MATH 076 - Computational Inverse Problems

Carter Kruse, August 9, 2023

Homework 3

Instructions

For the non-computational questions, you are **not** permitted to use a calculator or any other computational tools, unless it is to check your work. For the computational questions, you are asked to use a coding language of your choice to perform the tasks requested. You may make use any code posted to our course's Canvas page. **Please show all of your work.** If you have any questions or uncertainties, please reach out to your instructor.

The label [**CQ**] indicates a computational question that is to be solved using a computational tool such as MATLAB or Python.

Questions

1) Let $X \in \mathbb{R}$ be a uniformly distributed random variable between 0 and 1. Define a new random variable Z such that

$$Z = 2X^3 + 3$$

(a) What is the probability density function of Z?

Given $X \in \mathbb{R}$ is a uniformly distributed random variable between 0 and 1, the random variable Z is defined as follows:

$$Z = 2X^3 + 3$$

To determine the probability density function of Z, we must perform a change of variables, keeping in mind that we are working with probability densities. The method is outlined as follows:

Assume that we have two real-valued random variables $X \in \mathbb{R}$ an $Z \in \mathbb{R}$ that are related to each other through a functional relation $X = \phi(Z)$ where $\phi : \mathbb{R} \to R$ is a one-to-one mapping.

For simplicity, assume that ϕ is strictly increasing and differentiable, so that $\phi'(z) > 0$. If the probability density function $\pi(X)$ of X is given, what is the corresponding density $\pi(Z)$ of Z?

First, note that since ϕ is increasing, for any values $a < b$, we have

$$a < Z < b \text{ if and only if } a' = \phi(a) < \phi(z) = X < \phi(b) = b'$$

Therefore, $P\{a' < X < b'\} = P\{a < X < b\}$. Equivalently, the probability density of Z satisfies

$$\int_a^b \pi_Z(z)\, dz = \int_{a'}^{b'} \pi_X(x)\, dx$$

Performing a change of variables in the integral on the right $x = \phi(z)$ and $dx = \frac{d\phi}{dz}(z)\, dz$, we obtain

$$\int_a^b \pi_Z(z)\, dz = \int_a^b \pi_X(\phi(z)) \frac{d\phi}{dz}(z)\, dz$$

This holds for all a and b, and therefore we arrive at the conclusion that

$$\pi_Z(z) = \pi_X(\phi(z)) \frac{d\phi}{dz}(z)$$

In the derivation above, we assumed that ϕ was increasing. If it is decreasing, the derivative is negative. In general, since the density needs to be non-negative, we write

$$\pi_Z(z) = \pi_X(\phi(z)) \left| \frac{d\phi}{dz}(z) \right|$$

Using this method, we may determine the probability density function of Z. The functional relation $X = \phi(Z)$ may be determined using the expression $Z = 2X^3 + 3$, so $X = \phi(Z) = \left(\frac{1}{2}(Z-3)\right)^{\frac{1}{3}}$.

Given that $X \in \mathbb{R}$ is a uniformly distributed random variable between 0 and 1, $\pi_X(x) = \int_0^1 1\, dx$. Thus, we compute $\pi_Z(z) = \pi_X(\phi(z)) \frac{d\phi}{dz}(z) \rightarrow \pi_Z(z) = \frac{d\phi}{dz}(z)$, as $\pi_X(\phi(z)) = 1$.

Using our expression for $X = \phi(Z)$, we have

$$\pi_Z(z) = \frac{1}{6}\left(\frac{1}{2}(Z-3)\right)^{-\frac{2}{3}}$$

The bounds of the integral for the uniformly distributed random variable $X \in \mathbb{R}$ are $[0, 1]$, while the bounds of the integral for the random variable $Z \in \mathbb{R}$ are $[3, 5]$. Thus, we must adjust the probability density function to be piecewise.

$$\pi_Z(z) = \begin{cases} \frac{1}{6}\left(\frac{1}{2}(z-3)\right)^{-\frac{2}{3}} & \text{if } 3 \leq z \leq 5 \\ 0 & \text{else} \end{cases}$$

(b) **[CQ]** Generate random samples of Z and plot them in a histogram. Plot also the probability density function you calculated in (a). Compare these results.

View *Matlab* Code

2) Consider the forward model

$$Y = E \cdot \left(2X^2 + 1\right)$$

where $Y \in \mathbb{R}$ is the data, $X \in \mathbb{R}$ is the unknown, and $E \in \mathbb{R}^+$ is Gamma distributed noise with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$. Suppose X and E are mutually independent. Hence, the probability density function π_E is

$$\pi_E(\varepsilon) = \frac{\beta^\alpha}{\Gamma(\alpha)} \varepsilon^{\alpha-1} e^{-\beta\varepsilon}$$

where $\Gamma(\cdot)$ is the gamma function.

(a) What is the likelihood density function $\pi(y|x)$?

In the case where the forward model represents multiplicative noise, the following method is used to determine the likelihood density function.

Say we have a model $Y = Ef(X)$, where $Y, E, X \in \mathbb{R}$ and E and X are mutually independent.

It follows that

$$\begin{aligned}
\pi(y|x) &= \int_{\mathbb{R}} \delta(y - \varepsilon f(x))\, \pi_{noise}\, d\varepsilon \\
&= \frac{1}{f(x)} \int_{\mathbb{R}} \delta(y - v)\, \pi_{noise}\left(\frac{v}{f(x)}\right) dv \qquad\qquad v = \varepsilon f(x) \text{ and } dv = f(x)\, d\varepsilon \\
&= \frac{1}{f(x)} \pi_{noise}\left(\frac{y}{f(x)}\right)
\end{aligned}$$

Thus, given the forward model $Y = E \cdot \left(2X^2 + 1\right)$, we have $f(x) = \left(2x^2 + 1\right)$. Further, $Y \in \mathbb{R}$ is the data, $X \in \mathbb{R}$ is the unknown, and $E \in \mathbb{R}^+$ is Gamma distributed noise with shape parameter $\alpha > 0$ and rate parameter $\beta > 0$. The variables X and E are mutually independent.

The probability density function π_E is

$$\pi_E(\varepsilon) = \frac{\beta^\alpha}{\Gamma(\alpha)} \varepsilon^{\alpha-1} e^{-\beta\varepsilon}$$

where $\Gamma(\cdot)$ is the gamma function.

With this information, we may determine the likelihood density function as follows:

$$\pi(y|x) = \frac{1}{f(x)} \pi_{noise}\left(\frac{y}{f(x)}\right)$$

$$\pi(y|x) = \frac{1}{2x^2+1} \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{y}{2x^2+1}\right)^{\alpha-1} e^{-\beta\left(\frac{y}{2x^2+1}\right)} \right)$$

(b) **[CQ]** Given $\alpha = 2$ and $\beta = \frac{1}{2}$, plot $\pi(y|x)$ when $x = \frac{1}{10}$, $x = \frac{1}{2}$, and $x = 1$.

View *Matlab* Code

3) Let $X \in \mathbb{R}$ be a random variable with realization x whose probability density function is given as

$$\pi(x) = \frac{\alpha}{\pi + \pi\alpha^2 x^2}$$

where $\alpha > 0$ is some given constant.

(a) Derive an equation, not containing an integral, for the cumulative density function of X.

The cumulative density function of a real-valued random variable is defined as

$$\Phi_X(x) = \int_{-\infty}^{x} \pi_X(x')\, dx'$$

Thus, if $X \in \mathbb{R}$ is a random variable with realization x whose probability density function is given as $\pi(x) = \frac{\alpha}{\pi+\pi\alpha^2 x^2}$ where $\alpha > 0$ is some given constant, we may integrate this to find the cumulative density function.

To compute the integral

$$\Phi_X(x) = \int_{-\infty}^{x} \frac{\alpha}{\pi + \pi\alpha^2 t^2}\, dt$$

we may use a u-substitution, with $u = \alpha t$ and $du = \alpha\, dt$.

$$\Phi_X(x) = \int \frac{1}{\pi(1+u^2)}\, du = \frac{1}{\pi}\int \frac{1}{(1+u^2)}\, du$$

Evaluating the integral, we have

$$\Phi_X(x) = \frac{1}{\pi}\arctan(u) = \frac{1}{\pi}\arctan(\alpha t)$$

Considering the bounds of the integral ($-\infty$ and x), we have

$$\Phi_X(x) = \frac{1}{\pi}\arctan(\alpha x) - \frac{1}{\pi}\arctan(-\infty) = \frac{1}{\pi}\arctan(\alpha x) - \frac{1}{\pi}\left(-\frac{\pi}{2}\right)$$

This is equivalent to the following, which provides the cumulative density function of X:

$$\Phi_X(x) = \frac{1}{\pi}\arctan(\alpha x) + \frac{1}{2}$$

(b) Assuming only that you have the ability to generate uniformly distributed random numbers between 0 and 1, describe a method for generating random samples of X.

Assuming that we only have the ability to generate uniformly distribute random numbers between 0 and 1, we may generate random samples of X by taking the inverse of the cumulative density function.

Given the cumulative density function $\frac{1}{\pi}\arctan(\alpha x) + \frac{1}{2}$, which has a bounded range from 0 to 1, let us consider the following:

$$\frac{1}{\pi}\arctan(\alpha x) + \frac{1}{2} = u$$

where u is a random variable that is uniformly distribute between 0 and 1.

Solving for x, we find that

$$x = \frac{1}{\alpha}\tan\left(\pi\left(u - \frac{1}{2}\right)\right)$$

In other words, given a realization u of the uniformly distributed random variable U over the domain of $[0, 1]$, we may generate random samples of x by using the inverse of the cumulative density function.

(c) **[CQ]** Write code that generates random samples of X for $\alpha = 1$.

View *Matlab* Code

4) Suppose $f : [0,1] \to \mathbb{R}$ is an unknown function. Assume that we have prior knowledge that f is almost piecewise-linear and that $f(x) = 0$ for $0 \le x \le 0.2$. By saying f is almost piecewise-linear, we mean that the second derivative of f is relatively small with a few possible outliers. Let us discretize the interval $[0,1]$ by the points $t_j = \frac{j}{N}$, where $0 \le j \le N$ and $N \ge 10$, and write $x_j = f(t_j)$.

(a) Describe a probability density function $\pi_{pr}(x)$ that would incorporate the prior assumptions on x_j for $j = 1, \ldots, N$.

Hint: Consider the finite difference approximation of the second derivative given by $f''(t_j) = N^2 (f(t_{j-1}) - 2f(t_j) + f(t_{j+1}))$. Also, you may need to consider using a piecewise function for $\pi_{pr}(\boldsymbol{x})$.

Given the unknown function $f : [0,1] \to \mathbb{R}$, we may create a probability density function $\pi_{pr}(x)$ that incorporates the prior assumptions on x_j for $j = 1, \ldots, N$.

The prior knowledge is that f is almost piecewise-linear and that $f(x) = 0$ for $0 \le x \le 0.2$. By saying f is almost piecewise-linear, we mean that the second derivative of f is relatively small with a few possible outliers.

Let us discretize the interval $[0,1]$ by the points $t_j = \frac{j}{N}$, where $0 \le j \le N$ and $N \ge 10$, and write $x_j = f(t_j)$.

Further, let us consider the finite difference approximation of the second derivative given by $f''(t_j) = N^2 (f(t_{j-1}) - 2f(t_j) + f(t_{j+1}))$.

With this information, we may perform a similar method as done in class, specifically when we considered a value X with few discontinuities, where our goal was to estimate the function $f : [0,1] \to \mathbb{R}$, with $f(0) = 0$ from indirect observations.

By discretizing the interval $[0,1]$ by the points $t_j = \frac{j}{N}$, where $0 \le j \le N$ and $N \ge 10$, and writing $x_j = f(t_j)$, we are able to write the finite difference approximation of the derivative as

$$f''(t_j) = N^2 (x_{j-1} - 2x_j + x_{j+1})$$

The Cauchy density function is given as

$$\Phi(t) = \frac{2\alpha}{\pi} \int_0^t \frac{1}{1 + \alpha^2 s^2} \, dx = \frac{2}{\pi} \arctan(\alpha t)$$

Thus, to determine the value of x_j, we may use the expression

$$x_j = \Phi^{-1}(t_j) = \frac{1}{\alpha} \tan\left(\frac{\pi}{2} t_j\right)$$

As in class, we may input $x_{j-1} - 2x_j + x_{j+1}$ into the Cauchy density, which results in

$$\pi_{pr}(x) = \left(\frac{\alpha}{\pi}\right)^N \prod_{j=1}^{N} \frac{1}{1 + \alpha^2 (x_{j-1} - 2x_j + x_{j+1})^2}$$

This means that there is no longer independent densities, so we must account for this.

Further, recall that we were given the prior information that $f(x) = 0$ for $0 \leq x \leq 0.2$. This information may be encoded using a piecewise probability density function, or alternatively, using a combination of delta functions, as follows:

$$\pi_{pr}(x) = \delta(x_1)\,\delta(x_2)\cdots\delta(x_J)\left(\frac{\alpha}{\pi}\right)^N \prod_{j=J+1}^{N} \frac{1}{1+\alpha^2\left(x_{j-1}-2x_j+x_{j+1}\right)^2}$$

The number of J values for which the delta function is included depends on N, as we recall that $0 \leq x \leq 0.2$ is the condition, thus, the first 20% of the data should be encoded with zeros.

(b) Derive a method for generating samples from $\pi_{pr}(\boldsymbol{x})$.

To generating samples from $\pi_{pr}(\boldsymbol{x})$, consider the following. Given that the densities are no longer independent, we must define a value $\zeta_j = x_{j-1} - 2x_j + x_{j+1}$. The values ζ_j are independent, so we can sample as before and then use the expression for ζ_j to determine the vector x.

This method follows that covered from the textbook, which is provided below for reference:

> The impulse prior densities discussed above provide a tool for implementing this type of information in the prior. We consider here a one-dimensional example. Let us assume that the goal is to estimate a function $f : [0,1] \to \mathbb{R}$, $f(0) = 0$ from indirect observations. Our prior knowledge is that the function f may have a large jump at few places in the interval, but the locations of these places are unkown to us. One possible way to construct a prior in this case is to consider the finite difference approximation of the derivative of f and assume that it follows an impulse noise probability distribution. Let us discretize the interval $[0,1]$ by points $t_j = j/N$, $0 \leq j \leq N$ and write $x_j = f(t_j)$. Consider the density
>
> $$\pi_{\mathrm{pr}}(x) = \left(\frac{\alpha}{\pi}\right)^N \prod_{j=1}^{N} \frac{1}{1+\alpha^2(x_j - x_{j-1})^2}.$$
>
> It is again instructive to see what random draws from this distribution look like. To perform the drawing, let us define new random variables
>
> $$\xi_j = x_j - x_{j-1}, \quad 1 \leq j \leq N. \tag{3.6}$$
>
> The probability distribution of these variables is
>
> $$\pi(\xi) = \left(\frac{\alpha}{\pi}\right)^N \prod_{j=1}^{N} \frac{1}{1+\alpha^2\xi_j^2},$$
>
> that is, they are independent from each other. Hence each ξ_j can be drawn from the one-dimensional Cauchy density. Having these mutually independent increments at hand, the vector x is determined from (3.6). Figure 3.4 shows four random draws from this distribution with discretization $N = 1200$.

In other words, we may draw ζ_j from a one-dimensional Cauchy density, as the values are mutually independent. From this, we may determine the vector x from the expression $\zeta_j = f(x)$.

The discretization characteristic N determines our sampling, and with the prior knowledge that $f(x) = 0$ for $0 \leq x \leq 0.2$, this means that with $N = 100$, the first 20 samples of x_i should be 0, hence the piecewise/delta representation.

(c) **[CQ]** Write code to generate random samples from $\pi_{pr}(\boldsymbol{x})$ when $N = 100$ and $\alpha = 1$.

View *Matlab* Code