Artificial Intelligence 2 Supervisior: Marton Havasi
Lectures 5-8ish 08/05/2018

Core questions (expected)

1. Describe the gradient descent algorithm. Discuss its advantages and disadvantages. How can one set the step size? Is there any benefit to adapting the step size during optimization? (4 marks)

2. f is a twice differentiable function that we wish to minimize. The second order Taylor expansion of f at point x_n is

$$f(x) \approx f(x_n) + \nabla f(x_n)\Delta x + \frac{1}{2}\nabla^2 f(x_n)\Delta x^2$$

   Newton's optimization method uses the Hessian matrix (H) for the update rule:

$$x_{n+1} = x_n - H^{-1}\nabla f(x_n)$$

   Derive this update rule from the Taylor expansion. (4 marks)

   Relate this method to the gradient descent algorithm. (2 marks)

   How and why might one choose to include a step-size in the update equation? (2 marks)

3. CSP: 2014 paper 4 question 2 Link (20 marks)

4. KR&R: 2010 paper 4 question 2 Link (20 marks)

Tryhard questions (highly recommended)

1. KR&R: 2006 paper 4 question 4 Link (20 marks)