Exercise 4

Tuesday 6. October 2020 11.49

1. $$y = \underline{gx + r} \quad (1)$$

$$J(x) = (y - gx)^2 \quad (2)$$

a) $$x_{i+1} = x_i - \gamma \cdot \partial J/\partial x \Big|_{x_i}$$

$$= x_i - \gamma \; [\, 2(y - gx_i) \cdot -g \,]$$

$$= x_i + \gamma \; [\, 2 . g . \underbrace{(y - gx_i)} \,]$$

$$= x_i + \gamma \; [\, 2 \; g \; e_i \,] \quad \ldots (3)$$

b). $$J(x_{i+1}) = (y - g\, x_{i+1})^2$$

$$= (y^2 - 2g\, \underbrace{x_{i+1}}\, y + g^2\, \underbrace{x_{i+1}^2}) \quad \ldots (4)$$

by (3) x_{i+1} is a function of γ.

$$J(\gamma) = J \overset{\text{Composition}}{\circ} x_{i+1}(\gamma)$$

$$J(\gamma) = J(x_{i+1}(\gamma))$$

$$0 = \frac{\partial J}{\partial \gamma} = \frac{\partial J}{\partial x_{i+1}} \cdot \left(\frac{\partial x_{i+1}}{\partial \gamma} \right)$$

$$0 = \underbrace{[\, -2gy + 2g^2\, x_{i+1} \,]} \cdot [\, 2 . g . e_i \,] \quad \ldots (5)$$

$$0 = \underbrace{[\, -2gy + 2g^2\, [x_i + \gamma\, [\, 2ge_i \,]]\,]}_{\mathrm{I}} \cdot \underbrace{[\, 2 . g . e_i \,]}_{\mathrm{II}}$$

Since we cannot equate II to zero generally, after all part II is not a function of γ. then we equate I to zero.

From (5), term I.

$$2gy = 2g^2\, x_{i+1}$$

generally $g \neq 0$. hence we have:

$$y = g\,x_{i+1}$$
$$= g[x_i + \gamma[2ge_i]]$$
$$= gx_i + \gamma 2g^2 e_i$$
$$\gamma\,[2g^2 e_i] = \underbrace{(y - gx_i)}_{e_i}$$
$$\gamma\, 2g^2 e_i = e_i$$

if we select.

$$\gamma = \left(\frac{1}{2g^2}\right) \quad \ldots (6)$$

then we have the first term equal to zero. which means that it optimizes the cost function.

2. b).

$$x_{i+1} = x_i + [g^2]^{-1} g\,[y - gx_i] \quad \ldots (7)$$

the general formulation for GN iteration [WLS]

$$\underline{x}_{i+1} = \underline{x}_i + \left(G_x^T(\underline{x}_i)\, R^{-1} G_x(x)\right)^{-1} G_x^T(\underline{x}_i)\, R^{-1}(\underline{y} - g(\underline{x}_i))$$

G_x is the jacobian of $\underline{g}$

$g \neq 0.$

$$x_{i+1} = x_i + \frac{1}{g^2}\, g\,[y - gx_i] \quad \ldots (8)$$
$$= x_i + \left(\frac{1}{2g^2}\right) 2g\,[y - gx_i]$$
$$= x_i + \gamma^* \, 2g\,[y - gx_i]$$
$$= x_i + \gamma^* \, 2g\,[e_i].$$

So GN is using the ~~an~~ optimal γ for Gradient descent

a). let say that we start from some x_0. according to GN iteration (8), we have [first estimation]

$$x_1 = x_0 + \frac{1}{g^2}\, g\,[y - gx_0] \quad \ldots (9)$$

we would like to see whether x_0, $\mathbb{E}[x_1 - x] = 0$,

using probability density for r, that is $\mathcal{N}(0, R)$, $R = \sigma^2$

$$\boxed{\mathbb{E}[x_1 - x]} = \mathbb{E}\left[x_0 + \frac{1}{g^2} g[y - g x_0] - x\right]$$

remember $\underbrace{(g \neq 0)}$

$$= \mathbb{E}\left[x_0 + \left[\frac{1}{g} y - x_0\right] - x\right]$$

using eq. (1)

$$= \mathbb{E}\left[\frac{1}{g}[gx + r] - x\right]$$

$(g \neq 0)$

$$= \mathbb{E}\left[x + \frac{1}{g} r - x\right]$$

$$= \mathbb{E}\left[\frac{1}{g} r\right] = \frac{1}{g} \mathbb{E}[r] = \frac{1}{g} \cdot 0 = 0 \quad \ldots (10)$$

from the 2nd Exercise.

Exercise 3.

$$\underline{y}_n = \underline{g}(\underline{x}) + \underline{r}_n \quad \ldots (11)$$

from $n = 1, \ldots, N$. $\underline{r}_n \sim \mathcal{N}(\underline{0}, \underline{R})$

as an illustration

$$\underline{y}_1 = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix}, \quad \underline{y}_2 = \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}, \quad \ldots \quad \underline{y}_N = \begin{bmatrix} 0.0 \\ 0.1 \end{bmatrix}$$

$$\bar{y} = \begin{bmatrix} \underline{y}_1 \\ \vdots \\ \vdots \\ \underline{y}_N \end{bmatrix} = \begin{bmatrix} \underline{g}(\underline{x}) \\ \underline{g}(\underline{x}) \\ \vdots \\ \underline{g}(\underline{x}) \end{bmatrix} + \begin{bmatrix} \underline{r}_1 \\ \vdots \\ \vdots \\ \underline{r}_N \end{bmatrix} \quad \ldots (12)$$

$$\bar{y} = \bar{g}(\underline{x}) + \bar{r}, \quad \bar{y} \in \mathbb{R}^{2N}, \bar{r} \in \mathbb{R}^{2N}$$

$$\underline{g}(\underline{x}) = \begin{bmatrix} \alpha\sqrt{x_1} \\ \beta\sqrt{x_2} \end{bmatrix} \quad \ldots (13)$$

a) first $\bar{r} \sim N(0, \bar{R})$, $\mathbb{E}[r_i r_j] = 0$ if $i \neq j$

$$\bar{R} = \begin{bmatrix} R & 0 & 0 & 0 \\ & R & & \\ & & \ddots & \\ & & & R \end{bmatrix} \quad \bar{R}^{-1} = \begin{bmatrix} R^{-1} & & \\ & \ddots & \\ & & R^{-1} \end{bmatrix}$$

using $W = \bar{R}^{-1}$

$$J_{WLS}(x) = (\bar{y} - \bar{g}(x))^T W (\bar{y} - \bar{g}(x))$$

$$= (\bar{y} - \bar{g}(x))^T \bar{R}^{-1} (\bar{y} - \bar{g}(x))$$

$$= \sum_{i=1}^{N} (y_i - g(x))^T R^{-1} (y_i - g(x)) \quad \ldots (14)$$

b). what is the jacobian g

$$G^T(x) := \frac{\partial g}{\partial x} = \begin{bmatrix} \alpha \frac{1}{2\sqrt{x_1}} & 0 \\ 0 & \beta \frac{1}{2\sqrt{x_2}} \end{bmatrix} \quad \ldots (15)$$

for $\bar{g}$, we have

$$\bar{G}^T(x) = \underbrace{[\,G^T \ldots\ldots\ldots G^T\,]}_{N \text{ times}} \quad \ldots (16)$$

$$x_{i+1} = x_i + (\bar{G}^T(x_i) \bar{R}^{-1} \bar{G}(x_i))^{-1} \bar{G}^T \bar{R}^{-1} (\bar{y} - \bar{g}(x))$$

$$= x_i + [N \cdot G^T(x) R^{-1} G(x)]^{-1} \overbrace{[\,G^T \ldots\ldots G^T\,] \begin{bmatrix} R^{-1} & & \\ & \ddots & \\ & & R^{-1} \end{bmatrix} \bar{e}}^{e}$$

$$= x_i + \frac{1}{N} (\bar{G}^T(x) R^{-1} G(x))^{-1} \sum_{i=1}^{N} G^T R^{-1} e_i$$

$$= x_i + (\qquad)^{-1} \frac{1}{N} \sum_{i=1}^{N} G^T R^{-1} e_i \quad \ldots (17)$$

... block diagonal

$$\underbrace{\begin{bmatrix} G^T & \cdots & G^T \end{bmatrix}}_{N \text{ copy}} \overbrace{\begin{bmatrix} R^{-1} & & \\ & \ddots & \\ & & R^{-1} \end{bmatrix}}^{N \text{ block diagonals}} \begin{bmatrix} G \\ \vdots \\ G \end{bmatrix} \Big\} \quad \bar{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

where $e_i = (y_i - g(x))$.

$$= N \; G^T R^{-1} G$$

$$\begin{bmatrix} G^T & \cdots & G^T \end{bmatrix} \begin{bmatrix} R^{-1} & & \\ & \ddots & \\ & & R^{-1} \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

$$= \begin{bmatrix} G^T & \cdots & G^T \end{bmatrix} \begin{bmatrix} R^{-1} e_1 \\ \vdots \\ R^{-1} e_N \end{bmatrix} = \sum_{i=1}^{N} G^T R^{-1} e_i$$