c) NESTEROV gradient descent

$$x_{k+1} = x_k - \gamma \cdot \nabla f\left(x_k + \mu(x_k - x_{k-1})\right) + \mu(x_k - x_{k-1})$$

I) 2-periodic

Nesterov GD may result in non-constant 2-periodic sequence.

We pick a quadratic function: $f(x) = ax^2$ $\quad \nabla f(x) = 2ax$

$x_1 = 1$

$x_2 = -1$

$x_3 = 1$ ① $x_3 = x_2 - \gamma \cdot \nabla f(x_2 + \mu(x_2 - x_1)) + \mu(x_2 - x_1) = 1 \Rightarrow x_3 = -1 - \gamma \cdot \nabla f(-1 + \mu \cdot (-2)) + \mu \cdot (-2) = 1$

$x_4 = -1$ ② $x_4 = x_3 - \gamma \cdot \nabla f(x_3 + \mu(x_3 - x_2)) + \mu(x_3 - x_2) = -1 \Rightarrow x_4 = 1 - \gamma \cdot \nabla f(1 + \mu(2)) + 2\mu = -1$

Set $\mu = 1$

↓

Works for any $\mu > 0$, but then γ is different, also > 0

① $\gamma = \frac{-2\mu - 2}{\nabla f(\underbrace{-1-2\mu}_{-3})} = \frac{-2-2}{-6a} = \frac{2}{3a}$ ② $\gamma = \frac{2\mu + 2}{\nabla f(1+2\mu)} = \frac{2+2}{6a} = \frac{2}{3a}$ ✓

Similar example as for Polyak example:

$f(x) = x^2$, $\nabla f(x) = 2x$, $\mu = 1$, $\gamma = \frac{2}{3}$