Solution to analysis in Home Assignment 1

Yuhong Zhou + yuhong

Analysis

In this report I will present my independent analysis of the questions related to home assignment 1. I have discussed the solution with Xinying Wang but I swear that the analysis written here are my own.

1 Properties of random variables

(a) i)

Since we know that $x \sim \mathcal{N}(\mu, \sigma^2)$, then we can get the probability density function(pdf) of x:

$$p(x) = \mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Then we can show that:

$$\begin{aligned}
\mathbb{E}[x] &= \int_{-\infty}^{\infty} xp(x)\ dx \\
&= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\ dx
\end{aligned}$$

Assume that $t = \frac{(x-\mu)}{2\sigma}$ and $dx = 2\sigma dt$, then:

$$\begin{aligned}
\mathbb{E}[x] &= \int_{-\infty}^{\infty} x\frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\ dx \\
&= \int_{-\infty}^{\infty} (2\sigma t + \mu)\frac{2}{\sqrt{2\pi}} e^{(-2t^2)}\ dt \\
&= \frac{2}{\sqrt{2\pi}}\left(2\sigma\int_{-\infty}^{\infty} te^{(-2t^2)}\ dt + \mu\int_{-\infty}^{\infty} e^{(-2t^2)}\ dt\right)
\end{aligned}$$

Indeed, since te^{-2t^2} is an odd function, with t being odd and e^{-2t^2} being even, its integral over a symmetric interval $[-\infty, \infty]$ evaluates to zero due to symmetry. Thus we can get that $\int_{-\infty}^{\infty} te^{(-2t^2)}\ dt = 0$. Since $\int_{-\infty}^{\infty} e^{(-tx^2)}\ dx = \frac{\sqrt{\pi}}{\sqrt{t}}$, therefore:

$$\begin{aligned}\mathbb{E}[x] &= \frac{2\mu}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{(-2t^2)}\ dt \\ &= \mu\sqrt{\frac{2}{\pi}}\sqrt{\frac{\pi}{2}} \\ &= \mu\end{aligned}$$

ii)
Using the definition of expected value, we can show that:

$$\begin{aligned}\mathrm{Var}[x] &= \mathbb{E}[(x-\mu)^2] \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 p(x)\ dx \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\ dx\end{aligned}$$

Assume that $t = \frac{x-\mu}{\sqrt{2}\sigma}, \quad dx = \sqrt{2}\sigma dt$, then:

$$\begin{aligned}\mathrm{Var}[x] &= \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\ dx \\ &= \int_{-\infty}^{\infty} (\sqrt{2}\sigma t)^2 \frac{1}{\sqrt{2\pi}\ \sigma} e^{(-t^2)}\ \sqrt{2}\sigma dt \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{(-t^2)}\ dt\end{aligned}$$

Let $u = t^2, \quad dt = \frac{1}{2t}du$, we can get that:

$$\begin{aligned}\mathrm{Var}[x] &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{(-t^2)}\ dt \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{1}{2}\sqrt{u}e^{(-u)}du \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \left(\int_{0}^{\infty} \frac{1}{2}\sqrt{u}e^{(-u)}du + \int_{-\infty}^{0} \frac{1}{2}\sqrt{u}e^{(-u)}du\right) \\ &= \frac{\sigma^2}{\sqrt{\pi}} \left(\int_{0}^{\infty} \sqrt{u}e^{(-u)}du + \int_{0}^{\infty} \sqrt{-u}e^{(u)}du\right)\end{aligned}$$

We can use the definition of the gamma function: $\Gamma(z) = \int_0^\infty t^{z-1}e^{(-t)}dt$ and the properties of it: $\Gamma(\frac{3}{2}) = \Gamma(\frac{1}{2}+1) = \frac{(2*1)!}{4^1*1!}\sqrt{\pi} = \frac{\sqrt{\pi}}{2}$ to solve it:

$$\begin{aligned}
\mathrm{Var}[x] &= \frac{\sigma^2}{\sqrt{\pi}}\left(\int_0^\infty \sqrt{u}e^{(-u)}du + \int_0^\infty \sqrt{-u}e^{(u)}du\right) \\
&= \frac{\sigma^2}{\sqrt{\pi}}\left(\Gamma(\frac{3}{2}) + \Gamma(\frac{3}{2})\right) \\
&= \frac{\sigma^2}{\sqrt{\pi}}\left(\frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{2}\right) \\
&= \sigma^2
\end{aligned}$$

(b) i)

By definition of expected value with known probability density function $p(q)$ we can get that:

$$\mathbb{E}(z) = \int zp(q)dq$$

Since $z = Aq$ where **A** is a constant matrix, then:

$$\begin{aligned}
\mathbb{E}(z) &= \int zp(q)dq \\
&= \int Aqp(q)dq \\
&= A\int qp(q)dq \\
&= A\mathbb{E}(q)
\end{aligned}$$

ii)

Using the definition of the covariance, we can know that:

$$\mathrm{Cov}[z] = \mathbb{E}[(z - \mathbb{E}(z))(z - \mathbb{E}(z))^T]$$

Considering the result in i) and $z = Aq$ where **A** is a constant matrix, we can get:

$$\begin{aligned}
\mathrm{Cov}[z] &= \mathbb{E}[(z - \mathbb{E}(z))(z - \mathbb{E}(z))^T] \\
&= \mathbb{E}[(Aq - A\mathbb{E}(q))(Aq - A\mathbb{E}(q))^T] \\
&= \mathbb{E}[A(q - \mathbb{E}(q))(q - \mathbb{E}(q))^T A^T] \\
&= A\mathbb{E}[(q - \mathbb{E}(q))(q - \mathbb{E}(q))^T]A^T
\end{aligned}$$

By the definition of the Cov[q]:$\mathrm{Cov}[q] = \mathbb{E}[(q - \mathbb{E}(q))(q - \mathbb{E}(q))^T]$, then we get:

$$\begin{aligned}
\mathrm{Cov}[z] &= A\mathbb{E}[(q - \mathbb{E}(q))(q - \mathbb{E}(q))^T]A^T \\
&= A\mathrm{Cov}[q]A^T
\end{aligned}$$

(c) The $\mu_z and \Sigma_z$ are shown as follow:

$$\mu_z = \begin{bmatrix} 5 \\ 10 \end{bmatrix}, \quad \Sigma_z = \begin{bmatrix} 2.3000 & 4.0000 \\ 4.0000 & 8.0000 \end{bmatrix}$$

The Figure 1.1 illustrates the mean and covariance of **z** and **q**.

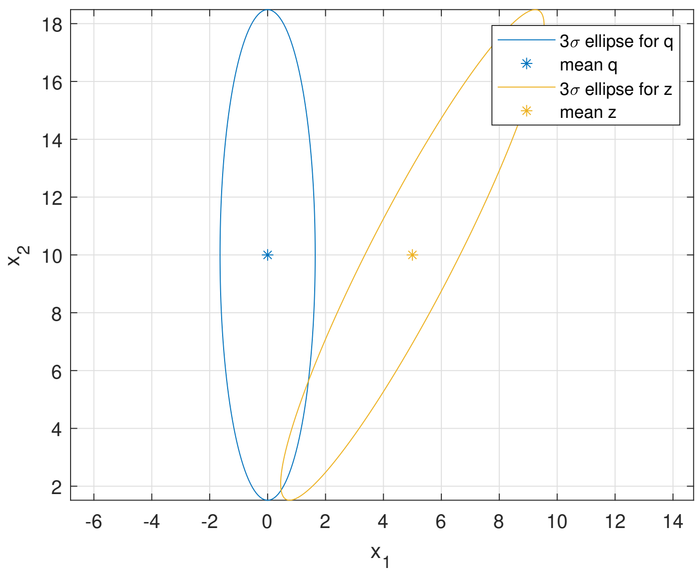

*Figure 1.1: The mean and covariance of **z** and **q**.*

The matrix **A** shifts the mean value of **q** by 5 units (by a factor of 0.5) in the x-direction while keeping it unchanged in the y-direction. The Figure 1.1 illustrates that after the transformation, the axis of the ellipse tilts at an angle towards the right. The variance of the second component remains unchanged after the transformation. Furthermore, the transformation has introduced a cross-covariance between the two components. The transformation can be traced back to the structure of **A** that the matrix **A** is upper diagonal, with diagonal elements all equal to 1. The cross-covariance introduced by the transformation arises from the off-diagonal elements he matrix **A**.

2 Transformation of random variables

(a) We assume that the $p(z)$ is a Gaussian, then we can calculate $\mathbb{E}[z]$ and $\text{Var}[z]$ analytically:

$$\mathbb{E}[z] = \mathbb{E}[3x] = 3\mathbb{E}[x] = 0$$
$$\text{Var}[z] = \text{Var}[3x] = 3^2 * \text{Var}[x] = 18$$

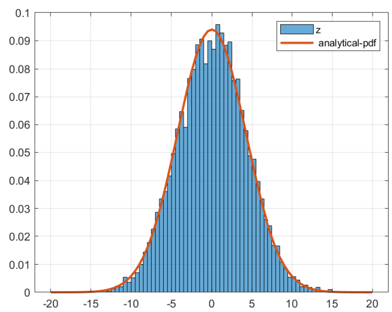

Figure 2.1: Analytical results.

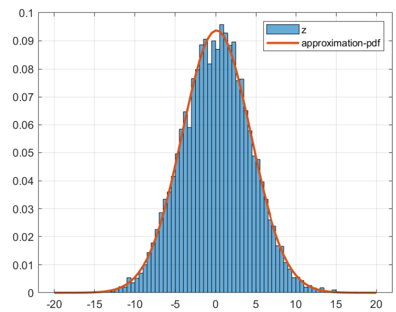

Figure 2.2: Gaussian approximation

The Figure 2.1 illustrates the analytical results and the Figure 2.2 shows the the numerical Gaussian approximation. By observing these two Figures, we can find that the histogram and the two probability density functions (PDFs) match well.

The figure illustrates that the analytical results and numerical approximations of the transformed Gaussian distribution exhibit remarkably similar shapes. This suggests that the numerical approximation accurately reflects the true distribution, providing a reliable estimate of its shape. Since it is a linear transformation, the z after transformation will still satisfies the properties of the Gaussian distribution.

With a small number of samples used in the approximation, we will get poor fitting results, significantly undermining the accuracy of numerical approximation. As the number of samples used in the approximation increases, the accuracy of the approximation improves, gradually converging towards the analytical results.

(b) Since $z = x^3$ is a non-linear transformation and $p(z)$ assumed to be a Gaussian, we could use the definition of the expected value and variance to calculate the analytical results. First, we can calculate the mean of the z.

$$\begin{aligned}\mathbb{E}[z] &= \mathbb{E}[x^3] \\ &= \int_{-\infty}^{\infty} x^3 p(x)\ dx \\ &= \int_{-\infty}^{\infty} x^3 \frac{1}{\sqrt{2\pi}\ \sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\ dx \\ &= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} x^3 e^{\left(-\frac{x^2}{4}\right)}\ dx\end{aligned}$$

Indeed, since $x^3 e^{\left(-\frac{x^2}{4}\right)}$ is an odd function, with x^3 being odd and $e^{\left(-\frac{x^2}{4}\right)}$ being even, its integral over a symmetric interval $[-\infty, \infty]$ evaluates to zero due to symmetry. Thus we can get:

$$\begin{aligned}\mathbb{E}[z] &= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} x^3 e^{\left(-\frac{x^2}{4}\right)}\ dx \\ &= 0\end{aligned}$$

Then we can use the result $\mathbb{E}[z] = 0$ to calculate the variance $\mathrm{Var}[z]$:

$$\begin{aligned}\mathrm{Var}[z] &= \mathbb{E}[(z - \mathbb{E}[z])^2] \\ &= \mathbb{E}[z^2] \\ &= \mathbb{E}[x^6] \\ &= \frac{1}{\sqrt{4\pi}} \int_{-\infty}^{\infty} x^6 e^{\left(-\frac{x^2}{4}\right)}\ dx\end{aligned}$$

Using **Matlab**, we can get the result that $\mathrm{Var}[z] = 120$.

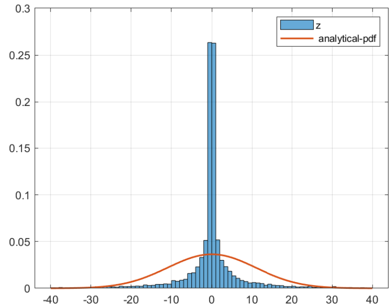

Figure 2.3: Analytical results.

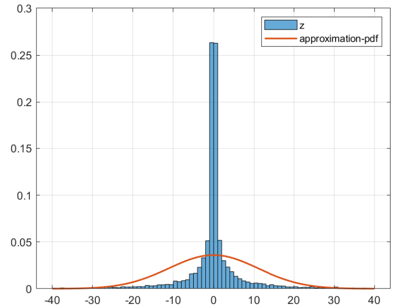

Figure 2.4: Gaussian approximation

The Figure 2.3 illustrates the analytical results and the Figure 2.4 shows the the numerical Gaussian approximation. Observing these two figures reveals a notable discrepancy between the histogram and the two probability density functions (PDFs).

(c) By comparing the results in 2a) and 2b), we can find that the linear transformation can keep some properties of the original random variable. Such that the linear transformation ensures that a scalar Gaussian random variable remains Gaussian even after the transformation. In contrast, a non-linear transformation can lead to significant alterations in the properties of the random variable.

3 Understanding the conditional density

(a) No,we can not describe $p(y)$ with the information given. According to the definition, we can get that:

$$p(y) = \int p(y|x) * p(x)dx$$

Since $y = h(x) + r$ and $h(x)$ is a deterministic, known, possibly non-linear function of x, we can know that $p(y|x)$ is determined. But x is just a random variable, thus the $p(x)$ is undetermined. Therefore we can not describe $p(y)$ with the information given.

(b) Yes, it is possible to describe $p(y|x)$ and we can determine that the $p(y|x)$ is a Gaussian distribution $\mathcal{N}(\mathbb{E}[h(x)], \sigma_r^2)$. Since $h(x)$ is a deterministic, known, possibly non-linear function of x and x is given (we assume it is a constant) , we can get the mean of $h(x)$:$\mu_h = \mathbb{E}[h(x)]$ and it has no variance.

Then we can get the mean and the variance of y with $y = h(x) + r$ and $r \sim \mathcal{N}(0, \sigma_r^2)$:

$$\mu_y = \mathbb{E}[h(x)] + \mathbb{E}[r] = \mathbb{E}[h(x)]$$
$$\mathrm{Var}[y] = \sigma_r^2$$

With the result, we can get:

$$p(y|x) = \frac{1}{\sqrt{2\pi}\ \sigma_r} exp\left(-\frac{(y - \mathbb{E}[h(x)])^2}{2\sigma_r^2}\right)$$

(c) As we previously motivated in 3a), since x remains undetermined, we are still unable to describe $p(y)$. However, for $p(y|x)$, we can provide a description. The $p(y|x)$ is a Gaussian distribution $\mathcal{N}(\mathbb{E}[Hx], \sigma_r^2)$.

$$p(y|x) = \frac{1}{\sqrt{2\pi}\ \sigma_r} exp\left(-\frac{(y - \mathbb{E}[Hx])^2}{2\sigma_r^2}\right)$$

(d) For 3a), we can describe $p(y)$ with the information given. Since we $x \sim \mathcal{N}(\mu_x, \sigma^2)$, we can determine the $p(x)$ and $p(y|x)$.

$$p(y|x) = \frac{1}{\sqrt{2\pi}\ \sigma_r} exp\left(-\frac{(y - \mathbb{E}[h(x)])^2}{2\sigma_r^2}\right)$$
$$p(x) = \frac{1}{\sqrt{2\pi}\ \sigma_x} exp\left(-\frac{(x - \mu_x)^2}{2\sigma_x^2}\right)$$

Then we can get the $p(y)$:

$$\begin{aligned}
p(y) &= \int p(y|x) * p(x) dx \\
&= \int \frac{1}{2\pi\sigma_x\sigma_r} exp\left(-\frac{(y - \mathbb{E}[h(x)])^2 * \sigma_x^2 + (x - \mu_x)^2 * \sigma_r^2}{2\sigma_x^2\sigma_r^2}\right) dx \\
&= \int \frac{1}{2\pi\sigma_x\sigma_r} exp\left(-\frac{(h(x) + r - \mathbb{E}[h(x)])^2 * \sigma_x^2 + (x - \mu_x)^2 * \sigma_r^2}{2\sigma_x^2\sigma_r^2}\right) dx
\end{aligned}$$

For b), we have got the description of $p(y|x)$ above:

$$p(y|x) = \frac{1}{\sqrt{2\pi}\ \sigma_r} exp\left(-\frac{(y - \mathbb{E}[h(x)])^2}{2\sigma_r^2}\right)$$

For c), we can describe both $p(y)$ and $P(y|x)$, they are both Gaussian distribution. We know that the deterministic function of x is a linear function $h(x) = Hx$, where H is a deterministic and known constant.. Thus we can get:

$$\begin{aligned}
\mathbb{E}[h(x)] &= \mathbb{E}[Hx] = H\mathbb{E}(x) = H * \mu_x \\
\mathrm{Var}[h(x)] &= \mathrm{Var}[Hx] = H * \sigma_x^2 * H^T
\end{aligned}$$

Therefore we can describe the $p(y|x)$:

$$p(y|x) = \frac{1}{\sqrt{2\pi}\ \sigma_r} exp\left(-\frac{(y - H * \mu_x)^2}{2\sigma_r^2}\right)$$

The $p(y|x)$ is a Gaussian distribution: $\mathcal{N}(H * \mu_x, \sigma_r^2)$. Then we can know that:

$$\begin{aligned}
\mu_y &= \mathbb{E}[Hx + r] = \mathbb{E}[Hx] + \mathbb{E}[r] = H * \mu_x \\
\mathrm{Var}[y] &= \mathrm{Var}[Hx + r] = H * \sigma_x^2 * H^T + \sigma_r^2
\end{aligned}$$

The $p(y)$ is also a Gaussian distribution: $\mathcal{N}(H * \mu_x, H * \sigma_x^2 * H^T + \sigma_r^2)$. We can describe $p(y)$ as:

$$p(y) = \frac{1}{\sqrt{2\pi(H * \sigma_x^2 * H^T + \sigma_r^2)}} exp\left(-\frac{(y - H * \mu_x)^2}{2(H * \sigma_x^2 * H^T + \sigma_r^2)}\right)$$

(e) First, we assume that $x \sim U(2, 8)$, a linear function is $h(x) = 6x$, an non-linear function is $h(x) = x^3$ and $r \sim \mathcal{N}(0, 9)$

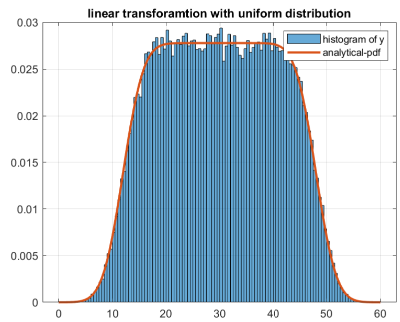

Figure 3.1: linear transformation with uniform distribution.

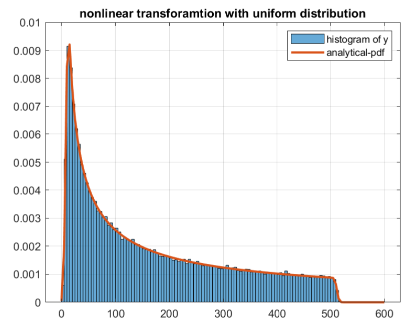

Figure 3.2: non-linear transformation with uniform distribution

Figure 3.1 shows the $p(y)$ after a linear transformation with uniform distribution for x. Figure 3.2 shows the $p(y)$ after a non-linear transformation with uniform distribution for x. It verifies that, if the distribution of x is determined, we are able to describe the $p(y)$.

Then we assume $x \sim \mathcal{N}(1, 4)$.

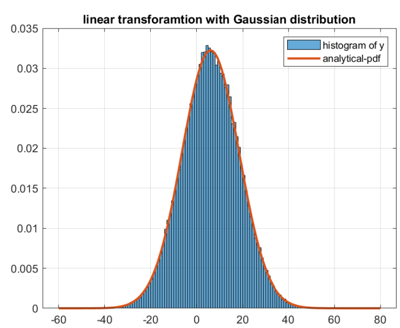

Figure 3.3: linear transformation with Gaussian distribution.

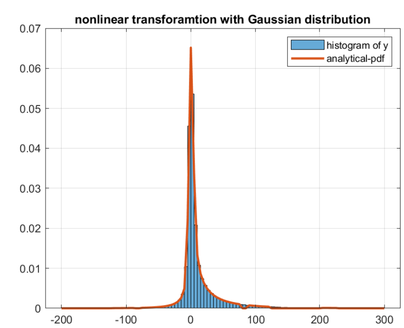

Figure 3.4: non-linear transformation with Gaussian distribution

Figure 3.3 shows the $p(y)$ after a linear transformation with Gaussian distribution for x. Figure 3.4 shows the $p(y)$ after a non-linear transformation with Gaussian distribution for x. It verifies that with $p(x)$ is a Gaussian distribution, we are able to describe the $p(y)$.

Now we make a simulation for $p(y|x)$, the results are shown as bellow.

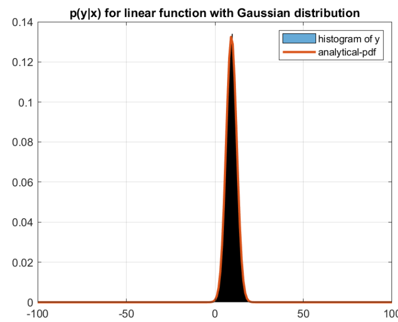

Figure 3.5: $p(y|x)$for linear function with Gaussian distribution.

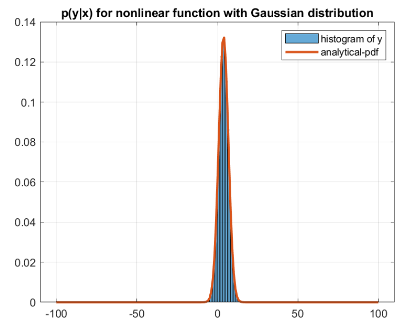

Figure 3.6: $p(y|x)$for non-linear function with Gaussian distribution.

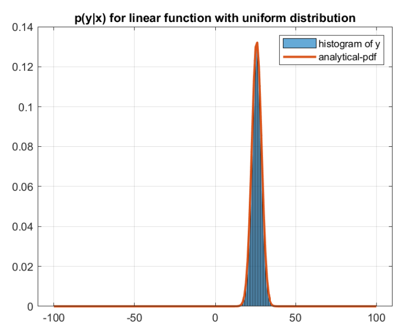

Figure 3.7: $p(y|x)$for linear function with uniform distribution.

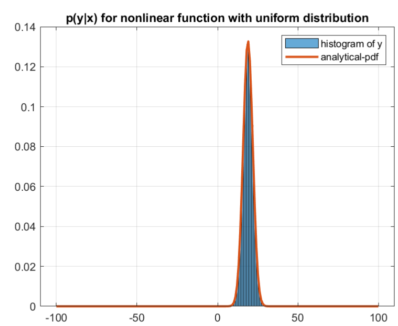

Figure 3.8: $p(y|x)$for non-linear function with uniform distribution.

The results of the simulation for $p(y|x)$, verify that for a given x, if the function is deterministic (whether it is linear or non-linear), we are able to describe the $p(y|x)$.

4 MMSE and MAP estimators

(a) Figure 3.1 illustrates the histogram of y. The histogram of y shows a mixed Gaussian distribution with two Gaussian distributions: $\mathcal{N}(-1, \sigma^2)$ and $\mathcal{N}(1, \sigma^2)$, with two peaks at -1 and 1. Given that θ has an equal likelihood of being either -1 or 1.

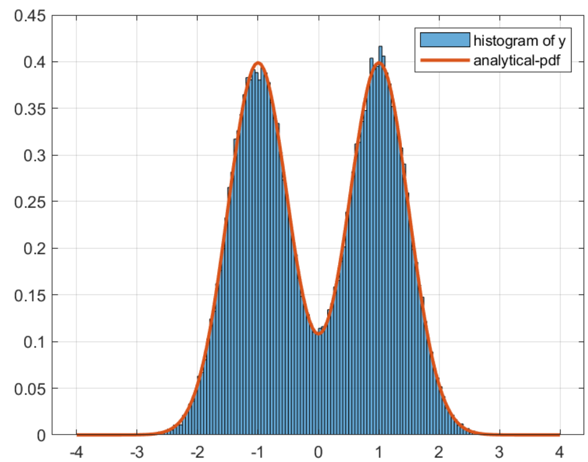

Figure 4.1: The histogram of y.

(b) From the given information and the Bayes' rule, we can get that:

$$p(\theta = 1|y = 0.7) = \frac{p(y = 0.7|\theta = 1) * P(\theta = 1)}{p(y = 0.7)}$$

$$p(\theta = -1|y = 0.7) = \frac{p(y = 0.7|\theta = -1) * P(\theta = -1)}{p(y = 0.7)}$$

Since θ has an equal likelihood of being either -1 or 1:$P(\theta = 1) = P(\theta = -1) = 0.5$, then we only need to consider the $p(y = 0.7|\theta = 1)$ and $p(y = 0.7|\theta = -1)$ to give a guess. Given that $y = \theta + w$, we can calculate:

$$p(y = 0.7|\theta = 1) = \frac{1}{\sqrt{2\pi} * 0.5} exp\left(-\frac{(0.7-1)^2}{2 * 0.5^2}\right) = 0.6664$$

$$p(y = 0.7|\theta = -1) = \frac{1}{\sqrt{2\pi} * 0.5} exp\left(-\frac{(0.7+1)^2}{2 * 0.5^2}\right) = 0.0025$$

By comparing the results, we can make a guess that $\theta = 1$.

(c) Given that $y = \theta + w$, where $W \sim \mathcal{N}(0, \sigma^2)$, we can know that $p(y|\theta)$ is a Gaussian distribution: $\mathcal{N}(\theta, \sigma^2)$ and θ has an equal likelihood of being either -1 or 1:$P(\theta = 1) = P(\theta = -1) = 0.5$. We can show that:

$$p(y) = \int p(y|\theta) * p(\theta) d\theta$$

$$= p(y|\theta = 1) * p(\theta = 1) + p(y|\theta = -1) * p(\theta = -1)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y - \theta = 1)^2}{2\sigma^2}\right) * 0.5 + \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y - \theta = -1)^2}{2\sigma^2}\right) * 0.5$$

$$= 0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} + 0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}$$

(d) By the Bayesian rule, we can get that:

$$p(\theta|y) = \frac{p(y|\theta) * p(\theta)}{p(y)}$$

As we motivated in 4c), we can show that:
For $\theta = 1$:

$$\begin{aligned} Pr\{\theta = 1|y\} &= \frac{p(y|\theta = 1) * p(\theta = 1)}{p(y)} \\ &= \frac{0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}}}{0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} + 0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}} \\ &= \frac{e^{-\frac{(y-1)^2}{2\sigma^2}}}{e^{-\frac{(y-1)^2}{2\sigma^2}} + e^{-\frac{(y+1)^2}{2\sigma^2}}} \\ &= \frac{1}{1 + e^{-\frac{2y}{\sigma^2}}} \end{aligned}$$

For $\theta = -1$:

$$\begin{aligned} Pr\{\theta = -1|y\} &= \frac{p(y|\theta = -1) * p(\theta = -1)}{p(y)} \\ &= \frac{0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}}{0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} + 0.5 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}} \\ &= \frac{e^{-\frac{(y+1)^2}{2\sigma^2}}}{e^{-\frac{(y-1)^2}{2\sigma^2}} + e^{-\frac{(y+1)^2}{2\sigma^2}}} \\ &= \frac{1}{1 + e^{\frac{2y}{\sigma^2}}} \end{aligned}$$

(e) As we motivated in 4d), we can show that:

$$
\begin{aligned}
\hat{\theta}_{MMSE} &= \sum_{\theta} \theta Pr\{\theta|y\} \\
&= \frac{1}{1+e^{-\frac{2y}{\sigma^2}}} - \frac{1}{1+e^{\frac{2y}{\sigma^2}}} \\
&= \frac{e^{\frac{2y}{\sigma^2}}}{e^{\frac{2y}{\sigma^2}}+1} - \frac{1}{1+e^{\frac{2y}{\sigma^2}}} \\
&= \frac{e^{\frac{2y}{\sigma^2}}-1}{e^{\frac{2y}{\sigma^2}}+1} \\
&= tanh(\frac{y}{\sigma^2})
\end{aligned}
$$

(f) By the definition, we can get that:

$$
\begin{aligned}
\hat{\theta}_{MAP} &= argmax_{\theta} p(\theta|y) \\
&= \begin{cases} 1 & if \;\; p(\theta=1|y) \geq p(\theta=-1|y) \\ -1 & if \;\; p(\theta=-1|y) > p(\theta=1|y) \end{cases}
\end{aligned}
$$

As we motivated in 4d), we can find that:

$$
\begin{aligned}
\frac{p(\theta=1|y)}{p(\theta=-1|y)} &= \frac{\frac{1}{1+e^{-\frac{2y}{\sigma^2}}}}{\frac{1}{1+e^{\frac{2y}{\sigma^2}}}} \\
&= \frac{1+e^{\frac{2y}{\sigma^2}}}{1+e^{-\frac{2y}{\sigma^2}}}
\end{aligned}
$$

Then we can observe that

$$
\begin{cases} \frac{p(\theta=1|y)}{p(\theta=-1|y)} \geq 1, & if \;\; y \geq 0 \\ \\ \frac{p(\theta=1|y)}{p(\theta=-1|y)} < 1, & if \;\; y < 0 \end{cases}
$$

Therefore we can simplify the MAP estimator as follow:

$$
\begin{aligned}
\hat{\theta}_{MAP} &= argmax_{\theta} p(\theta|y) \\
&= \begin{cases} 1, & if \;\; y \geq 0 \\ -1, & if \;\; y < 0 \end{cases}
\end{aligned}
$$

(g) With the condition in 4b), we can get that $\hat{\theta}_{MMSE} = 0.9926$ and $\hat{\theta}_{MAP} = 1$. The results indicate that our guess at 4b) that $\theta = 1$ coincides with both MMSE and MAP estimator.

The MMSE estimator aims to minimize the expected value of the squared error between the estimated value of θ and the true value of θ. It achieves this by averaging over all possible values of θ weighted by their respective probabilities. While the MAP estimator aims to maximize the posterior probability of the θ given the observation y.

In this particular example, both the MMSE and MAP estimators make the same decision when $y > 0$ *or* $y < 0$. If $y > 0$, both the MMSE and MAP estimators make the estimation that $\hat{\theta} = 1$. If $y < 0$, both the MMSE and MAP estimators make the estimation that $\hat{\theta} = -1$. But when $y = 0$, the MMSE and MAP estimators will make different decision. The MMSE estimator make the estimation that $\hat{\theta} = -1$, while the MAP estimator make the estimation that $\hat{\theta} = 1$.