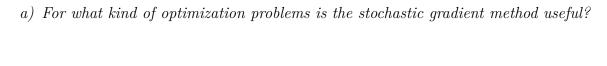
Optimization and Data Science 10. Presence exercises

Theoretical exercise 1: Stochastic gradient method



b) How does the stochastic gradient method work?

Theoretical exercise 2: Stochasti	c, minibatch ar	nd full	gradient	method
-----------------------------------	-----------------	---------	----------	--------

a) What is the difference between the stochastic gradient, the minibatch gradient and the full gradient method.

b) How would you quantify the total computational effort of the stochastic gradient, the minibatch gradient and the full gradient method?

Theoretical exercise 3: Loss function

a) What are typical loss functions?

b) Are there any disadvantages when choosing certain loss functions?

Theoretical exercise 4: Evolution Strategies

- a) How does the $(\mu + \lambda)$ -Evolution Strategy work using the following mutability?
 - constant mutability
 - different mutabilities
 - mutated mutabilities
 - Mutation Strategy Parameter Control (MSC)
- b) What is the difference between $(\mu + \lambda)$ -ES and (μ, λ) -ES?

c) What are the main differences between CMA-ES and $(\mu + \lambda)$ -ES?