Exercise session 1

Petteri Pulkkinen

October 15, 2020

Estimate signal gain using maximum likelihood method

Consider a measurement model

$$y(n) = gx(n) + v(n),$$

where g is gain, $x(n)$ is known signal and $v(n)$ is measurement noise. The measurement noise is i.i.d and Gaussian such that for all n

$$v(n) \in \mathcal{N}(0, \sigma^2)$$

and $\mathrm{Cov}\,(v(n), v(m)) = 0$ for all $n \neq m$ and σ^2 is measurement noise power.

The following vector notation will be utilized in the derivations

$$\mathbf{y} = g\mathbf{x} + \mathbf{v}$$

where $\mathbf{y} = [y(1), ..., y(N)]^T$, $\mathbf{x} = [y(1), ..., y(N)]^T$, $\mathbf{v} = [v(1), ..., v(N)]^T$, and N is number of samples measured. The energy of the signal $\|\mathbf{x}\|_2^2 > 0$.

a) Find maximum likelihood estimate for gain parameter g.

The Gaussian probability density function (pdf) can be written as follows

$$f(y(n)|g) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}[y(n) - gx(n)]^2\right)$$

where mean of the distribution is $gx(n)$ and variance σ^2. The likelihood function is

$$f(\mathbf{y}|g) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[y(i) - gx(i)]^2\right)$$

It is equivalent to solve ML problem after taking a logarithm. Therefore, natural logarithm of the likelihood function is calculated

$$l(\mathbf{y}|g) = -\frac{N}{2}(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}[y(i) - gx(i)]^2$$

which is called the log-likelihood function.

In order to find the maximum of the log-likelihood function, lets differentiate $l(\mathbf{y}|g)$

$$\begin{aligned}\frac{\partial}{\partial g} l(\mathbf{y}|g) &= -\frac{1}{2\sigma^2}\frac{\partial}{\partial g}\sum_{i=1}^{N}[y(i) - gx(i)]^2 \\ &= \frac{1}{\sigma^2}\sum_{i=1}^{N}[y(i)x(i) - gx(i)x(i)],\end{aligned}$$

which is also called the score function.

Now let's find roots of the derivative to caculate the ML estimate

$$\begin{aligned}\frac{1}{\sigma^2}\sum_{i=1}^{N}[y(i)x(i) - g_{\mathrm{ML}}x(i)x(i)] = 0 \\ \sum_{i=1}^{N} y(i)x(i) &= g_{\mathrm{ML}}\sum_{i=1}^{N} x(i)x(i) \\ g_{\mathrm{ML}} &= \frac{\sum_{i=1}^{N} y(i)x(i)}{\sum_{i=1}^{N} x(i)x(i)} \\ g_{\mathrm{ML}} &= \frac{\mathbf{y}^T\mathbf{x}}{\|\mathbf{x}\|_2^2}\end{aligned}$$

which is a candidate ML estimator since we have not yet proved that it is the maximum of the log-likelihood function.

To ensure that the g_{ML} is actually maximum of the function $l(\mathbf{y}|g)$, the 2nd derivative is calculated

$$\begin{aligned}\frac{\partial}{\partial g}\frac{1}{\sigma^2}\sum_{i=1}^{N}[y(i)x(i) - gx(i)x(i)] &= -\frac{1}{\sigma^2}\sum_{i=1}^{N} x(i)x(i) \\ &= -\frac{\|\mathbf{x}\|_2^2}{\sigma^2} < 0.\end{aligned}$$

Since the second derivative is negative for all g, the log-likelihood is concave and thus the maximum was found.

b) Prove that g_{ML} is unbiased.

For deterministic parameter g, unbiased estimator means that

$$\mathbb{E}[g_{\mathrm{ML}}] = g.$$

Therefore, let's calculate the expected value of the obtained ML estimate

$$\begin{aligned}
\mathbb{E}\left[\frac{\mathbf{y}^T\mathbf{x}}{\|\mathbf{x}\|_2^2}\right] &= \frac{1}{\|\mathbf{x}\|_2^2}\mathbb{E}\left[\mathbf{y}^T\mathbf{x}\right] \\
&= \frac{1}{\|\mathbf{x}\|_2^2}\left(x(1)\mathbb{E}[y(1)] + x(2)\mathbb{E}[y(2)] + \cdots + x(N)\mathbb{E}[y(N)]\right) \\
&= \frac{1}{\|\mathbf{x}\|_2^2}\left(gx(1)^2 + gx(2)^2 + \cdots + gx(N)x(N)^2\right) \\
&= \frac{g\,\|\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = g.
\end{aligned}$$

Thus the ML estimate is unbiased.

c) Find variance of the ML estimate.

The following properties of variance are used used in the derivations

$$\begin{aligned}
\mathrm{Var}[cX] &= c^2\mathrm{Var}[X] \\
\mathrm{Var}[X+Y] &= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}\left(X, Y\right),
\end{aligned}$$

where X and Y are random variables and c is a constant.

The variance of the ML estimator can be calculated as follows

$$\begin{aligned}
\mathrm{Var}[g_{\mathrm{ML}}] &= \mathrm{Var}\left[\frac{\mathbf{y}^T\mathbf{x}}{\|\mathbf{x}\|_2^2}\right] \\
&= \frac{1}{\|\mathbf{x}\|_2^4}\mathrm{Var}\left[\sum_{i=1}^{N} x(i)y(i)\right] \\
&= \frac{1}{\|\mathbf{x}\|_2^4}\sum_{i=1}^{N}\mathrm{Var}[x(i)y(i)] \\
&= \frac{1}{\|\mathbf{x}\|_2^4}\sum_{i=1}^{N} x(i)^2\mathrm{Var}[y(i)] = \frac{1}{\|\mathbf{x}\|_2^4}\sum_{i=1}^{N} x(i)^2\sigma^2 \\
&= \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^4}\sigma^2 = \frac{\sigma^2}{\|\mathbf{x}\|_2^2}.
\end{aligned}$$

c) Is there a complete sufficient statistic for the observations $\mathbf{y}$?

Fisher–Neyman factorization theorem states that

$$f(\mathbf{y}|g) = h(\mathbf{y})u(T(\mathbf{y})|g)$$

where $T(\mathbf{y})$ is the sufficient statistic and $h(\mathbf{y})$, $u(T(\mathbf{y})|g)$ are arbitrary functions.

The likelihood function was written as follows

$$
\begin{aligned}
f(\mathbf{y}|g) &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[y(i) - gx(i)]^2\right) \\
&= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}[y(i)^2]\right) \\
&= \underbrace{(2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}y(i)^2\right)}_{h(\mathbf{y})} \underbrace{\exp\left(-\frac{1}{2\sigma^2}g^2\sum_{i=1}^{N}x(i)^2\right)\exp\left(\frac{1}{\sigma^2}g\overbrace{\sum_{i=1}^{N}[y(i)x(i)]}^{T(\mathbf{y})}\right)}_{u(T(\mathbf{y})|g)}
\end{aligned}
$$

where we see that

$$T(\mathbf{y}) = \sum_{i=1}^{N}[y(i)x(i)] = \mathbf{y}^T\mathbf{x}.$$

To prove that the statistic is complete, the following lemma is utilized: the sufficient statistic $T(\mathbf{y})$ is complete if

$$\mathbb{E}[r(T(\mathbf{y}))] = 0$$

for all r only if the function is $r(T(\mathbf{y})) = 0$.

The statistic $T(\mathbf{y})$ is distributed as follows $\mathcal{N}(g\,\|\mathbf{x}\|_2^2\,, g^2\sigma^2\,\|\mathbf{x}\|_2^2)$. The proof is based on the fact that a linear combination of Gaussian variables is also Gaussian with mean and variance calculated similarly to in questions b) and c).

Thus,

$$\mathbb{E}[g(T(\mathbf{y}))] = 0$$

only if $r(T(\mathbf{y})) = 0$ for all $T(\mathbf{y}) \in \mathbb{R}$ and $\|x\|_2^2 > 0$.

The ML estimate can be written using the sufficient statistic

$$g_{\text{ML}} = \frac{T(\mathbf{y})}{\|\mathbf{x}\|_2^2}.$$

Therefore according to Lehmann–Scheffé theorem, the ML estimator is minimum variance unbiased estimator (MVUE) since the complete sufficient statistic exists, the ML estimator is unbiased and is dependent on the sufficient statistic.

d) Find Cramér-Rao lower bound (CRLB) of the parameter g. Is the estimator efficient?

The CRLB is defined as follows

$$\text{Var}[\hat{g}] >= \frac{1}{-\mathbb{E}\left[\frac{\partial^2 l(\mathbf{y}|g)}{\partial g^2}\right]}.$$

By substituting the second derivative calculated in a) we obtain the CRLB

$$\text{Var}[\hat{g}] >= \frac{\sigma^2}{\|\mathbf{x}\|_2^2}.$$

The CRLB is equal to variance calculated in c). Therefore, the ML estimator attains the CRLB and is efficient. However, note that in general MVUE is not necessarily efficient.