Solution of Sensor Fusion Homework 3

1 Homework 3

The key point (or trick) of this homework is to realize that even though we use symbols $x_1, \ldots, x_N$ in the problem setting, it does not imply that they would denote the unknown $\mathbf{x}$ in the problem.

(a)

We have

$$J(a,b) = \sum_{n=1}^{N} (y_n - a\,x_n - b)^2. \tag{1}$$

The derivative w.r.t a is

$$\begin{aligned} \frac{\partial J}{\partial a} &= -2 \sum_{n=1}^{N} x_n\ (y_n - a\,x_n - b) \\ &= -2 \left(\sum_{n=1}^{N} x_n\,y_n - a \sum_{n=1}^{N} x_n^2 - b \sum_{n=1}^{N} x_n \right) \end{aligned} \tag{2}$$

and w.r.t b it is

$$\begin{aligned} \frac{\partial J}{\partial b} &= -2 \sum_{n=1}^{N} (y_n - a\,x_n - b) \\ &= -2 \left(\sum_{n=1}^{N} y_n - a \sum_{n=1}^{N} x_n - b \underbrace{\sum_{n=1}^{N} 1}_{=N} \right). \end{aligned} \tag{3}$$

Setting the derivatives to zero thus gives

$$\begin{aligned} -2 \left(\sum_{n=1}^{N} x_n\,y_n - a \sum_{n=1}^{N} x_n^2 - b \sum_{n=1}^{N} x_n \right) &= 0 \\ -2 \left(\sum_{n=1}^{N} y_n - a \sum_{n=1}^{N} x_n - b\,N \right) &= 0, \end{aligned} \tag{4}$$

that is,

$$
\begin{aligned}
a \sum_{n=1}^{N} x_n^2 + b \sum_{n=1}^{N} x_n &= \sum_{n=1}^{N} x_n\, y_n \\
a \sum_{n=1}^{N} x_n + b\, N &= \sum_{n=1}^{N} y_n.
\end{aligned}
\tag{5}
$$

This can be written as matrix equation

$$
\begin{bmatrix} \sum_{n=1}^{N} x_n^2 & \sum_{n=1}^{N} x_n \\ \sum_{n=1}^{N} x_n & N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N} x_n\, y_n \\ \sum_{n=1}^{N} y_n \end{bmatrix}
\tag{6}
$$

which can be solved as

$$
\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{N} x_n^2 & \sum_{n=1}^{n} x_n \\ \sum_{n=1}^{N} x_n & N \end{bmatrix}^{-1} \begin{bmatrix} \sum_{n=1}^{N} x_n\, y_n \\ \sum_{n=1}^{N} y_n \end{bmatrix}.
\tag{7}
$$

If we define

$$
\begin{aligned}
S_x^2 &= \frac{1}{N} \sum_{n=1}^{N} x_n^2, \quad \bar{x} = \frac{1}{N} \sum_{n=1}^{N} x_n \\
S_{x,y} &= \frac{1}{N} \sum_{n=1}^{N} x_n\, y_n, \quad \bar{y} = \frac{1}{N} \sum_{n=1}^{N} y_n
\end{aligned}
\tag{8}
$$

Equation (7) can be written as

$$
\begin{aligned}
\begin{bmatrix} a \\ b \end{bmatrix} &= \begin{bmatrix} S_x^2 & \bar{x} \\ \bar{x} & 1 \end{bmatrix}^{-1} \begin{bmatrix} S_{x,y} \\ \bar{y} \end{bmatrix} = \frac{1}{S_x^2 - \bar{x}^2} \begin{bmatrix} 1 & -\bar{x} \\ -\bar{x} & S_x^2 \end{bmatrix} \begin{bmatrix} S_{x,y} \\ \bar{y} \end{bmatrix} \\
&= \frac{1}{S_x^2 - \bar{x}^2} \begin{bmatrix} S_{x,y} - \bar{x}\, \bar{y} \\ -\bar{x}\, S_{x,y} + S_x^2\, \bar{y} \end{bmatrix} = \frac{1}{S_x^2 - \bar{x}^2} \begin{bmatrix} S_{x,y} - \bar{x}\, \bar{y} \\ -\bar{x}\, S_{x,y} + (S_x^2 - \bar{x}^2)\, \bar{y} + \bar{x}^2\, \bar{y} \end{bmatrix} \\
&= \begin{bmatrix} \frac{S_{x,y} - \bar{x}\, \bar{y}}{S_x^2 - \bar{x}^2} \\ \bar{y} - \frac{S_{x,y} - \bar{x}\, \bar{y}}{S_x^2 - \bar{x}^2} \bar{x} \end{bmatrix}
\end{aligned}
\tag{9}
$$

If we further define

$$
\begin{aligned}
s_x^2 &= S_x^2 - \bar{x}^2 \\
s_{x,y} &= S_{x,y} - \bar{x}\, \bar{y},
\end{aligned}
\tag{10}
$$

then we can write the estimates as

$$
\begin{aligned}
a &= \frac{s_{x,y}}{s_x^2}, \\
b &= \bar{y} - c\, \bar{x},
\end{aligned}
\tag{11}
$$

which matches the form e.g. given in `https://en.wikipedia.org/wiki/Simple_linear_regression#Fitting_the_regression_line` although we should never trust Wikipedia.

(b)

As it is said in the problem hint, the unknown vector is

$$\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix} \tag{12}$$

We now need to figure out a matrix $\mathbf{G}$ such that if we define

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} y_1 & \dots & y_N \end{bmatrix}^\top, \\ \mathbf{r} &= \begin{bmatrix} r_1 & \dots & r_N \end{bmatrix}^\top, \end{aligned} \tag{13}$$

the following:

$$\mathbf{y} = \mathbf{G}\,\mathbf{x} + \mathbf{r} \tag{14}$$

is equivalent to the following:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} a\,x_1 + b \\ \vdots \\ a\,x_N + b \end{bmatrix} + \begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix}. \tag{15}$$

We now notice that

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a\,x_1 + b \\ \vdots \\ a\,x_N + b \end{bmatrix} \tag{16}$$

and thus our $\mathbf{G}$ should be

$$\mathbf{G} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix}. \tag{17}$$

which reduces our problem into the canonical form (14) and we can also notice that

$$J(\mathbf{x}) = (\mathbf{y} - \mathbf{G}\,\mathbf{x})^\top (\mathbf{y} - \mathbf{G}\,\mathbf{x}) = \sum_{n=1}^{N} (y_n - a\,x_n - b)^2 = J(a, b). \tag{18}$$

Using the least squares solution from course material now gives

$$
\begin{aligned}
\hat{\mathbf{x}}_{\mathrm{LS}} &= (\mathbf{G}^\top \mathbf{G})^{-1} \mathbf{G}^\top \mathbf{y} \\
&= \left(\begin{bmatrix} x_1 & \dots & x_N \\ 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1 & \dots & x_N \\ 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \\
&= \begin{bmatrix} \sum_{n=1}^N x_n^2 & \sum_{n=1}^N x_n \\ \sum_{n=1}^N x_n & N \end{bmatrix}^{-1} \begin{bmatrix} \sum_{n=1}^N x_n \, y_n \\ \sum_{n=1}^N y_n \end{bmatrix}
\end{aligned} \tag{19}
$$

which is now the same as (7) and can be simplified in the same way.