MA574. Pre class Assignment - 25.

1) $f(x) = x^3 - 3x^2 - x + 3$ where $x \geq 0$

→ 1) constant step size $r = \frac{1}{50}$

$x1 = x0 - \frac{1}{50} \cdot f'(x0)$

$x2 = x1 - \frac{1}{50} \cdot f'(x1)$

$x3 = x2 - \frac{1}{50} \cdot f'(x2)$

2) Variable step size $r_t = \frac{1}{25t}$

$x1 = x0 - \frac{1}{25 \times 1} f'(x0)$

$x2 = x1 - \frac{1}{25 \times 2} \cdot f'(x1)$

$x3 = x2 - \frac{1}{25 \times 3} f'(x2)$

3) variable step size $rt = \frac{1}{t}$.

$x1 = x0 - \frac{1}{1} \cdot f'(x0)$

$x2 = x1 - \frac{1}{2} \cdot f'(x1)$

$x3 = x2 - \frac{1}{2} \cdot f'(x2)$

4) variable stepsize $\gamma_t = \frac{1}{50t}$

$$x_1 = x_0 - \frac{1}{50 \times 1} f'(x_0)$$

$$x_2 = x_1 - \frac{1}{50 \times 2} \cdot f'(x_1)$$

$$x_3 = x_2 - \frac{1}{50 \times 3} \cdot f'(x_3)$$

Solving for x_1, x_2, x_3, we get:

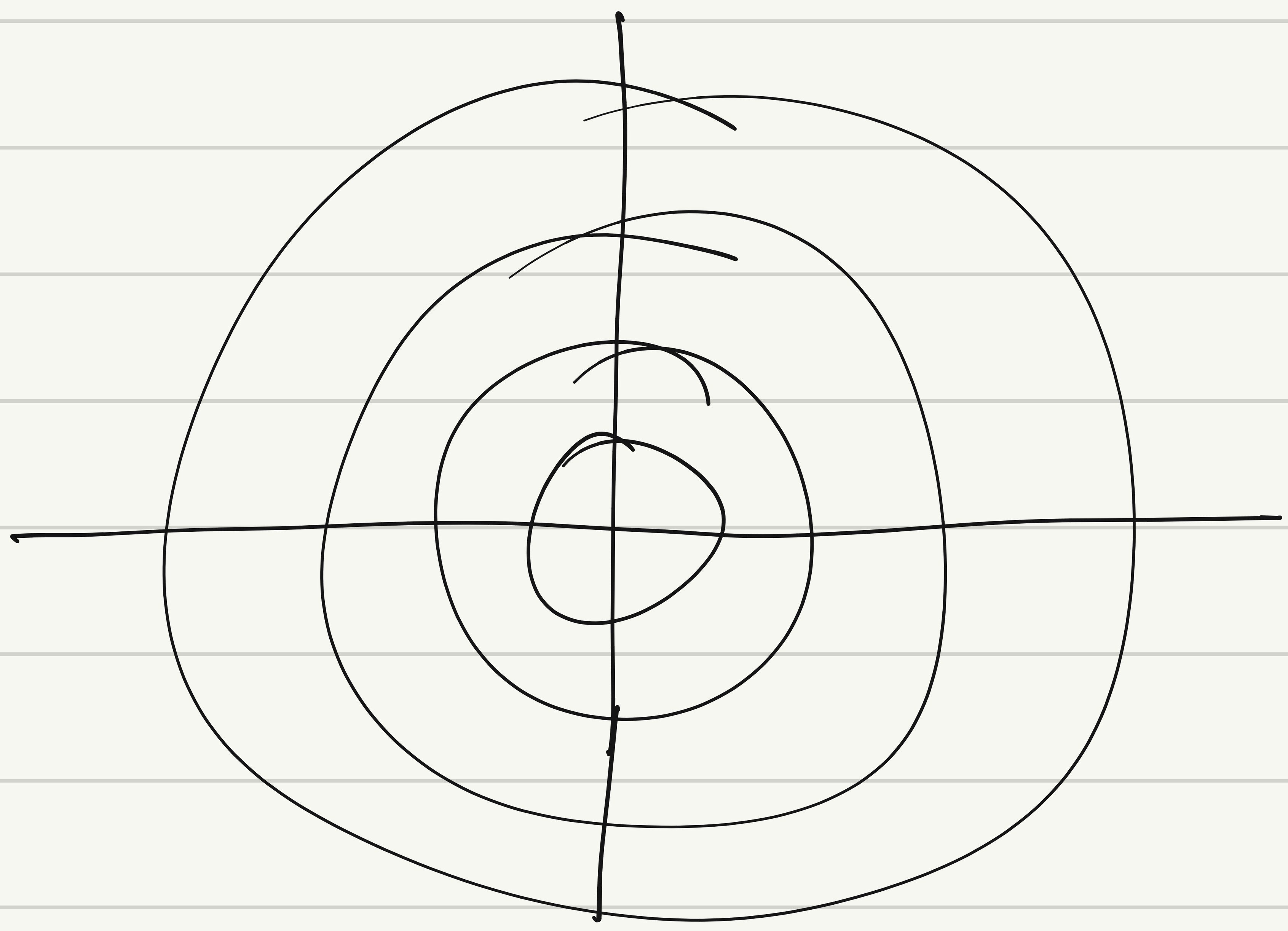
$$x_1 = \frac{27}{25}, \quad x_2 = \frac{537.45}{500}, \quad x_3 = \frac{537.45}{500}.$$

Q2. $f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$

a is subsequently larger than b the level curves are ellipses centered at the origin $f(x,y)=0$

Now for the gradient descent without momentum might converge slowly towards the origin due to highly elongated nature of the ellipses.

1) Contour diagram.



$f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ where $a > b$.

2) Gradient descent without momentum.

point $(x_i y_i)$ may converge slowly due to the elongated nature of the ellipses with gradient along longer axis (a)

3) $\alpha = 0$

$x_0 = -1$

round 1:

$$x_1 = -1 - \frac{1}{10}\left(4(-1)^3 - 3(-1)^2 - 2(-1)\right) = -0.9.$$

Round 2 :

$x_2 = -0.81$

$x_3 = -0.729$

$x_4 = -0.656$

approx value

1) $\alpha = 0.3$: $x = -0.656$

2) $\alpha = 0.6$: $x = -0.409$

3) $\alpha = 0$ no momentum.

Q3. $f(x) = x^4 - x^3 - x^2 + 1$

$x_0 = -1$

$\gamma = \frac{1}{10}$

1) $\alpha = 0.3$.

round 1 : $x_1 = -0.9$

round 2 = $x_2 = -0.81$

round 3 = $x_3 = -0.729$

round 4 = $x_4 = -0.656$.

2) $\alpha = 0.8$

$x_0 = -1$, $v_0 = 0$.

round 1 : $x_1 = 0.8$

round 2 = $x_2 = -0.64$

round 3 = $x_3 = -0.512$

round 4 = $x_4 = -0.409$

3) Point (x_i, y_i) will converge efficiently by exploiting the geometric properties.