$$

Let us verify the limits by checking that the PDF integrates to one:

```
syms y1 y2 real
int(int( exp(-y2)/2,y1,-y2,y2),y2,0,Inf)
ans =
1
```

III. Linear algebra

A. Question 1

Assume that matrix $\mathbf{H}$

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

and vector $\mathbf{x}$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Calculate by hand

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x}$$

Answer:

$$\mathbf{H}^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\mathbf{H}^T\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = 4\mathbf{I}$$

$$\left(\mathbf{H}^T\mathbf{H}\right)^{-1} = \tfrac{1}{4}\mathbf{I}$$

$$\mathbf{H}^T\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ -2 \end{bmatrix}$$

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x} = \tfrac{1}{4} \begin{bmatrix} 10 \\ -2 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -0.5 \end{bmatrix}$$

B. Question 2

Assume that received samples $x[n] = \theta_1 + \theta_2 n + \theta_3 n^2 + w[n]$, $n = 1, 2, 3, 4$. Here $\theta_1, \theta_2, \theta_3$ are the unknown parameters to be estimated, $w[n]$ is white Gaussian noise and n is the time index. Let us define $\boldsymbol{\theta}$ as a vector containing the unknown parameters

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

and let us collect the received samples to a vector

$$\mathbf{x} = \begin{bmatrix} x[1] \\ x[2] \\ x[3] \\ x[4] \end{bmatrix}$$

Find observation matrix $\mathbf{H}$ and vector $\mathbf{w}$ such that $\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$.

Answer: The observation matrix is:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}$$

and

$$\mathbf{w} = \begin{bmatrix} w[1] \\ w[2] \\ w[3] \\ w[4] \end{bmatrix}$$

C. Question 3

Assume that received samples $x[n] = \theta_1 + \theta_2 n + \theta_3 \cos(n)$, $n = 1, 2, 3, 4$. Here $\theta_1, \theta_2, \theta_3$ are the unknown parameters to be estimated and n is the time index. Let us define $\boldsymbol{\theta}$ as a vector containing the unknown parameters

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

and let us collect the received samples to a vector

$$\mathbf{x} = \begin{bmatrix} x[1] \\ x[2] \\ x[3] \\ x[4] \end{bmatrix}$$

Find observation matrix $\mathbf{H}$ such that $\mathbf{x} = \mathbf{H}\boldsymbol{\theta}$.

Answer: The observation matrix is:

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & \cos(1) \\ 1 & 2 & \cos(2) \\ 1 & 3 & \cos(3) \\ 1 & 4 & \cos(4) \end{bmatrix}$$

D. *Question 4*

Assume that the observed data (collected in a vector $\mathbf{x}$) follows the Gaussian distribution where both mean and covariance can depend on the unknown parameters (collected in a vector $\boldsymbol{\theta}$)

$$\mathbf{x} \sim N\left(\mu\left(\boldsymbol{\theta}\right), \mathbf{C}\left(\boldsymbol{\theta}\right)\right)$$

Let us assume that there are p unknown parameters. We know that for this kind of problem the $p \times p$ Fisher Information matrix is given by:

$$\left[\mathbf{I}\left(\boldsymbol{\theta}\right)\right]_{ij} = \left[\frac{\partial \mu\left(\boldsymbol{\theta}\right)}{\partial \theta_i}\right]^T \mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\left[\frac{\partial \mu\left(\boldsymbol{\theta}\right)}{\partial \theta_j}\right] + \frac{1}{2} tr\left[\mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\frac{\partial \mathbf{C}\left(\boldsymbol{\theta}\right)}{\partial \theta_i}\mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\frac{\partial \mathbf{C}\left(\boldsymbol{\theta}\right)}{\partial \theta_j}\right]$$

where tr is the matrix trace. Now assume that we observed data

$$x[n] = w[n],\ n = 0, 1, \cdots, N-1$$

where $w[n]$ is white Gaussian noise with unknown variance $\theta = \sigma^2$. Find the Fisher Information matrix for this problem.

Answer: We notice that

$$\mathbf{x} \sim N\left(0, \mathbf{C}\left(\boldsymbol{\theta}\right)\right)$$

where

$$\mathbf{C}\left(\boldsymbol{\theta}\right) = \theta \mathbf{I}$$

Since mean is always zero (no matter what the unknown parameters are), the first term of the Fisher Information matrix is zero and can be ignored. Also, since there is only one unknown parameter the Fisher Information matrix is a scalar which is:

$$\mathbf{I}\left(\theta\right) = \frac{1}{2} tr\left[\frac{1}{\theta}\mathbf{I}\frac{\partial \theta \mathbf{I}}{\partial \theta}\frac{1}{\theta}\mathbf{I}\frac{\partial \theta \mathbf{I}}{\partial \theta}\right] = \frac{1}{2} tr\left[\frac{1}{\theta^2}\mathbf{I}\right] = \frac{N}{2\theta^2} = \frac{N}{2\sigma^4}$$

E. *Question 5*

We know that when

$$\mathbf{x} \sim N\left(\mu\left(\boldsymbol{\theta}\right), \mathbf{C}\left(\boldsymbol{\theta}\right)\right)$$

the Fisher Information matrix is given by:

$$\left[\mathbf{I}\left(\boldsymbol{\theta}\right)\right]_{ij} = \left[\frac{\partial \mu\left(\boldsymbol{\theta}\right)}{\partial \theta_i}\right]^T \mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\left[\frac{\partial \mu\left(\boldsymbol{\theta}\right)}{\partial \theta_j}\right] + \frac{1}{2} tr\left[\mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\frac{\partial \mathbf{C}\left(\boldsymbol{\theta}\right)}{\partial \theta_i}\mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\frac{\partial \mathbf{C}\left(\boldsymbol{\theta}\right)}{\partial \theta_j}\right]$$

where tr is the matrix trace. Now assume that observed data vector of size 2×1 follows the Gaussian distribution so that

$$\mathbf{x} \sim N\left(0, \mathbf{C}\left(\rho\right)\right)$$

where ρ is the unknown parameters and the covariance matrix is

$$\mathbf{C}\left(\rho\right) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Find the Fisher Information matrix for this problem.

Answer: We know that

$$\mathbf{C}^{-1}\left(\rho\right) = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$$

The mean does not depend on the unknown parameter so the first term of the Fisher Information matrix can be ignored. Also, since there is only one unknown parameter $\theta = \rho$ the second term can be expressed in terms of squaring:

$$\mathbf{I}\left(\boldsymbol{\theta}\right) = \tfrac{1}{2} tr\left[\left(\mathbf{C}^{-1}\left(\boldsymbol{\theta}\right)\tfrac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta}\right)^2\right] = \tfrac{1}{2} tr\left[\left(\tfrac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)^2\right]$$
$$= \tfrac{1}{2} tr\left[\left(\tfrac{1}{1-\rho^2}\begin{bmatrix} -\rho & 1 \\ 1 & -\rho \end{bmatrix}\right)^2\right] = \tfrac{1}{2} tr\left[\left(\tfrac{1}{(1-\rho^2)^2}\begin{bmatrix} \rho^2+1 & -2\rho \\ -2\rho & \rho^2+1 \end{bmatrix}\right)\right]$$
$$= \tfrac{2\left(1+\rho^2\right)}{2(1-\rho^2)^2} = \tfrac{1+\rho^2}{(1-\rho^2)^2}$$

F. *Question 6*

Assume that $N \times N$ covariance matrix is

$$\mathbf{C}\left(\sigma_A^2\right) = \sigma_A^2 \mathbf{1}\mathbf{1}^T + \sigma^2 \mathbf{I}$$

where

$$\mathbf{1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}^T$$

Find inverse of the covariance matrix using Woodbury's identity:

$$\left(\mathbf{A} + \mathbf{u}\mathbf{u}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{u}^T\mathbf{A}^{-1}}{1+\mathbf{u}^T\mathbf{A}^{-1}\mathbf{u}}$$

Answer: We notice that

$$\mathbf{C}^{-1}\left(\sigma_A^2\right) = \left(\sigma^2\mathbf{I} + \sigma_A^2\mathbf{1}\mathbf{1}^T\right)^{-1} = \left(\mathbf{A} + \mathbf{u}\mathbf{u}^T\right)^{-1}$$

where

$$\mathbf{A} = \sigma^2 \mathbf{I}$$
$$\mathbf{u} = \sigma_A \mathbf{1}$$

Using the Woodbury's identity we get

$$\mathbf{C}^{-1}\left(\sigma_A^2\right) = \tfrac{1}{\sigma^2}\mathbf{I} - \frac{\frac{1}{\sigma^4}\mathbf{u}\mathbf{u}^T}{1+\frac{1}{\sigma^2}\mathbf{u}^T\mathbf{u}} = \tfrac{1}{\sigma^2}\mathbf{I} - \frac{\frac{\sigma_A^2}{\sigma^4}\mathbf{1}\mathbf{1}^T}{1+\frac{\sigma_A^2}{\sigma^2}\mathbf{1}^T\mathbf{1}} = \tfrac{1}{\sigma^2}\mathbf{I} - \frac{\frac{\sigma_A^2}{\sigma^4}\mathbf{1}\mathbf{1}^T}{1+\frac{N\sigma_A^2}{\sigma^2}}$$
$$= \tfrac{1}{\sigma^2}\mathbf{I} - \frac{\sigma_A^2\mathbf{1}\mathbf{1}^T}{\sigma^4+N\sigma^2\sigma_A^2} = \tfrac{1}{\sigma^2}\left(\mathbf{I} - \frac{\sigma_A^2}{\sigma^2+N\sigma_A^2}\mathbf{1}\mathbf{1}^T\right)$$

G. Question 7

We want to find

$$\widehat{\boldsymbol{\theta}} = \frac{\mathbf{s}^T\mathbf{C}^{-\mathbf{1}}\mathbf{x}}{\mathbf{s}^T\mathbf{C}^{-\mathbf{1}}\mathbf{s}}$$

where

$$\mathbf{s} = \begin{bmatrix} s[0] & s[1] & \cdots & s[N-1] \end{bmatrix}^T$$

and

$$\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \cdots & x[N-1] \end{bmatrix}^T$$

What happens if s is eigenvector of the $N \times N$ covariance matrix $\mathbf{C}$? Notice that eigenvectors of matrix and its inverse are the same (and the eigenvalues are inverse of each other). Also notice that covariance matrix is symmetric.

Answer: We know that

$$\mathbf{C}\mathbf{s} = \lambda\mathbf{s}$$

where λ is the corresponding eigenvalue. Therefore

$$\mathbf{C}^{-\mathbf{1}}\mathbf{s} = \frac{1}{\lambda}\mathbf{s}$$

By noticing that transpose of a scalar is the original value we get that

$$\widehat{\boldsymbol{\theta}} = \frac{\mathbf{s}^T\mathbf{C}^{-\mathbf{1}}\mathbf{x}}{\frac{1}{\lambda}\mathbf{s}^T\mathbf{s}} = \frac{\left(\mathbf{s}^T\mathbf{C}^{-\mathbf{1}}\mathbf{x}\right)^T}{\frac{1}{\lambda}\mathbf{s}^T\mathbf{s}} = \frac{\mathbf{x}^T\mathbf{C}^{-\mathbf{1}}\mathbf{s}}{\frac{1}{\lambda}\mathbf{s}^T\mathbf{s}} = \frac{\frac{1}{\lambda}\mathbf{x}^T\mathbf{s}}{\frac{1}{\lambda}\mathbf{s}^T\mathbf{s}} = \frac{\mathbf{x}^T\mathbf{s}}{\mathbf{s}^T\mathbf{s}}$$

References

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.