Solution to analysis in Home Assignment 1

Daniel Söderqvist (danisode)

05/04-2023

Analysis

In this report I will present my independent analysis of the questions related to home assignment 1. I swear that the analysis written here are my own.

1 Properties of random variables

(a) We want to prove:

(i) $\mathbb{E}(x) = \mu$

To prove this expression we start by defining the probability density function which we know is since we have a Gaussian distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

For our case with the mean μ and the variance σ^2.
We then use the hints to firstly substitute the expression mentioned in the task description and then take the integral from $-\inf$ to $\inf$ of x times the PDF since that is the definition of the expected value of a continuous random variable. By calculating and simplifying we land in the final expression:

$$\begin{aligned}
\mathbb{E}(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu) \exp\left(-t^2\right) \mathrm{d}t \\
&= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma \int_{-\infty}^{\infty} t \exp\left(-t^2\right) \mathrm{d}t + \mu \int_{-\infty}^{\infty} \exp\left(-t^2\right) \mathrm{d}t\right) \\
&= \frac{1}{\sqrt{\pi}} \left(\sqrt{2}\sigma \left[-\frac{1}{2} \exp\left(-t^2\right)\right]_{-\infty}^{\infty} + \mu\sqrt{\pi}\right) \\
&= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}} \\
&= \mu
\end{aligned} \tag{2}$$

Which would be proved.

(ii) $\mathrm{Var}(X) = \mathbb{E}[(x-\mu)^2] = \sigma^2$
To prove this expression we start by defining the same PDF as before since the argument of having mean μ and variance σ^2 is the same as before and we have a Gaussian distribution.
The expression for variance is the following:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2 f_X(x) \mathrm{d}x - (\mathbb{E}(X))^2 \tag{3}$$

And by as before use the hints some calculations and simplifications we finally ends up with the following expression:

$$
\begin{aligned}
\mathrm{Var}(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x - \mu^2 \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu)^2 \exp\left(-t^2\right) \mathrm{d}t - \mu^2 \\
&= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp\left(-t^2\right) \mathrm{d}t + 2\sqrt{2}\sigma\mu \int_{-\infty}^{\infty} t \exp\left(-t^2\right) \mathrm{d}t + \mu^2 \int_{-\infty}^{\infty} \exp\left(-t^2\right) \mathrm{d}t\right) - \mu^2 \\
&= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp\left(-t^2\right) \mathrm{d}t + 2\sqrt{2}\sigma\mu \left[-\frac{1}{2}\exp\left(-t^2\right)\right]_{-\infty}^{\infty} + \mu^2\sqrt{\pi}\right) - \mu^2 \\
&= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp\left(-t^2\right) \mathrm{d}t + 2\sqrt{2}\sigma\mu \cdot 0\right) + \mu^2 - \mu^2 \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 \exp\left(-t^2\right) \mathrm{d}t \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \left(\left[-\frac{t}{2}\exp\left(-t^2\right)\right]_{-\infty}^{\infty} + \frac{1}{2}\int_{-\infty}^{\infty} \exp\left(-t^2\right) \mathrm{d}t\right) \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \int_{-\infty}^{\infty} \exp\left(-t^2\right) \mathrm{d}t \\
&= \frac{2\sigma^2\sqrt{\pi}}{2\sqrt{\pi}} \\
&= \sigma^2
\end{aligned}
\tag{4}
$$

Which then is proved.

(b) The integral form of expectation for a continuous random multi-variate variable $\mathbf{q}$ is given by following when knowing the probability density function $p(q)$ and that we let $\mathbf{z} = \mathbf{Aq}$:

$$\mathbb{E}[g(q)] = \int g(q)p(q)dq \tag{5}$$

And we want to prove that:

(i)

$$\mathbb{E}[z] = A\mathbb{E}[q] \tag{6}$$

Rewriting the expression for the expected value z gives:

$$\mathbb{E}[z] = \mathbb{E}[Aq] = \int Aqp(q)dq \tag{7}$$

Rewriting expression for Aq gives:

$$Aq = [a_1 q_1 + a_2 q_2 + \cdots + a_n q_n] \tag{8}$$

This is then substituted into:

$$\mathbb{E}[z] = \int [a_1 q_1 + a_2 q_2 + \cdots + a_n q_n] \, p(q) dq \tag{9}$$

And expanded with linearity of integration:

$$\mathbb{E}[z] = \int a_1 q_1 p(q) dq + \int a_2 q_2 p(q) dq + \cdots + \int a_n q_n p(q) dq \tag{10}$$

We recognize that we can rewrite into as an integral of each component:

$$\begin{aligned} &\int a_1 q_1 p(q) dq = E\left[a_1 q_1\right] \\ &\int a_2 q_2 p(q) dq = E\left[a_2 q_2\right] \\ &\ldots \\ &\int a_n q_n p(q) dq = E\left[a_n q_n\right] \end{aligned} \tag{11}$$

We can rewrite once again and use linearity of expectation:

$$E[z] = a_1 E\left[q_1\right] + a_2 E\left[q_2\right] + \cdots + a_n E\left[q_n\right] \tag{12}$$

Which is equivalent to:

$$\mathbb{E}[z] = A\mathbb{E}[q] \tag{13}$$

Which then is proved.

(ii) The covariance for the same variable z can be expressed as the following:

$$\mathrm{Cov}[z] = \mathbb{E}\left[(z - \mathbb{E}[z])(z - \mathbb{E}[z])^T\right] \tag{14}$$

Then by firstly expanding using the term inside the expectation then rewriting we get:

$$\mathrm{Cov}[z] = \mathbb{E}\left[A(q - \mathbb{E}[q])(q - \mathbb{E}[q])^T A^T\right] \tag{15}$$

Expanding again:

$$\mathrm{Cov}[z] = A\mathbb{E}\left[(q - \mathbb{E}[q])(q - \mathbb{E}[q])\hat{}T\right] A\hat{}T \tag{16}$$

We can see that from the expression in equation (14) which simply is the covariance is partly the same as the expression on the right hand side. We just substitute and rewrite and end up with:

$$\mathrm{Cov}[z] = A\,\mathrm{Cov}[\mathrm{q}] A\hat{}T \tag{17}$$

Which then is proved.

(c) The function sigmaEllipse2D() is copied from HA1.1.1. Then i wrote a main script for plotting and using both the results from task 1b) and the given data which can be seen in the following figure (1.1).

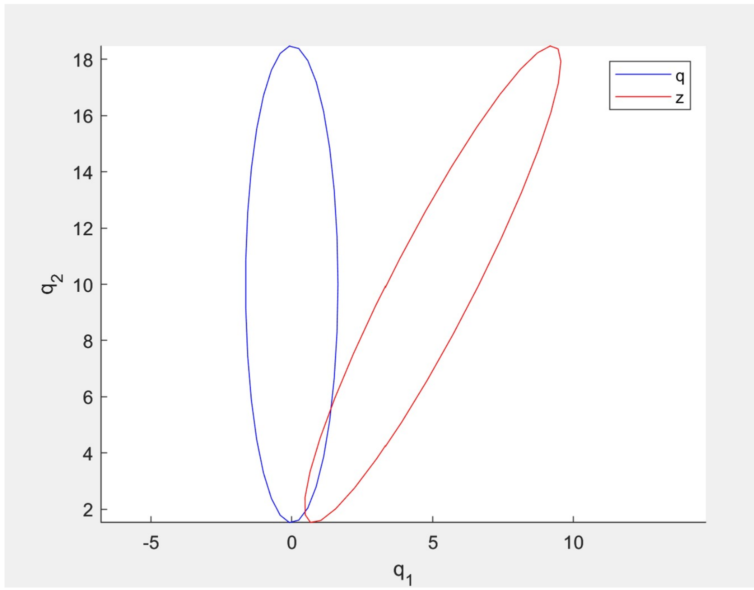


Figure 1.1: Plot of the sigma ellipses for q and z.

We can see that due to the multiplication with A the mean of z is slightly shifted but also the correlation between the components of q.

2 Transformation of random variables

(a) We start by calculating the mean and variance for z which can be done with the following given data and formulas:

$$\mathbb{E}[z] = \mathbb{E}[3x] = 3\mathbb{E}[x] = 3 * 0 = 0 \tag{18}$$

$$\text{Var}[z] = \text{Var}[3x] = 9\,\text{Var}[x] = 9 * 2 = 18 \tag{19}$$

This gives us the Gaussian random variable z with mean 0 and variance 18.

$z \sim \mathcal{N}(0, 18)$ To be able to compare to analytical Gaussian distribution and the numerical approximation that we get from the function approxGaussianTransform() which is just copied from task HA1.1.3 we need to define the function f. Which in our case simply is.

$f = @(x)3 * x$ Then we use the script (see attached main.m script in the zipped file) to evaluate and plot the results. The results are shown in figure (2.1). We can see that the results matches the analytical solution

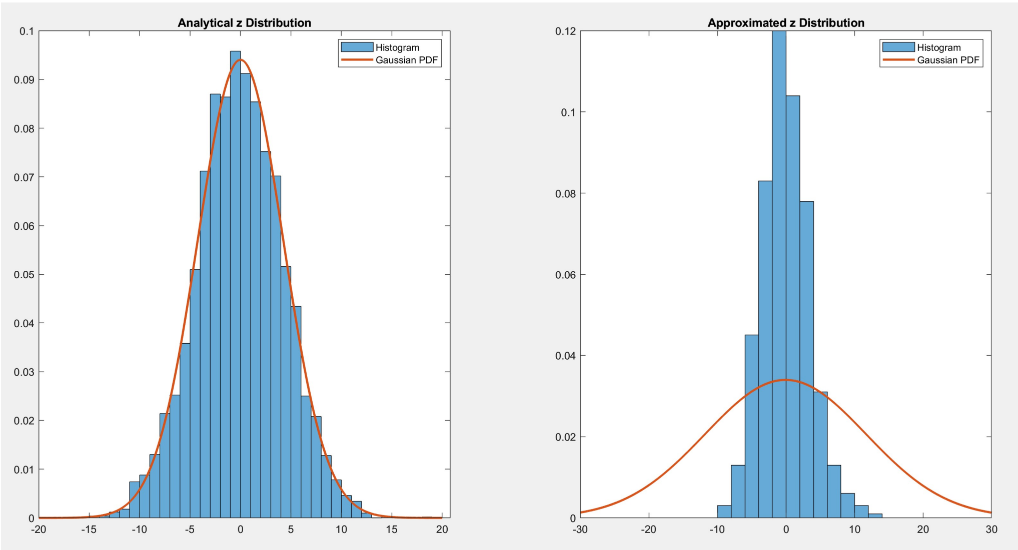


Figure 2.1: Plot of the analytical and approximated distribution of z.

which indicates that approxGaussianTransform() is a accurate method for approximating the distribution z of a linear function.

(b) Since the non-linear transformation x^3 is not a smooth function, one explanation for this is that the samples taken from the changed density might not accurately reflect the actual density. The covariance of the changed variable cannot be precisely estimated by a linear transformation of the covariance of the original variable because the transformation is also nonlinear.

As a result, it is not appropriate to use approxGaussianTransform in this instance to estimate the distribution of the changed variable z. Other approaches, such as numerical integration or Monte Carlo modeling, may be more suitable.

(c) Since the function $3x$ in a) is linear we will be able to decide accurately using approxGaussianTransform the distribution of z where we also confirm this by looking at the figure above (2.1) that the results match well with the analytical solution as well. Therefore the function works well with linear functions. However, in b) the function is not linear anymore but high non-linear which means that the function doesn't work accurately at all in these cases other methods such as Monte Carlo modeling can more accurately estimate the distribution in these cases. To note is that the method might work with a non-linear function but it has to be smooth and well-behaved in order to estimate a good result.

3 Understanding the conditional density

(a) No. Since we don't know the distribution of x or the shape of the deterministic function h, it is impossible to characterize the distribution of y with the knowledge that is currently available.

(b) Yes. We have $y|x \sim N(h(x), \sigma_r^2)$ because the conditional distribution of y given x is just a normal distribution with mean $h(x)$ and variance σ_r^2.

(c) With the function $h(x)$ known we can decide the distribution both when knowing just $p(y)$ and $p(y|x)$. Which simply gives us a distribution with the mean $h(x)$ and variance σ_r^2

(d) In this task both the mean, variance's of x and r are known which mean that we can define the distribution with the law of probability for all the cases even a) in this task.

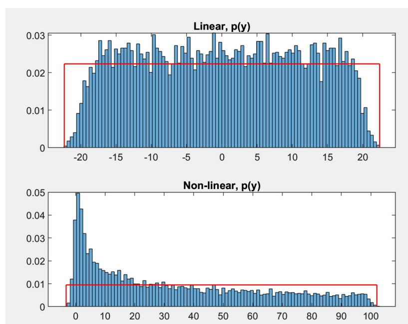
(e) (i) 

Figure 3.1: Linear and non-linear example of a uniform distribution of x, when knowing p(y).

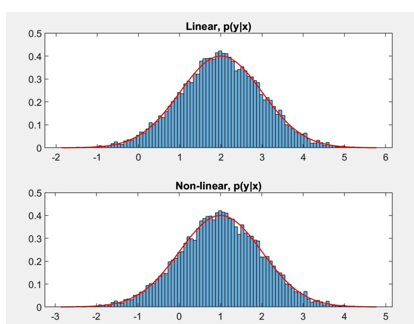
(ii) 

Figure 3.2: Linear and non-linear example of a uniform distribution of x, when knowing p(y—x).

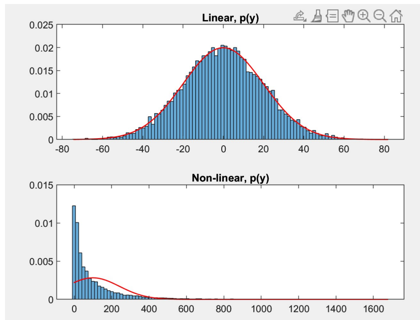
(iii) 

Figure 3.3: Linear and non-linear example of a distribution with known mean and variance of x, when knowing p(y).

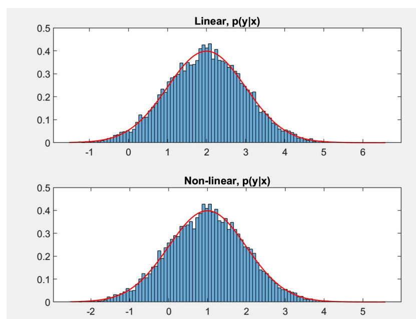
(iv) 

Figure 3.4: Linear and non-linear example of a distribution with known mean and variance of x, when knowing p(y—x).

4 MMSE and MAP estimators

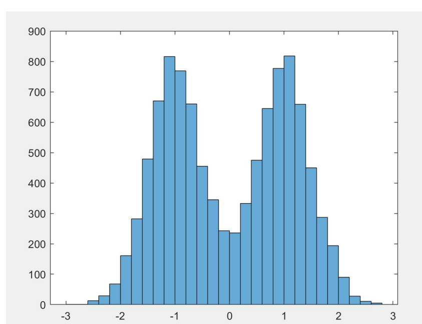
(a) 

Figure 4.1: Histogram of the distribution y.

The graph of y appears to be a combination of two Gaussian distributions, each with a central value of -1 and 1. This is due to the fact that the noise ω is normally distributed with a zero mean and a variation of 0.5^2, while the distribution of θ is a discontinuous uniform distribution over the range of -1, 1.

(b) I would make a guess on either $\theta = -1$ or $\theta = 1$ cause they have the highest probability since the max is around there.

(c) We start by using the given equation (3) from task to be able to describe the probability density function for y given θ.

$$p(y \mid \theta) = \left(1/\sqrt{}(2\pi\sigma\hat{}2)\right) \exp\left(-\left((y-\theta)\hat{}2\right) / \left(2\sigma\hat{}2\right)\right) \tag{20}$$

Then we use the given hint to integrate over all possible θ's and then substitute the expression above for the PDF which gives us after some somplification the expression we were looking for:

$$p(y) = 0.5\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(y-1)^2}{2\sigma^2}} + 0.5\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(y+1)^2}{2\sigma^2}} \tag{21}$$

(d) In this task we use the Bayes' rule to define the θ that most likely corresponds to $y = 0.7$. We write a matlab code to calculate the highest probability of y being 0.7 for both $\theta = -1$ and $\theta = 1$, where we used to answer from previous question and where we then defined the marginal and posterior probability to finally decide which θ that was most likely. The result shows that the highest probability is when $\theta = -1$.

(e) In this task we are supposed to find the MMSE estimator. In the previous task we calculated the posterior probabilities which now simply can be used to define the MMSE estimator. We just then take the weighted sum of the 2 θ's with their correpsdoning propabilties. We end up with $MMSEestimator = -0.99263$.