ML Sheet 7

7.3.3) Levenberg-Marquardt algorithm for circles

The goal is to solve the following optimization problem:

$$\hat{c}, \hat{r} = \underset{c,r}{\arg\min} \sum_{i=1}^{n} \big(\underbrace{\|x_i - c\|_2}_{=d_i} - r \big)^2 = \underset{c,r}{\arg\min}\, L(X,c,r)$$

x_i: data points, c: center of circle, r: radius of circle
d_i: euclidean distance betw. x_i and c

The optimal radius $\hat{r}$ can be expressed as a function of c and the data points x_i:

$$0 \overset{!}{=} \frac{\partial}{\partial r} L(X,c,r) = \sum_i \frac{\partial}{\partial r} (d_i - r)^2 = \sum_i 2(d_i - r)(-1)$$

$$\Leftrightarrow \quad 0 = \sum_i -d_i + \sum_i r = \sum_i -d_i + nr$$

$$\Leftrightarrow \quad \hat{r} = \frac{1}{n} \sum_i d_i =: \bar{r} \quad \text{(mean distance from } c\text{)}$$

This reduces the problem to

$$\hat{c} = \underset{\hat{c}}{\arg\min}\, L(X,c,\bar{r}) = \underset{\hat{c}}{\arg\min} \sum_i (d_i - \bar{r})^2$$

$L(X,c,\bar{r})$ can be simplified:

$$L(X,c,\bar{r}) = \sum_i (d_i^2 - 2d_i\bar{r} + \bar{r}^2)$$

using: $\sum_i (\bar{r}^2 - 2d_i\bar{r}) = \bar{r} \sum_i (\bar{r} - 2d_i) = \bar{r} \big(\underbrace{\sum_i \bar{r}}_{=n\bar{r}} - 2\sum_i d_i \big)$

$$= n\bar{r} \big(\bar{r} - 2 \underbrace{\tfrac{1}{n} \sum_i d_i}_{=\bar{r}} \big) = -n\bar{r}^2$$

$$\Rightarrow L(X,c,\bar{r}) = \sum_i (d_i^2 - n\bar{r}^2)$$

which can be minimized w.r.t c directly where $c = (x,y)$

$$0 \overset{!}{=} \frac{\partial L}{\partial x} = \sum_i \frac{\partial}{\partial x}(d_i^2 - n\bar{r}^2) = \sum_i \left[2d_i\left(\frac{\partial}{\partial x} d_i\right) - 2n\bar{r}\left(\frac{\partial}{\partial x}\bar{r}\right)\right]$$

$$= \sum_i 2d_i\left(\frac{\partial}{\partial x} d_i\right) - 2n\bar{r}\,\frac{1}{n}\sum_j\left(\frac{\partial}{\partial x} d_j\right)\Big]$$

$$\left[\text{using: } \frac{\partial}{\partial x} d_i = \frac{1}{2}\frac{1}{d_i}\frac{\partial}{\partial x}(x_i - x)^2 = \frac{-(x_i - x)}{d_i}\right]$$

$$= \sum_i \left[2\cancel{d_i}\frac{-(x_i - x)}{\cancel{d_i}} - 2\bar{r}\sum_j\left(\frac{-(x_j - x)}{d_j}\right)\right]$$

$$\Leftrightarrow 0 \overset{!}{=} \sum_i \left[\bar{r}\sum_j\left(\frac{x_j - x}{d_j}\right) - \frac{(x_i - x)}{d_i}\right]$$

$\hookrightarrow$ analogous for $\frac{\partial L}{\partial y}$

In the Levenberg-Marquardt algorithm, however, the minimization of

$$L(x_i, c, \bar{r}) = \sum_i (0 - (d - \bar{r}))^2 = \sum_i (0 - f(x_i; c))^2$$

is done by approximating $f(x_i, c)$ around a current guess for c by linearization and iteratively adjusting c by δ as follows:

$$f(x_i | c + \delta) \approx f(x_i | c) + J_i \delta$$

$$\text{with} \quad J_i = \frac{\partial f(x_i | c)}{\partial c}$$

plugging this into the loss function leads an approximation of the loss for the parameters $(c + \delta)$

$$\Rightarrow L(c + \delta) = \sum_i (0 - f(x_i | c + \delta))$$

$$\approx \sum_i (0 - f(x_i | c) - J_i \delta)^2 = \sum_i (d_i - \bar{r} - J_i \delta)^2$$

This can be minimized w.r.t. δ:

→ in matrix notation:

$$L(c+\delta) = \| f - J\delta \|^2 = (f - J\delta)^T (f - J\delta)$$

$$= f^T f \underbrace{- f^T J\delta - (J\delta)^T f}_{= -2\delta^T J^T f} + \underbrace{(J\delta)^T J\delta}_{\delta^T J^T J \delta}$$

$$0 \overset{!}{=} \frac{\partial}{\partial \delta} = \frac{\partial}{\partial \delta} \left(f^T f - 2\delta^T J^T f + \delta^T J^T J \delta \right)$$

$$= -2J^T f + 2J^T J\delta$$

$$\Leftrightarrow \quad J^T f = J^T J \delta \quad \Leftrightarrow \quad \boxed{\begin{aligned} \delta &= (J^T J)^{-1} J^T f \\ &= (J^T J)^{-1} J^T (d_i - \bar{r}) \end{aligned}}$$

with the Jacobian:

$$\frac{\partial f}{\partial x} (d_i - \bar{r}) = \frac{\partial}{\partial x} d_i - \frac{\partial}{\partial x} \frac{1}{n} \sum_j d_j = \frac{-(x_i - x)}{d_i} - \frac{1}{n} \left(\sum_j \frac{-(x_j - x)}{d_j} \right)$$

$$= \frac{1}{n} \left(\sum_j \frac{x_j - x}{d_j} \right) - \frac{(x_i - x)}{d_i}$$

$$\frac{df}{dy} = \frac{1}{n} \left(\sum_j \frac{(y_j - y)}{d_j} \right) - \frac{(y_i - y)}{d_i}$$