ML ERASMUS+, NOV 28, 2022

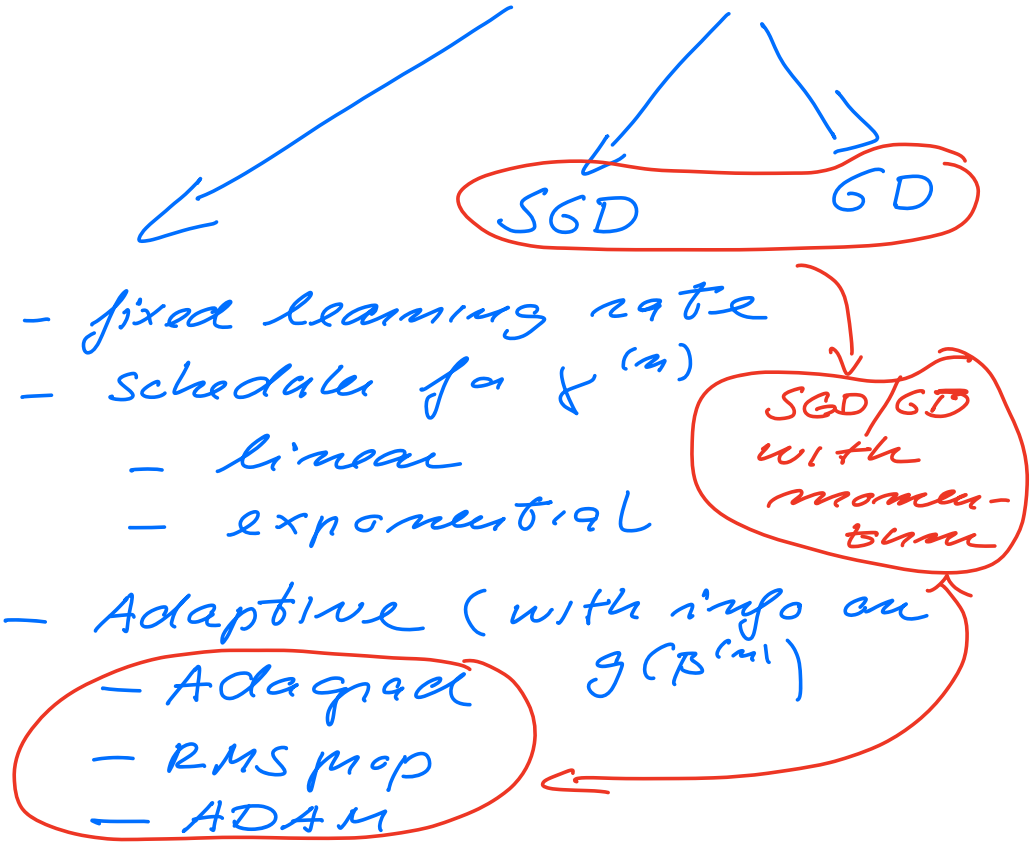
Gradients and optimization

$$\frac{\partial C(\beta)}{\partial \beta} = 0 = g(\beta)$$

Newton's method

$$\beta^{(m+1)} = \beta^{(m)} - H^{-1}(\beta^{(m)}) g(\beta^{(m)})$$

$$\simeq \beta^{(m)} - \gamma^{(m)} g(\beta^{(m)})$$



- Automatic differentiation
  - Difference methods-

$$\frac{df(x)}{dx} \approx \frac{f(x+\Delta x) - f(x-\Delta x)}{2\Delta x}$$

$$(O(\Delta x^2))$$

  - Symbolic manipulations-
  - analytical derivations-