



# Logistic Regression

## Introduction to Data Science Algorithms

Jordan Boyd-Graber and Michael Paul

ABC

## Logistic Regression: Objective Function

---

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

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

## Logistic Regression: Objective Function

---

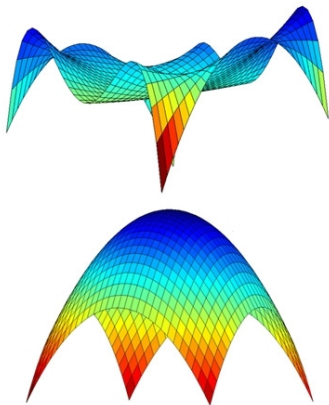
$$\ell \equiv \ln p(Y|X, \beta) = \sum_j \ln p(y^{(j)} | x^{(j)}, \beta) \quad (1)$$

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

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

## Convexity

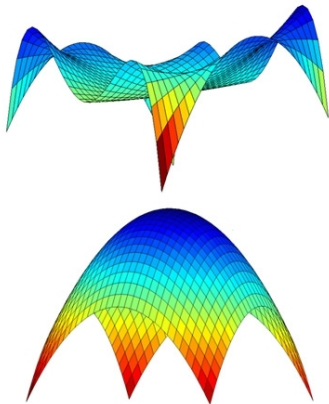
---



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

## Convexity

---



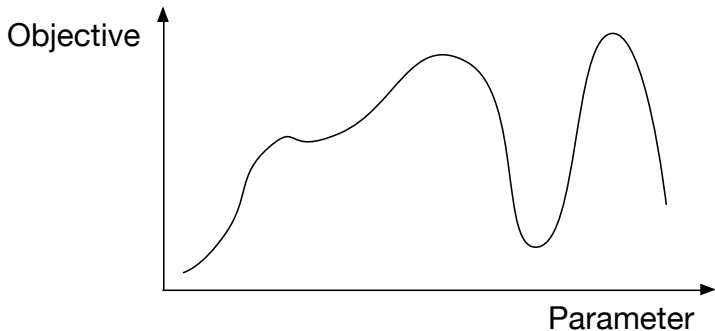
- Convex function
- Doesn't matter where you start, if you “walk up” objective
- Gradient!

## Gradient Ascent (non-convex)

---

### Goal

Optimize log likelihood with respect to variables  $\beta$

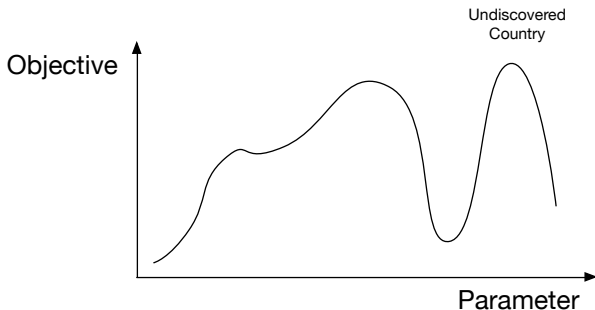


## Gradient Ascent (non-convex)

---

### Goal

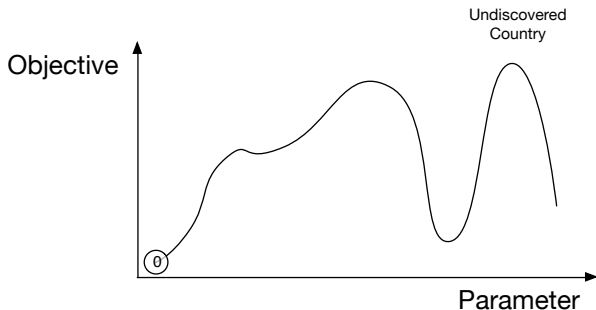
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

Optimize log likelihood with respect to variables  $\beta$

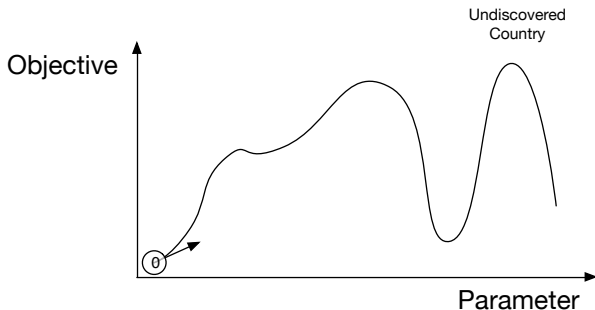




## Gradient Ascent (non-convex)

### Goal

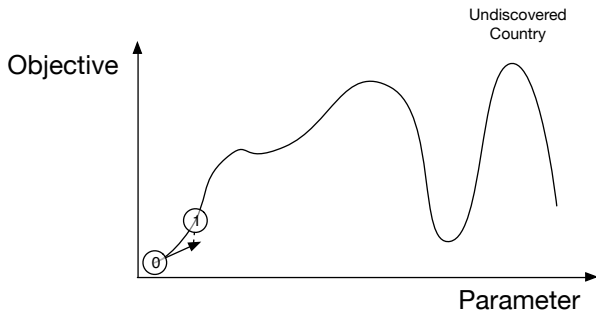
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

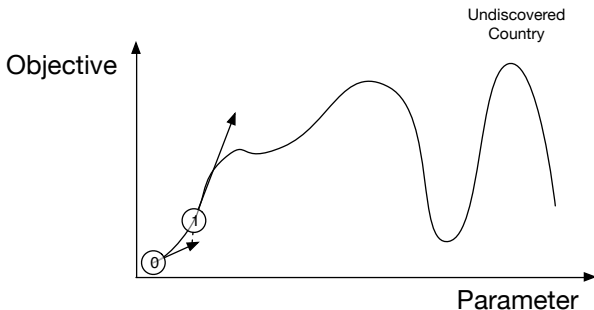
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

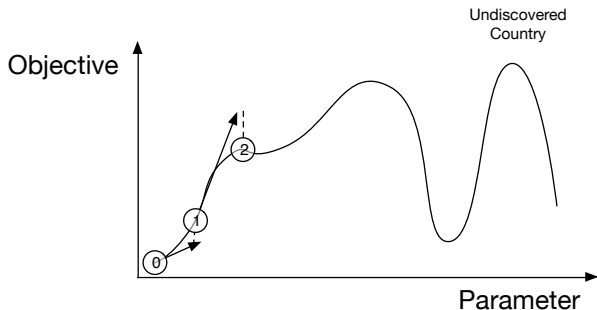
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

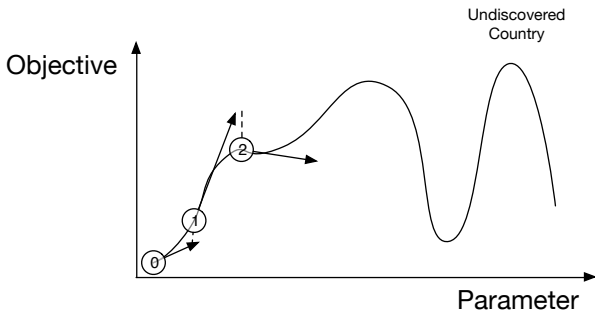
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

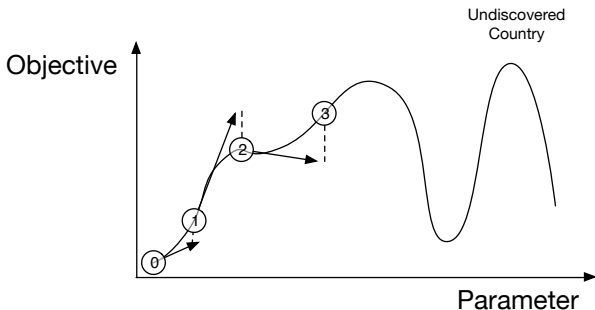
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

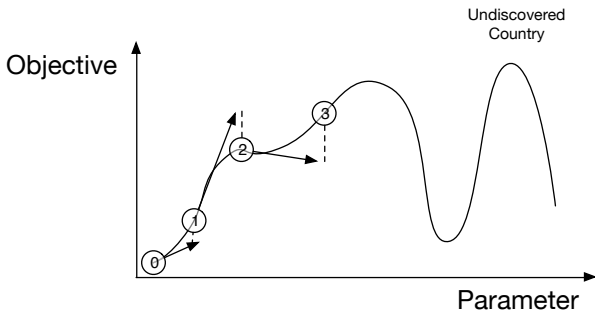
Optimize log likelihood with respect to variables  $\beta$



## Gradient Ascent (non-convex)

### Goal

Optimize log likelihood with respect to variables  $\beta$



Luckily, (vanilla) logistic regression is convex