Logistic Regression

Introduction to Data Science Algorithms

Jordan Boyd-Graber and Michael Paul
ABC

Logistic Regression: Objective Function

$$\ell \equiv \ln p(Y \mid X, \beta) = \sum_j \ln p(y^{(j)} \mid x^{(j)}, \beta) \tag{1}$$

$$= \sum_j y^{(j)} \left(\beta_0 + \sum_i \beta_i x_i^{(j)} \right) - \ln \left[1 + \exp \left(\beta_0 + \sum_i \beta_i x_i^{(j)} \right) \right] \tag{2}$$

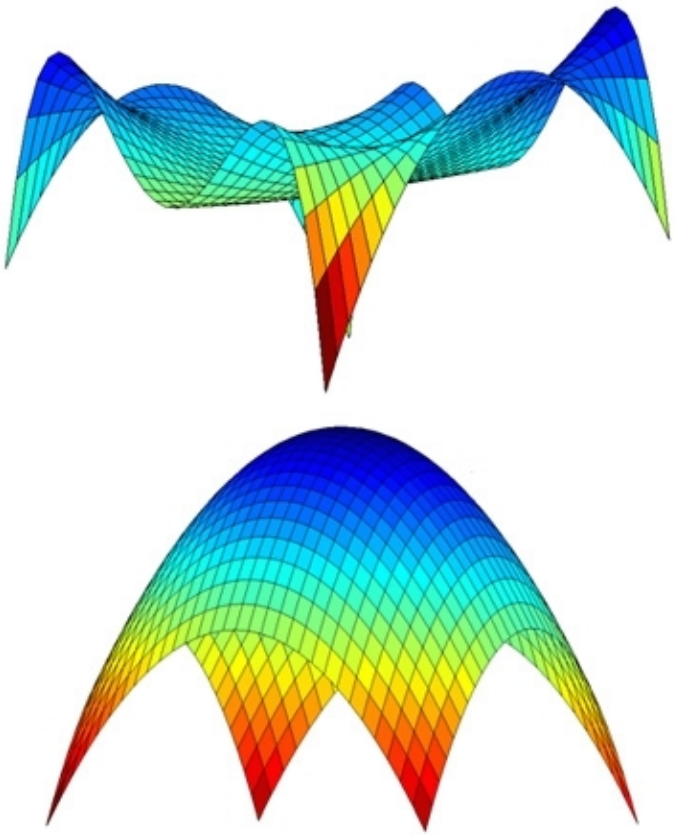
Logistic Regression: Objective Function

$$\ell \equiv \ln p(Y \mid X, \beta) = \sum_j \ln p(y^{(j)} \mid x^{(j)}, \beta) \tag{1}$$

$$= \sum_j y^{(j)} \left(\beta_0 + \sum_i \beta_i x_i^{(j)} \right) - \ln \left[1 + \exp \left(\beta_0 + \sum_i \beta_i x_i^{(j)} \right) \right] \tag{2}$$

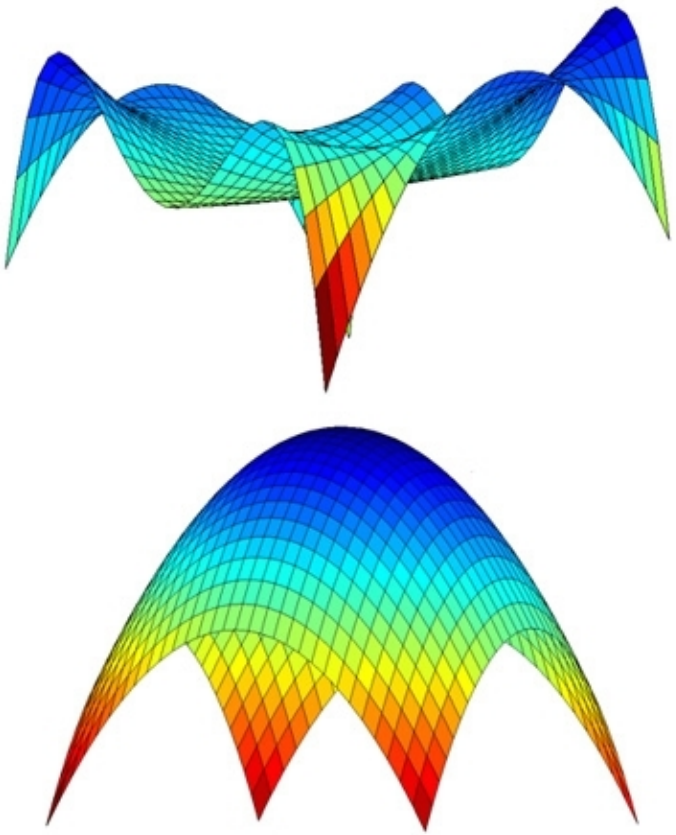
Training data (y, x) are fixed. Objective function is a function of β ... what values of β give a good value.

Convexity



- Convex function
- Doesn't matter where you start, if you "walk up" objective

Convexity

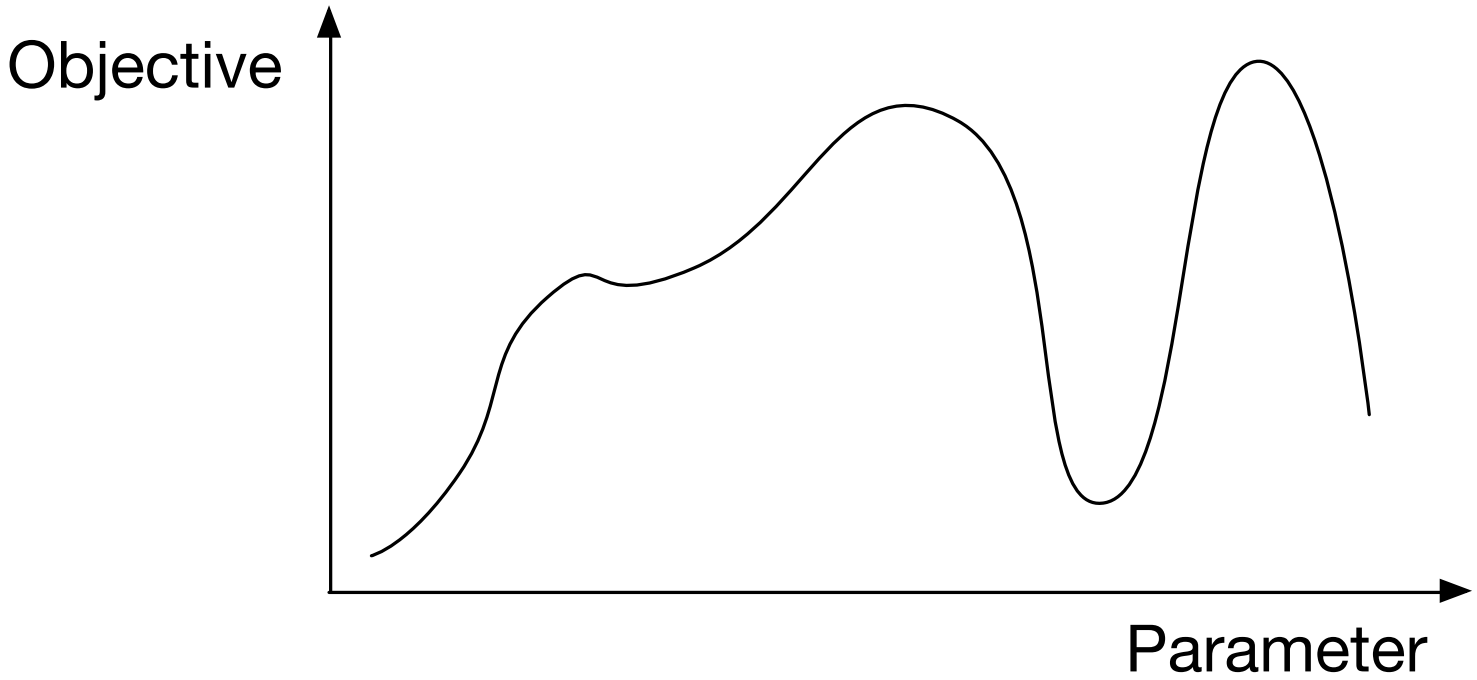


- Convex function
- Doesn't matter where you start, if you "walk up" objective
- Gradient!

Gradient Ascent (non-convex)

Goal

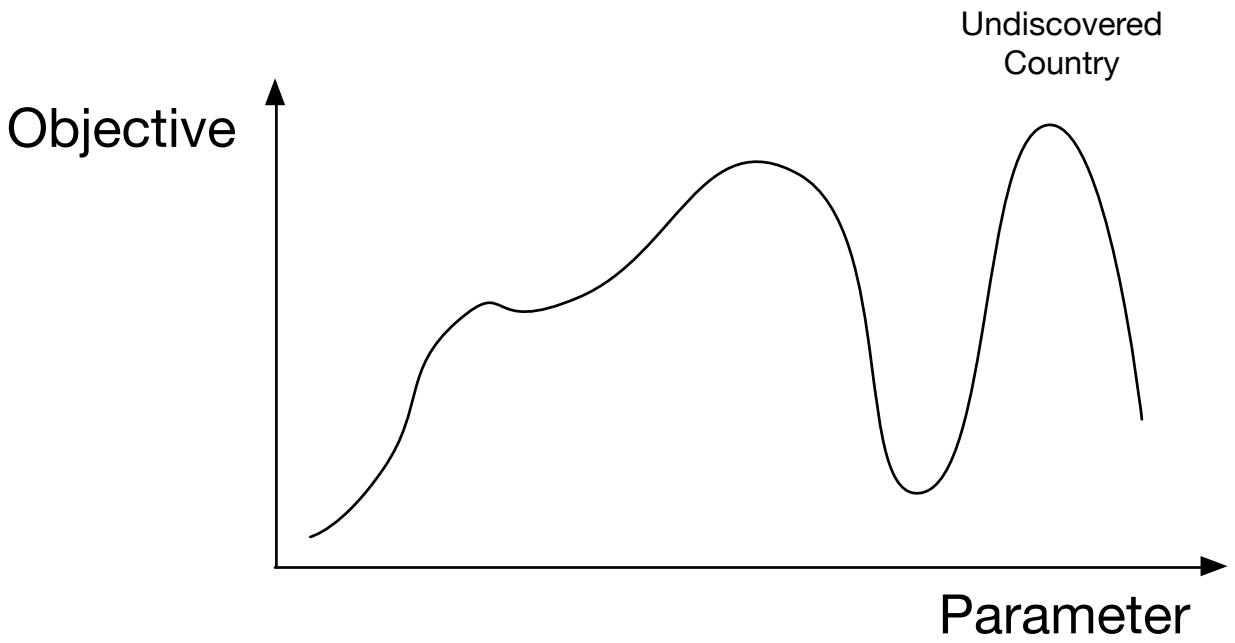
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

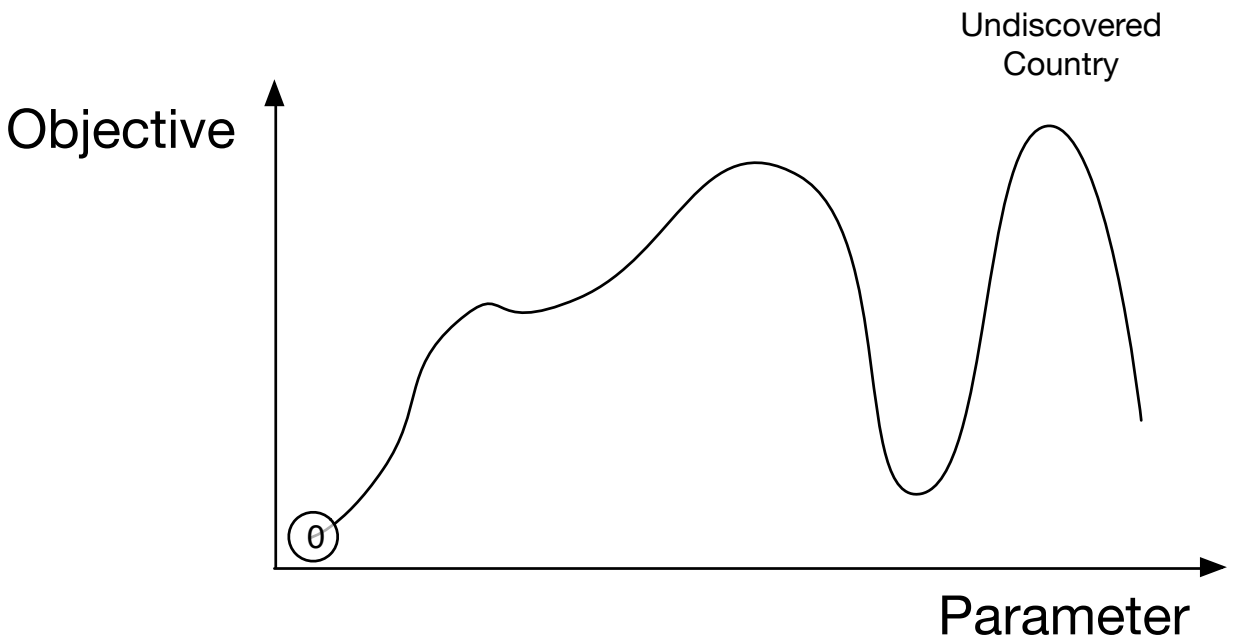
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

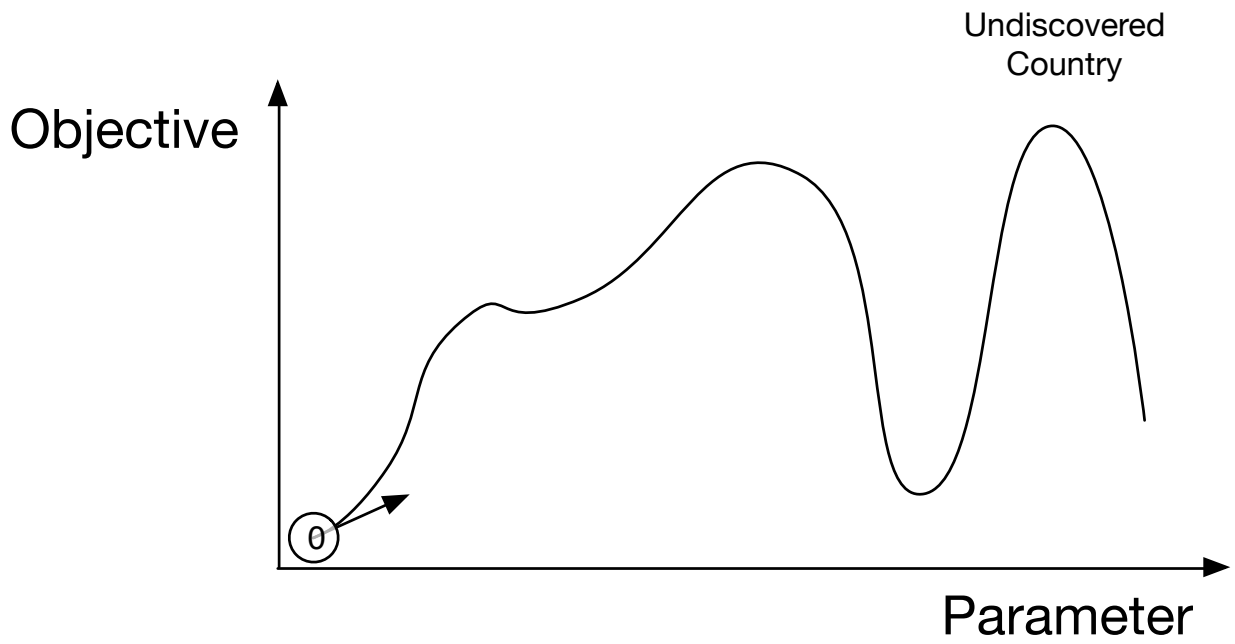
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

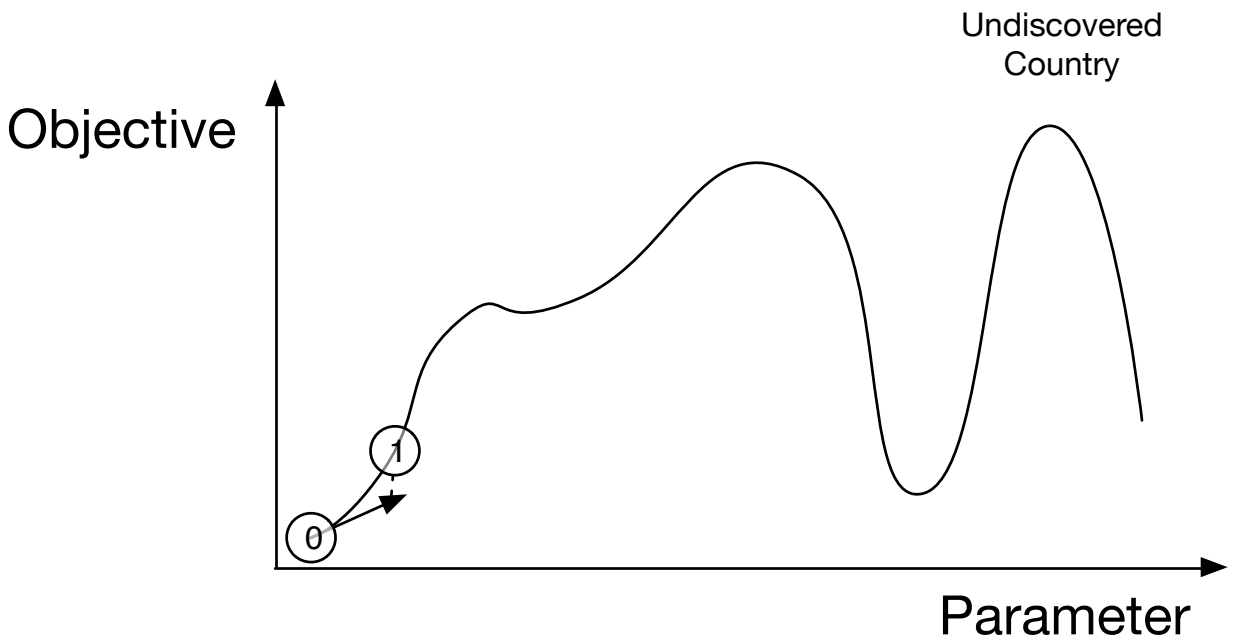
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

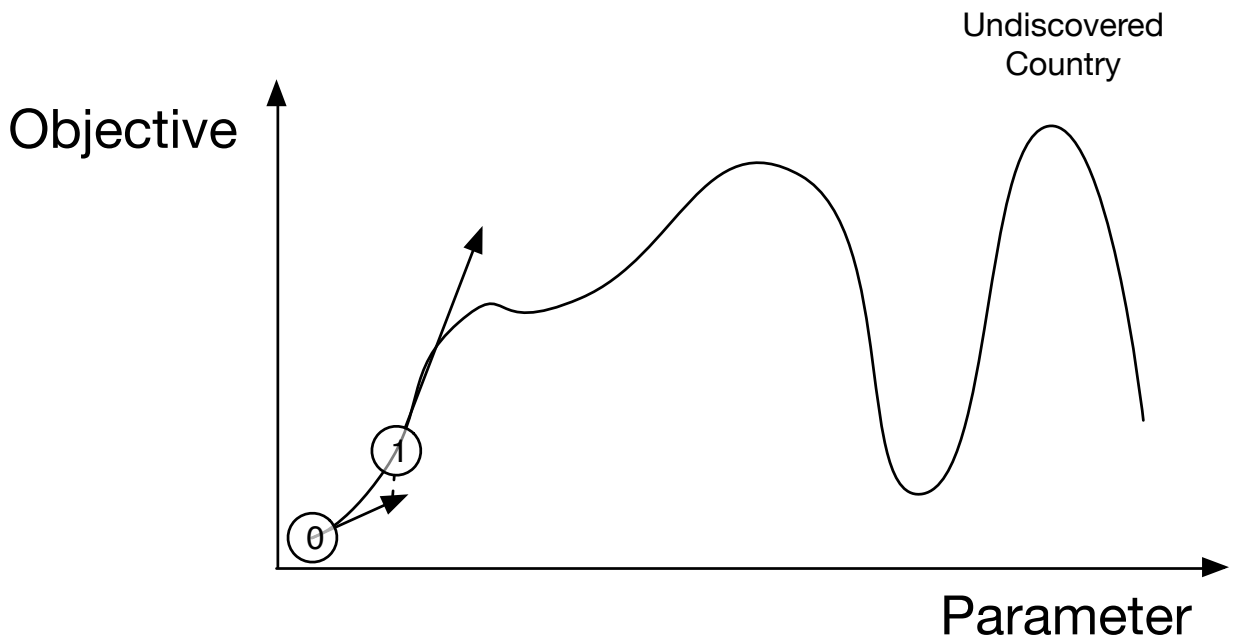
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

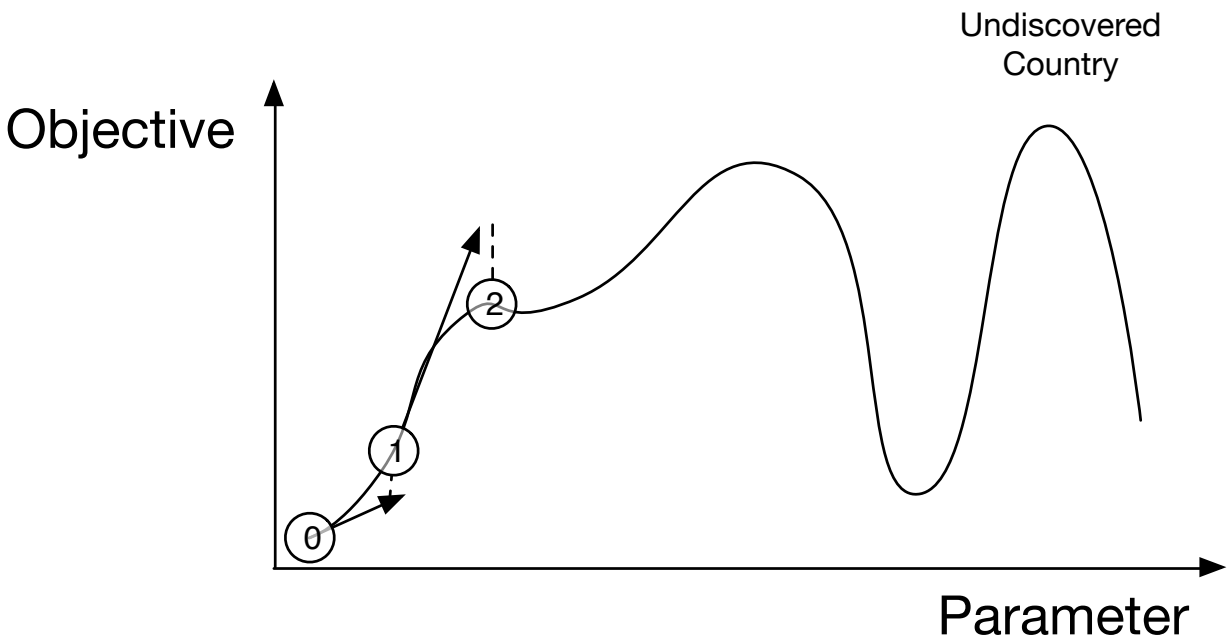
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

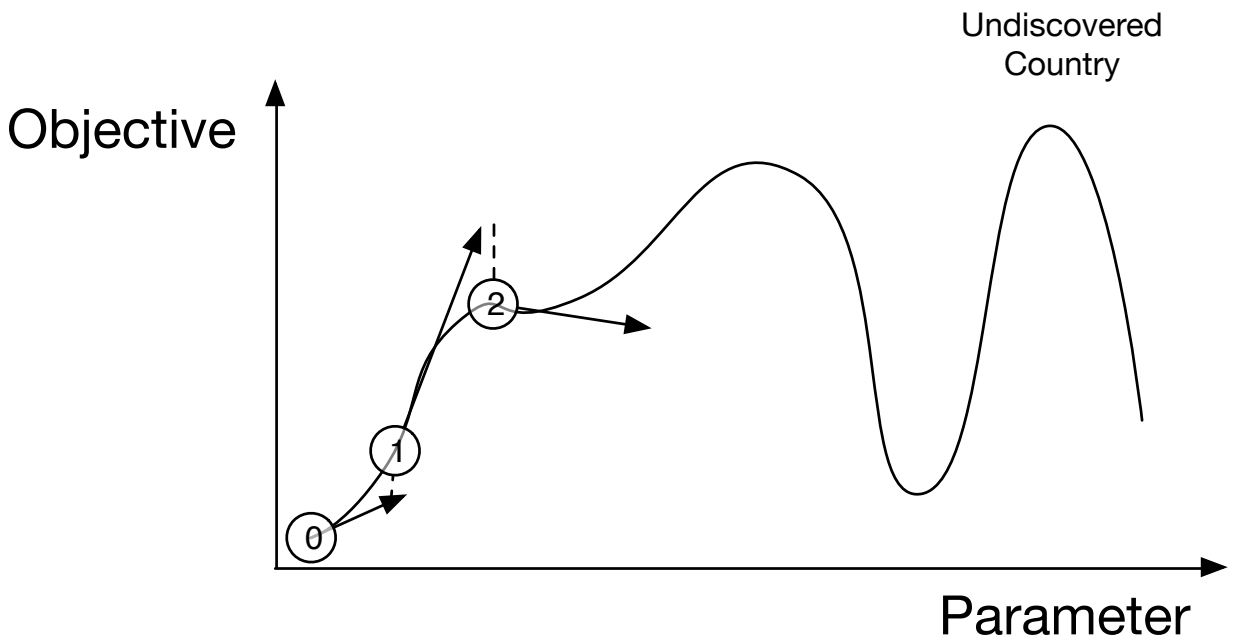
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

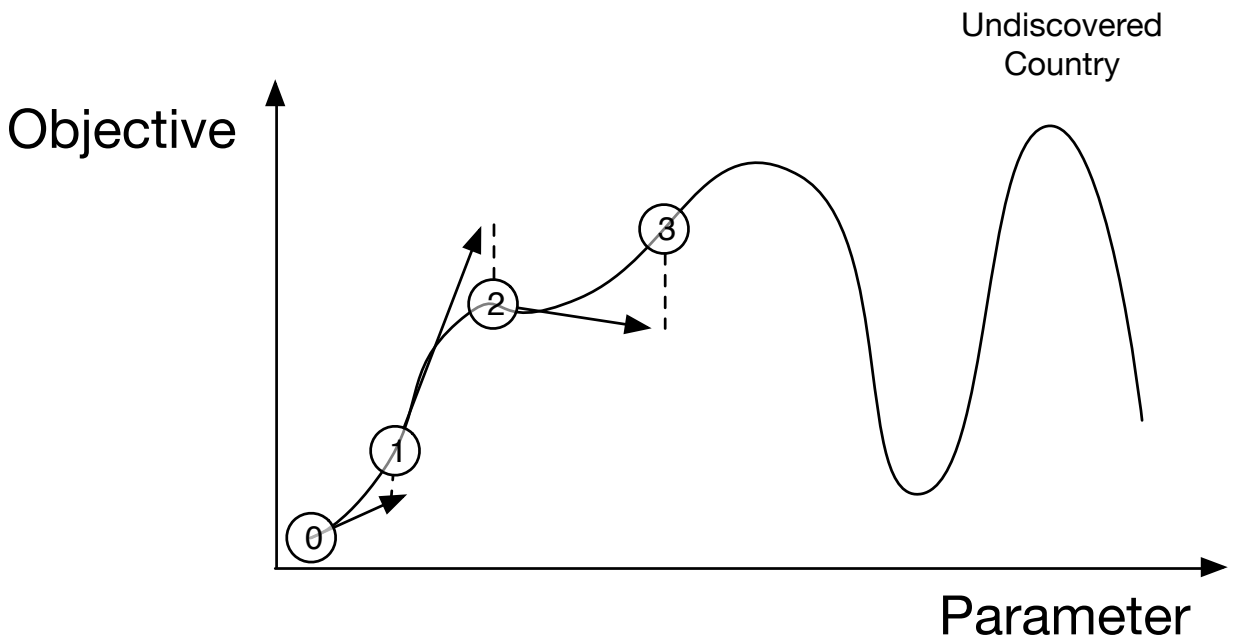
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

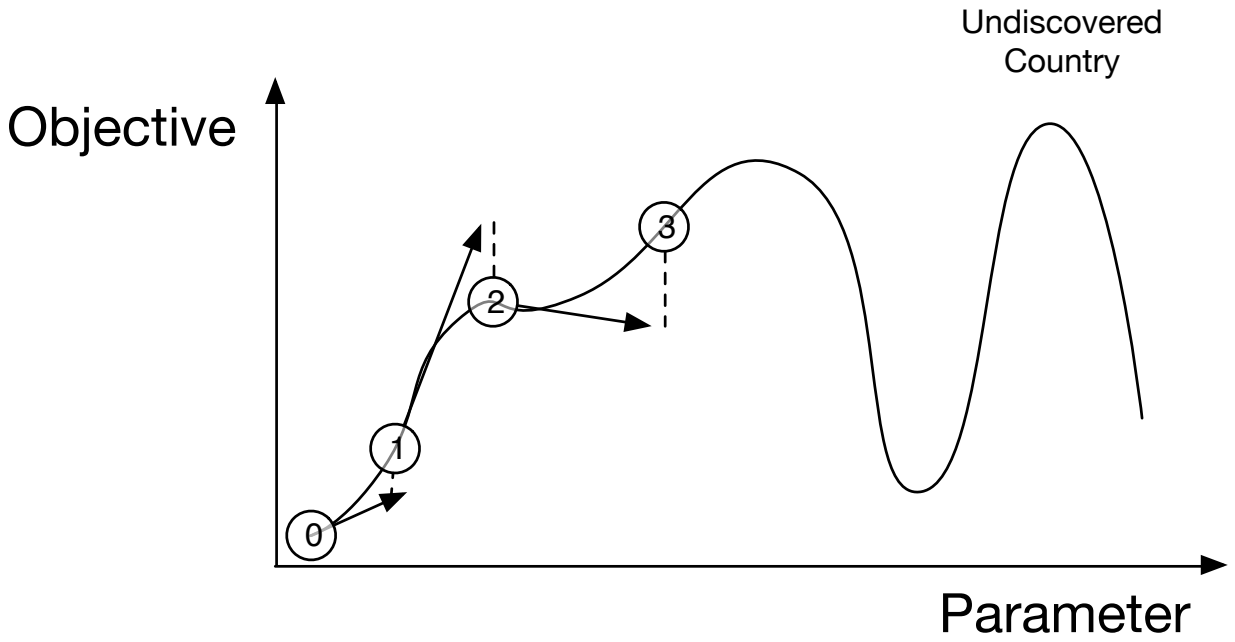
Optimize log likelihood with respect to variables β



Gradient Ascent (non-convex)

Goal

Optimize log likelihood with respect to variables β



Luckily, (vanilla) logistic regression is convex