# Solutions of Basics of Sensor Fusion Exercise Round 4

Simo Särkkä, Fatemeh Yaghoobi

October 2024

## Exercise 1

Model and and cost function are

$$\begin{aligned} y &= g\,x + r, \\ J(x) &= (y - g\,x)^2. \end{aligned} \tag{1}$$

Thus we have $\mathrm{var}\{r\} = R = 1$. The Jacobian is

$$G_x(x) = g. \tag{2}$$

(a) So the GD takes the form

$$\begin{aligned} \Delta x^{(i+1)} &= g\,(y - g\,x^{(i)}) \\ x^{(i+1)} &= x^{(i)} + \gamma\,\Delta x^{(i+1)} \end{aligned} \tag{3}$$

(b) Let us fix $x^{(i)}$ and consider

$$\begin{aligned} J(x^{(i+1)}) &= J(x^{(i)} + \gamma\,\Delta x^{(i+1)}) \\ &= (y - g\,[x^{(i)} + \gamma\,\Delta x^{(i+1)}])^2 \\ &= (y - g\,[x^{(i)} + \gamma\,g\,(y - g\,x^{(i)})])^2 \\ &= (y - g\,x^{(i)} - \gamma\,g^2\,(y - g\,x^{(i)})])^2 \end{aligned} \tag{4}$$

Now differentiate with respect to $\gamma$:

$$\begin{aligned} &\frac{\partial}{\partial\gamma} J(x^{(i)} + \gamma\,\Delta x^{(i+1)}) \\ &= \frac{\partial}{\partial\gamma}(y - g\,x^{(i)} - \gamma\,g^2\,(y - g\,x^{(i)}))^2 \\ &= -2g^2\,(y - g\,x^{(i)})\,(y - g\,x^{(i)} - \gamma\,g^2\,(y - g\,x^{(i)})) = 0 \end{aligned} \tag{5}$$

When $y - g\,x^{(i)} = 0$, then optimal $\gamma = 0$. Otherwise

$$
\begin{aligned}
y - g\,x^{(i)} - \gamma\, g^2\,(y - g\,x^{(i)}) &= 0 \\
(1 - \gamma\, g^2)(y - g\,x^{(i)}) &= 0 \\
\Rightarrow \gamma &= 1/g^2.
\end{aligned} \tag{6}
$$

So the method with above $\gamma$ becomes

$$
x^{(i+1)} = x^{(i)} + (1/g)\,(y - g\,x^{(i)}) \tag{7}
$$

# Exercise 2

(a) Gauss–Newton takes the form

$$
\begin{aligned}
\Delta x^{(i+1)} &= (1/g^2)\,g\,(y - g\,x^{(i)}) \\
x^{(i+1)} &= x^{(i)} + \Delta x^{(i+1)}
\end{aligned} \tag{8}
$$

The optimum is given by

$$
\begin{aligned}
\frac{\partial J(x)}{\partial x} &= 0 \\
-2g(y - g\,x) &= 0 \\
\Rightarrow x &= y/g.
\end{aligned} \tag{9}
$$

The Gauss–Newton after one step is

$$
\begin{aligned}
x^{(1)} &= x^{(0)} + (1/g^2)\,g\,(y - g\,x^{(0)}) \\
&= x^{(0)} + y/g - x^{(0)} \\
&= y/g.
\end{aligned} \tag{10}
$$

We see for this model, after one step the Gauss–Newton method converges.
(b) We have

$$
x^{(i+1)} = x^{(i)} + (1/g)\,(y - g\,x^{(i)}) \tag{11}
$$

which is equal to (7). So, for this model, Gauss–Newton is equivalent to the gradient-descent method with the optimal step size.

# Exercise 3

Vector form:

$$
\begin{aligned}
\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix} &= \begin{bmatrix} \mathbf{g}(\mathbf{x}) \\ \vdots \\ \mathbf{g}(\mathbf{x}) \end{bmatrix} + \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_n \end{bmatrix} \\
\bar{\mathbf{y}} &= \bar{\mathbf{g}}(\mathbf{x}) + \bar{\mathbf{r}}
\end{aligned} \tag{12}
$$

where $\mathbf{g}(\mathbf{x}) = \begin{bmatrix} \alpha\sqrt{x_1} \\ \beta\sqrt{x_2} \end{bmatrix}$ and $r_n = \mathcal{N}(0, \mathbf{R})$

a) First, we need to to find out the covariance of $\bar{\mathbf{r}}$. We assumed that noises are independent from each other, meaning that $E[r_i r_j] = 0$ for $i \neq j$.
Then

$$cov(\bar{\mathbf{r}}) = \bar{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & 0 & \cdots & 0 \\ 0 & \mathbf{R} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{R} \end{bmatrix} \tag{13}$$

and $\bar{\mathbf{R}}^{-1}$ is

$$\bar{\mathbf{R}}^{-1} = \begin{bmatrix} \mathbf{R}^{-1} & 0 & \cdots & 0 \\ 0 & \mathbf{R}^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{R}^{-1} \end{bmatrix} \tag{14}$$

Then we have

$$\begin{aligned} J_{\text{WLS}} &= (\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x}))^\top W (\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x})) \\ &= (\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x}))^\top \bar{\mathbf{R}}^{-1} (\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x})) \\ &= \sum_{i=1}^{N} (\mathbf{y}_i - \mathbf{g}(\mathbf{x}))^\top \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{g}(\mathbf{x})) \end{aligned} \tag{15}$$

b) First we find the Jacobian of $\mathbf{g}(\mathbf{x})$

$$\mathbf{G}_{\mathbf{x}}^\top(\mathbf{x}) = \begin{bmatrix} \frac{\alpha}{2\sqrt{x_1}} & 0 \\ 0 & \frac{\beta}{2\sqrt{x_2}} \end{bmatrix} \tag{16}$$

Now, for the Jacobian of $\bar{g}(\mathbf{x})$ we have

$$\bar{\mathbf{G}}_{\mathbf{x}}^\top(\mathbf{x}) = \underbrace{\begin{bmatrix} \mathbf{G}_{\mathbf{x}}^\top(\mathbf{x}) & \mathbf{G}_{\mathbf{x}}^\top(\mathbf{x}) & \cdots & \mathbf{G}_{\mathbf{x}}^\top(\mathbf{x}) \end{bmatrix}}_{\text{N times}} \tag{17}$$

Gauss–Newton algorithm

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + (\bar{\mathbf{G}}_{\mathbf{x}}^\top(\mathbf{x})\bar{\mathbf{R}}^{-1}\bar{\mathbf{G}}_{\mathbf{x}}(\mathbf{x}))^{-1}\bar{\mathbf{G}}_{\mathbf{x}}^\top(\mathbf{x})\bar{\mathbf{R}}^{-1}\,(\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x})) \tag{18}$$

which can be simplified to:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + (\underbrace{\bar{\mathbf{G}}_{\mathbf{x}}^\top(\mathbf{x})\bar{\mathbf{R}}^{-1}\bar{\mathbf{G}}_{\mathbf{x}}(\mathbf{x})}_{N\mathbf{G}_{\mathbf{x}}^\top(\mathbf{x})\mathbf{R}^{-1}\mathbf{G}_{\mathbf{x}}(\mathbf{x})})^{-1}\underbrace{\bar{\mathbf{G}}_{\mathbf{x}}^\top(\mathbf{x})\bar{\mathbf{R}}^{-1}\,(\bar{\mathbf{y}} - \bar{\mathbf{g}}(\mathbf{x}))}_{\sum_{i=1}^{N}\mathbf{G}_{\mathbf{x}}^\top(\mathbf{x})\mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{g}(\mathbf{x}))} \tag{19}$$

So, after simplification we have:

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \frac{1}{N}(\mathbf{G}_{\mathbf{x}}^\top(\mathbf{x})\mathbf{R}^{-1}\mathbf{G}_{\mathbf{x}}(\mathbf{x}))^{-1}\sum_{i=1}^{N}\mathbf{G}_{\mathbf{x}}^\top(\mathbf{x})\mathbf{R}^{-1}(\mathbf{y}_i - \mathbf{g}(\mathbf{x})) \tag{20}$$