Homework 2

① GD: $x_{k+1} = x_k - \gamma$

$x_{N+J} = x_{N+J+2} = x_{N+J} - \gamma \nabla f(x_{N+J}) - \gamma \nabla f(x_{N+J+1})$

$\gamma \nabla f(x_{N+J}) = -\gamma \nabla f(x_{N+J+1})$ ① $\boxed{\nabla f(x_{N+J}) = -\nabla f(x_{N+J+1})}$

Lets say that $f(x) = x^2$ $\nabla f(x) = 2x$ $\nabla^2 f(x) = 2$

$x_1 = -1$

$x_2 = -1 - \gamma \cdot 2 \cdot (-1) = -1 + 2\gamma$

$x_3 = (-1+2\gamma) - \gamma \cdot 2 \cdot (-1+2\gamma)$

Using condition ① $\nabla f(x_2) = -\nabla f(x_3)$

$x_3 = -1 + 2\gamma + 2\gamma - 4\gamma^2$

$2 \cdot (-1+2\gamma) = -(2 \cdot (-1 + 4\gamma - 4\gamma^2))$

$-2 + 4\gamma = +2 - 8\gamma + 8\gamma^2$

$8\gamma^2 - 12\gamma + 4 = 0$

$\boxed{\gamma_1 = 1}$ $\gamma_2 = 0{,}5$

Check $f(x) = x^2$ $x_1 = -1$ $\gamma = 1$ $\Rightarrow$ We found 2-periodic sequence

$x_2 = -1 - 1 \cdot (-2) = 1$

$x_3 = 1 - 1 \cdot (2 \cdot 1) = -1$

PGD: $x_{k+1} = x_k - \gamma \nabla f(x_k) + \mu (x_k - x_{k-1})$

Let try to get the following sequence $x_1 = 0$ $x_2 = 0$ $x_3 = -1$ $x_4 = 0$ $x_5 = 0$...

$x_3 = x_2 - \gamma \nabla f(x_2) + \mu (x_2 - x_1) = -\gamma \nabla f(x_2) = -1$ $\boxed{\gamma \nabla f(0) = 1}$

$x_4 = x_3 - \gamma \nabla f(x_3) + \mu (x_3 - x_2) = -1 - \gamma \nabla f(-1) - \mu = 0$ $\boxed{-\gamma \nabla f(-1) - \mu = 1}$

$x_5 = x_4 - \gamma \nabla f(x_4) + \mu (x_4 - x_3) = 0 - \gamma \nabla f(0) + \mu = 0$ $\boxed{\gamma \nabla f(0) = \mu}$

3 conditions

$\nabla f(0) = \frac{1}{\gamma} = \frac{\mu}{\gamma}$ $\boxed{\mu = 1}$ $-\gamma \nabla f(-1) - 1 = 1$ $\nabla f(-1) = \frac{2}{-\gamma}$ $\nabla f(0) = \frac{1}{\gamma}$

$f(x) = ax^2 + bx + c$

$\nabla f(x) = 2ax + b$

$\nabla^2 f(x) = 2a$

$a > 0$

$\nabla f(-1) = -2a + b = \frac{2}{-\gamma}$ $\gamma = \frac{2}{2a-b}$

$\nabla f(0) = b = \frac{1}{\gamma}$ $\gamma = \frac{1}{b}$

$\frac{1}{b} = \frac{2}{2a-b} \Rightarrow 2a - b = 2b$, $2a = 3b$

$\boxed{\gamma = \frac{1}{2}}$ $\boxed{b = 2 \mid a = 3}$

Check: $x_1 = 0$ $x_2 = 0$ $f = 3x^2 + 2x$ $\nabla f = 6x + 2$

$x_3 = -\frac{1}{2} \cdot (+2) = -1$

$x_4 = -1 - \frac{1}{2} \cdot (-4) + 1(-1-0) = -1 + 2 - 1 = 0$

$x_5 = 0 - \frac{1}{2} \cdot +2 + 1(0+1) = 0$

NGD: $x_{k+1} = x_k - \gamma \nabla f(x_k + \mu(x_k - x_{k-1})) + \mu(x_k - x_{k-1})$

lets try to form seq: $x_1 = 0 \quad x_2 = 1 \quad x_3 = 0 \quad x_4 = 1 \ldots$

$x_1 = 0 \quad x_2 = 1$

$x_3 = x_2 - \gamma \nabla f(x_2 + \mu(x_2 - x_1)) + \mu(x_2 - x_1) = 1 - \gamma \nabla f(1+\mu) + \mu = 0$

$\boxed{\gamma \nabla f(1+\mu) - \mu = 1}$ 1

$x_4 = 0 - \gamma \nabla f(0 + \mu(0-1)) + \mu(0-1) = -\gamma \nabla f(-\mu) - \mu = 1$

$\boxed{\gamma \nabla f(-\mu) + \mu = -1}$ 2

$x_5 = 1 - \gamma \nabla f(1+\mu) + \mu = 0 \qquad \boxed{\gamma \nabla f(1+\mu) - \mu = 1}$ 3

From 1 $\longrightarrow$ 2 $\longrightarrow$ 3

$\gamma = \frac{1+\mu}{\nabla f(1+\mu)} \qquad \frac{(1+\mu)\nabla f(-\mu)}{\nabla f(1+\mu)} + \mu = -1 \qquad \frac{(1+\mu)\cdot \cancel{\nabla f(1+\mu)}}{\cancel{\nabla f(1+\mu)}} - \mu = 1$

$(1+\mu) - \mu = 1$ ✓

$f(x) = ax^2 + bx + c$

$\nabla f(x) = 2ax + b$

$\nabla^2 f(x) = 2a \quad \boxed{a > 0}$

$\mu > 0$

from $\gamma = \frac{1+\mu}{\nabla f(1+\mu)} > 0 \qquad \nabla f(1+\mu) > 0$

$2a + 2a\mu > 0 \qquad 2a\mu > -2a \qquad \mu > -1$

Take $\mu = \qquad a =$

$\frac{(1+\mu)\cdot(-2a\mu + b)}{(2a + 2a\mu + b)} + \mu = -1$

$(\cancel{1+\mu})(-2a\mu + b) = -(\cancel{1+\mu})(2a + 2a\mu + b)$

$-\cancel{2a\mu} + b = -2a - \cancel{2a\mu} - b \qquad +2b = -2a \qquad \boxed{a = 1 \quad b = -1 \quad \mu = 1}$

$\boxed{\gamma = \frac{2}{4-1} = \frac{2}{3}}$

$f(x) = x^2 - x \qquad \gamma = \frac{2}{3} \quad \mu = 1$

$x_3 = 1 - \frac{2}{3}\cdot 3 - 1 = 0$

$x_4 = 0 - \frac{2}{3}\cdot(-2-1) + 1(0-1) = 2 - 1 = 2$

$x_5 = 1 - \frac{2}{3}(3) + 1 = 2 - 2 = 0$

(2) Determine optimal learning rates γ, μ for the Polyak GD on

$$f(x,y,z) = x^2 + 2y^2 - 2yz + 4z^2 + 3x - 4y + 5z$$

For Polyak GD we choose: $\sqrt{\mu} = \frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}}$ $\gamma = \frac{4}{(\sqrt{\alpha} + \sqrt{\beta})^2}$

$$\nabla f = \begin{bmatrix} 2x+3 \\ 4y-2z-4 \\ -2y+8z+5 \end{bmatrix} \qquad \nabla^2 f = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & -2 \\ 0 & -2 & 8 \end{bmatrix}$$

Searching for the eigen-values:

$$\det(\nabla^2 f - \lambda I) = 0 = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & -2 \\ 0 & -2 & 8-\lambda \end{vmatrix} = (2-\lambda) \cdot \begin{vmatrix} 4-\lambda & -2 \\ -2 & 8-\lambda \end{vmatrix} = (2-\lambda)\left((4-\lambda)(8-\lambda) - 4\right) =$$

$$= (2-\lambda)(32 - 8\lambda - 4\lambda + \lambda^2 - 4) = 0$$

We see that the function is complex since $\nabla^2 f$ is PD.

$\lambda_1 = 2$ $\quad \lambda^2 - 12\lambda + 28 = 0$

$$\lambda_{2/3} = \frac{12 \pm \sqrt{144 - 112}}{2} = \frac{12 \pm 4\sqrt{2}}{2} = 6 \pm 2\sqrt{2}$$

$\alpha = 2 = \lambda_1$. $\quad \beta = 6 + 2\sqrt{2} = \lambda_3$

$$\sqrt{\mu} = \frac{\sqrt{6+2\sqrt{2}} - \sqrt{2}}{\sqrt{6+2\sqrt{2}} + \sqrt{2}} = 0{,}355 \qquad \mu = 0{,}126 \qquad \gamma = \frac{4}{(\sqrt{2} + \sqrt{6+2\sqrt{2}})^2} = 0{,}208$$

(5)

(a) $f(x,y,z) = (x-z)^2 + (2y+z)^2 + (4x-2y+z)^2 + x + y$

$$\nabla f = \begin{bmatrix} 2(x-z) + 2(4x-2y+z) \cdot 4 + 1 \\ 2(2y+z) \cdot 2 + 2(4x-2y+z) \cdot (-2) + 1 \\ -2(x-z) + 2(2y+z) + 2(4x-2y+z) \end{bmatrix} = \begin{bmatrix} 34x - 16y + 6z + 1 \\ -16x + 16y + 1 \\ 6x + 6z \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} 34 & -16 & 6 \\ -16 & 16 & 0 \\ 6 & 0 & 6 \end{bmatrix}$$

(b) $f(x,y,z) = (x-1)^2 + (y-1)^2 + 100(y-x^2)^2 + 100(z-y^2)^2$

$$\nabla f = \begin{bmatrix} 2(x-1) + 200(y-x^2) \cdot (-2) \cdot x \\ 2(y-1) + 200(y-x^2) + 200(z-y^2) \cdot (-2) \cdot y \\ 200 \cdot (z-y^2) \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} -400(y-x^2) + 800x^2 + 2 & -400x & 0 \\ -400x & -400(z-y^2) + 800y^2 + 202 & -400y \\ 0 & -400y & 200 \end{bmatrix}$$