Exercise 5

Tuesday 13. October 2020 11.49

Question 1.

$$\underline{y} = g(\underline{x}) + \underline{r} \quad \dots \quad (1)$$

where $\underline{y} \in \mathbb{R}^{m\times 1}$, $\underline{r} \sim \mathcal{N}(0, R)$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. $g \in C^1$
(it has first derivative).

$W = I$

The cost function : $J(x) = (\underline{y} - g(\underline{x}))^T \underline{R}^{-1} (\underline{y} - g(\underline{x}))$ (2)

The standart Gradient descent:

$$\hat{\underline{x}}_{i+1} = \hat{\underline{x}}_i - \gamma \left.\frac{\partial J}{\partial x}\right|_{x=\hat{x}_i}, \nabla \quad (3)$$

let say that $-\left.\frac{\partial J}{\partial x}\right|_{x=\hat{x}_i} = \Delta \hat{x}_i$, $G_i = \left.\frac{\partial g}{\partial x}\right|_{x=\hat{x}_i}$

Pseudo code then:

given: $\hat{x}_0$, g, J, G, $N_{iteration}$, N_{grid}.

output : $\hat{x}_i$ [also the history of it].

$\gamma = [0, \dots, 1] \longleftarrow$ in linear space with N_{grid} element

for $i = 0$; $i < N_{iteration}$; $i{+}{+}$

$$\Delta \hat{x}_i = G^T(\hat{\underline{x}}_i) \underline{R}^{-1} (\underline{y} - \underline{g}(x_i))$$

$\gamma_{min} = 1$

$J_{min} = J(\hat{x}_i + \gamma_{min} \Delta \hat{x}_i)$

for $j = N_{grid} - 1$; $j > 0$; $j{-}{-}$

$\gamma_{prop} = \gamma[j]$

$J_{prop} = J(\hat{x}_i + \gamma_{prop} \cdot \Delta \hat{x}_i)$

if $J_{min} > J_{prop}$:

$\gamma_{min} = \gamma_{prop}$

$J_{min} = J_{prop}$

$$\gamma_{min} = \gamma_{prop}$$
$$J_{min} = J_{prop}$$

$$\hat{x}_{i+1} = \hat{x}_i + \gamma_{min} \Delta \hat{x}_i$$

Question 2: $$y = g(x) + r \quad (1),$$

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

$y, x \in \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$, differentiable. assume $r \sim \mathcal{N}(0,1)$

assumption that $J(x)$ has unique global minimum. $R = 1$.

a). scaled G-N, we have: $g'(x) = \frac{\partial g}{\partial x}$ $\qquad \frac{\partial J}{\partial x} = 2g'(x)(y-g(x))$

$$\hat{x}_{i+1} = \hat{x}_i + \gamma \frac{g(x_i)}{g'(x_i)\, g'(x_i)} [y - g(x_i)] \qquad \gamma^* = \frac{1}{2(g'(x))^2} \quad (3).$$

handout

direct substitution to eq. in slide no. 22_3 in lecture 4 for measurement model (1) & cost function (2)

---

b). Assume that we start from point $\hat{x}_0$, and $g'(\hat{x}_0) \neq 0$. Show that there exists γ such that the algorithm reach $\hat{x}_*$ on single step.

$$x_* = \arg\min J(x) \quad \text{is unique}$$

it means that there only one x_* and for any other $x \in \mathbb{R}$, $x \neq x^*$
$J(x) > J(x_*)$

$$J(x) = (y - g(x))^2$$

g is differentiable, at x^*

$$0 = \frac{\partial J}{\partial x}\Big|_{x^*} = 2(y - g(x_*))\, g'(x^*) = 0$$

either $(y - g(x^*)) = 0$, or $g'(x_*) = 0$ or both.

if $y - g(x^*) = 0$ then $(x \neq x_*$ $y - g(x) \neq 0$

why? $J(x) = 0$ a contradiction.

$$\underline{\hat{x}_1} = \underline{\hat{x}_0} + \gamma \frac{g'(x_0)}{g'(x_0)\, g'(x_0)} (y - g(x_0)) \quad ..(4)$$

- if $\boxed{\hat{x}_0 \neq x_*}$ $g'(x_0) \neq 0$

$$\hat{x}_1 = x_* = \hat{x}_0 + \gamma \frac{1}{g'(x_0)} (y - g(x_0)) \quad .(5)$$

now since $\hat{x}_0 \neq x_*$, then $y - g(x_0) \neq 0$

then from (5)

$$x_* - \hat{x}_0 = \gamma \frac{1}{g'(x_0)} (y - g(x_0)) \quad ..(6)$$

$$\gamma_* = \boxed{\frac{g'(x_0)(x_* - \hat{x}_0)}{(y - g(x_0))}} \quad ..(7)$$

- what if $\hat{x}_0 = x_*$?

From (5)

$$x_* - \hat{x}_0 = \gamma \frac{1}{g'(x_0)} (y - g(x_0))$$

$$0 = \gamma \frac{1}{g'(x_0)} (y - g(x_0))$$

Should choose $\gamma_* = 0$ $g'(x_0) \neq 0$.

---

c). LS, G-N. $\gamma = [\gamma_0, \; .... \; \gamma_N = 1]$, $x_0 \neq x_*$,

if youare lucky to set γ_* inside γ. then

your $\underline{\hat{x}_1 = x_*}$

for example if your $x_0 \neq x^*$ and you choose

$\gamma = \gamma^*$, then, $x_1 = x^*$.

$$\hat{x}_2 = \hat{x}_1 + \gamma \frac{g'(x_1)}{g'(x_1)^2} [y - g(x_1)] \dots$$

since $\hat{x}_1 = x^*$

$$\hat{x}_2 = x_* + \gamma \frac{g'(x_*)}{g'(x_*)^2} [y - g(x_*)] \quad \dots (8)$$

if your x_* is such that $g'(x_*) \neq 0$ and $y - g(x_*) = 0$

then your $\hat{x}_2 = x_*$, and so on $\hat{x}_3, \hat{x}_4 \dots$

if $g'(x_*) = 0$, then (8) is not defined.