

This handout includes space for every question that requires a written response. Please feel free to use it to handwrite your solutions (legibly, please). If you choose to typeset your solutions, the `README.md` for this assignment includes instructions to regenerate this handout with your typeset L^AT_EX solutions.

1.a

1.b

1.c

2.a

$$J(\theta) =$$

Differentiating this objective, we get:

$$\nabla_{\theta} J(\theta) =$$

The gradient descent update rule is

$$\theta := \theta - \lambda \nabla_{\theta} J(\theta)$$

which reduces here to:

2.c

2.e

2.g

3.a

3.b

3.c

3.d