Part a

A screenshot of a computer program

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

Part b

A blackboard with math equations

AI-generated content may be incorrect.A screenshot of a computer screen

AI-generated content may be incorrect.

Part c

A blackboard with math equations

AI-generated content may be incorrect.

Part d

A blackboard with math equations

AI-generated content may be incorrect.

A blackboard with math equations

AI-generated content may be incorrect.

Part e

A screen shot of a computer program

AI-generated content may be incorrect.A graph of a function

AI-generated content may be incorrect.

based on the plots, alpha = 2 is the best.

Part f

A screen shot of a computer program

AI-generated content may be incorrect.

A graph of a function

AI-generated content may be incorrect.

best is alpha = 1.2. SGD only uses one data point's gradient, so sometimes that estimate will make delta k go up.

Part g

A blackboard with math equations

AI-generated content may be incorrect.

A blackboard with math equations

AI-generated content may be incorrect.

Part h

A screenshot of a computer program

AI-generated content may be incorrect.A graph of a graph showing the difference between a certain and a certain value

AI-generated content may be incorrect.

Part i

if you can afford to do GD (O(p^2) cost), then it will be faster. coordinate descent costs O(p) so with larger p, it would be better than GD.