

## Problem 2

1. Batch gradient descent:

$$\nabla J = 2\lambda \beta_K - X^T(y - \mu)$$

$$\mu = \frac{1}{1 + e^{-\beta_K^T x}}$$

$$\beta_{K+1} = \beta_K - 2\nabla$$

$$NLL = \lambda \|\beta\|_2^2 - \sum_{i=1}^n [y_i \log(\mu_i) + (1 - y_i) \log(1 - \mu_i)]$$

2

$$\mu_i = \frac{1}{1 + e^{-\beta^T x_i}}$$

$$\nabla = 2\lambda \beta - X_i^T(y_i - \mu)$$

$$\beta_{K+1} = \beta_K - 2\nabla$$

$$NLL = \lambda \|\beta\|_2^2 - \sum_{i=1}^N [y_i (\log(\mu_i)) + (1 - y_i) \log(1 - \mu_i)]$$