1. No, this is FALSE. Depending on the starting point, the method might converge to a minimum or a maximum. Thus, once an optimum has been reached, one must check whether it is a minimum or a maximum (for example by considering the second derivative of f(x)). As an example, repeat Example 2.4, but start instead at $x_0=0$ (instead of 2). Which point will then be reached? Is it a maximum or a minimum?

2. No. It is necessary, but not sufficient, that f(x) should be convex. In order for an optimization problem to be convex, the constraints defining the feasible set S must fulfil certain criteria. More specifically, the inequality constraints must be convex, and the equality constraints must be affine. See also pp. 24-25 in the book.

3. The gradients of f and h are parallel at the local optima.

4. Yes, and this is one, among several, advantages with such methods. Stochastic optimization methods do not explicitly make use of gradients (or higher-order derivatives), and they can therefore handle non-differentiable objective functions.