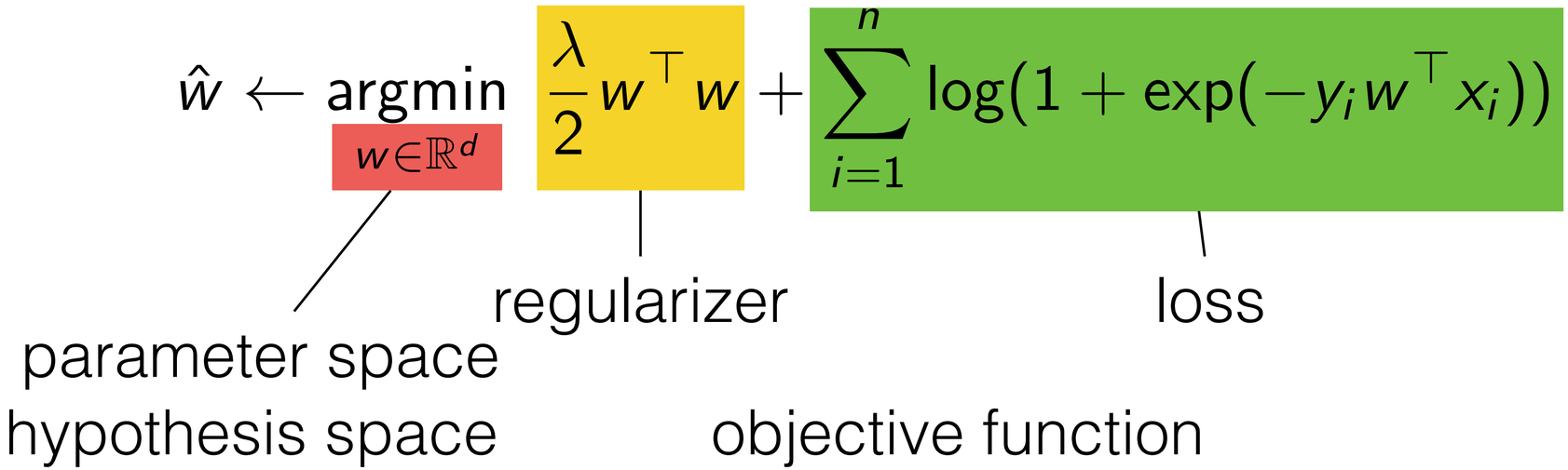
Learning Objective Functions

Machine Learning
CSx824/ECEx242
Bert Huang
Virginia Tech

Outline

- Optimization and machine learning
- Types of optimization problems
  - Unconstrained, constrained
  - Lagrange multipliers for constrained optimization
  - **Convex**, non-convex, discrete

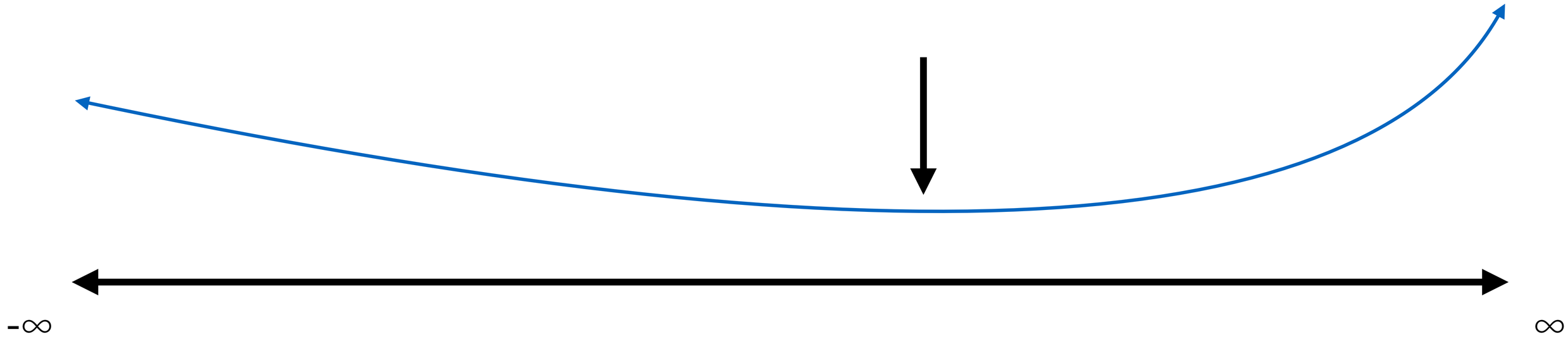
Optimization and Machine Learning



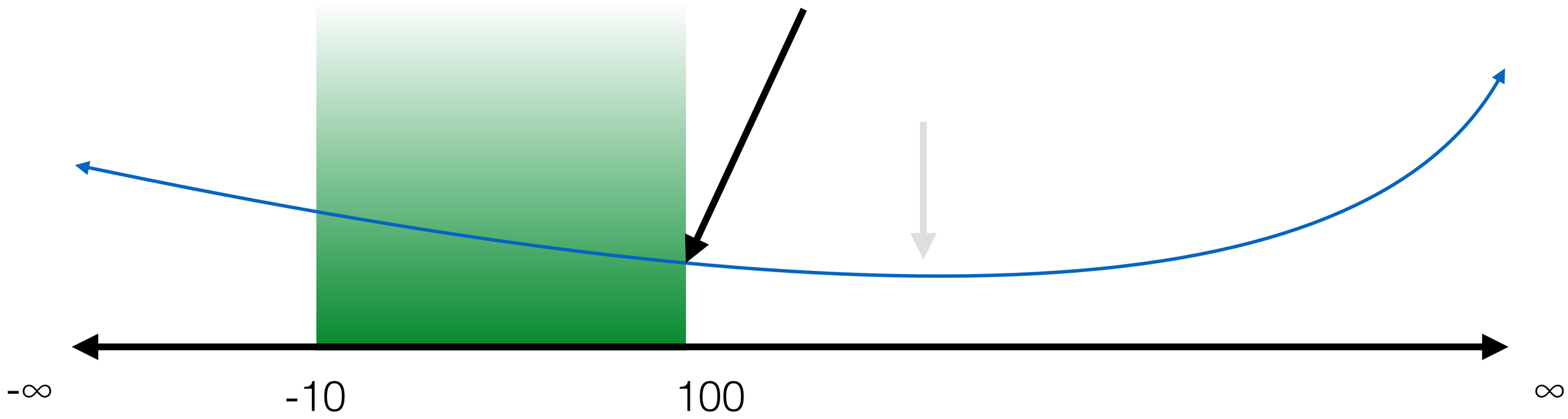
Types of Optimization Problems

- Unconstrained, constrained
- Lagrange multipliers for constrained optimization
- Convex, non-convex, discrete

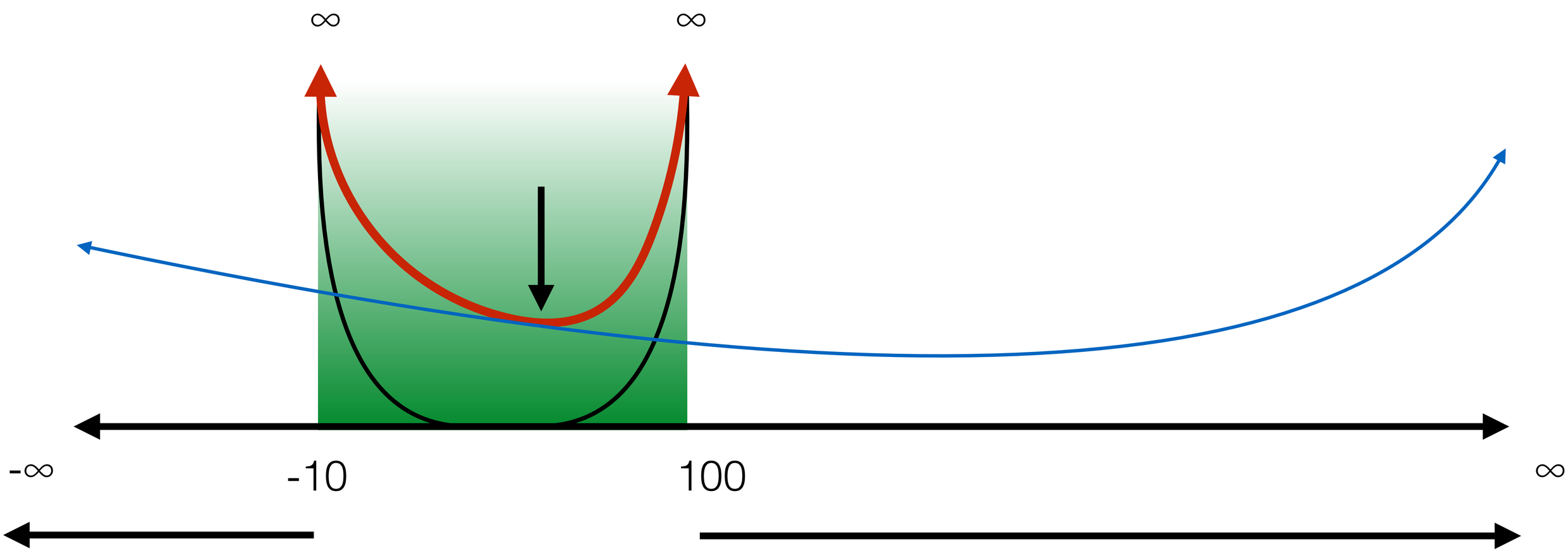
Unconstrained Optimization



Constrained Optimization



Penalty Functions



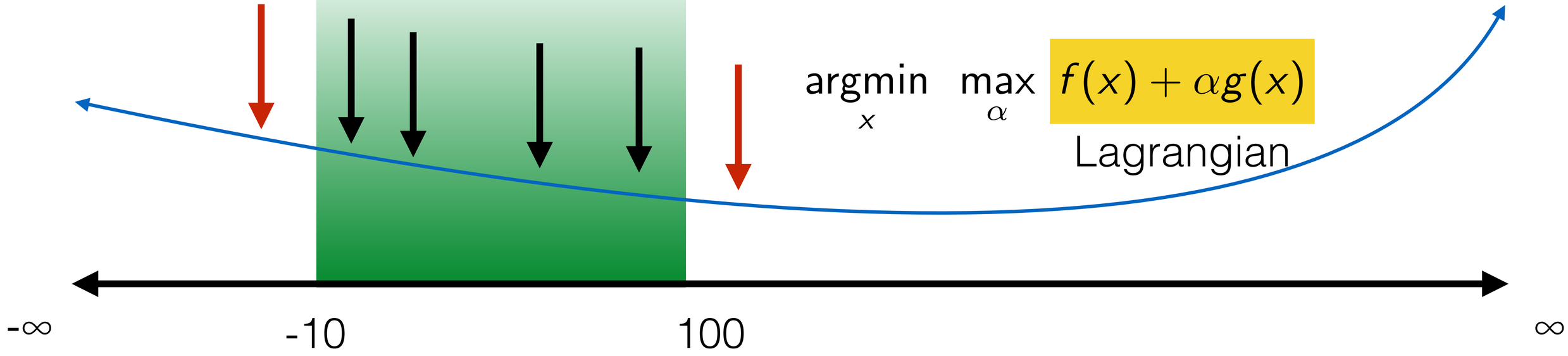
Lagrange Multipliers

$$\underset{x}{\text{argmin}}\, f(x) \text{ subject to } g(x) = 0$$

$$g(x) = I(x \in [-10, 100])$$

$$\underset{x}{\text{argmin}} \quad \max_{\alpha} \quad f(x) + \alpha g(x)$$

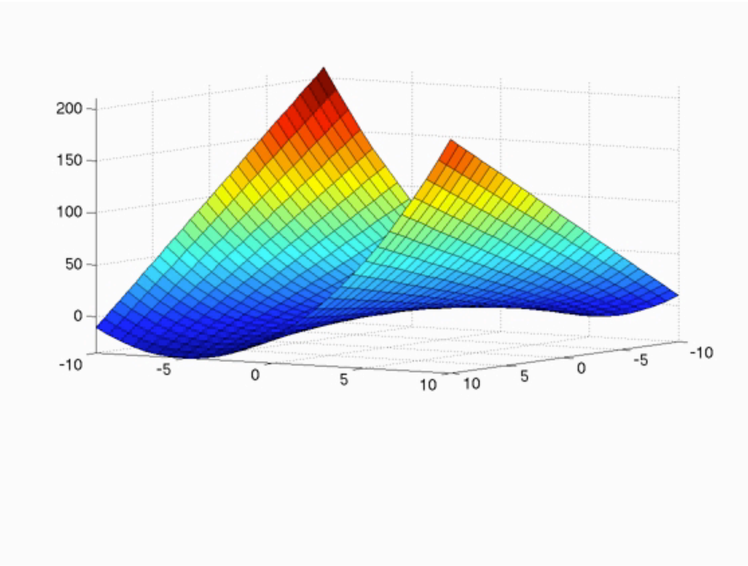
Lagrangian



Saddle Point Optimization

$$\underset{x}{\text{argmin}} \quad \underset{\alpha}{\max} \quad f(x) + \alpha g(x)$$

$x^2 + (1 - x)\, y$



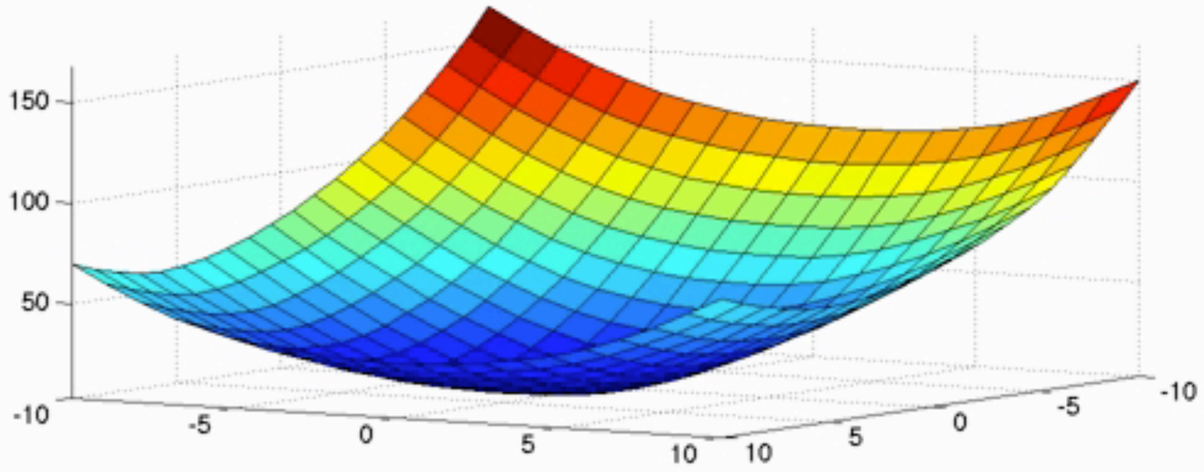
Convexity

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$$

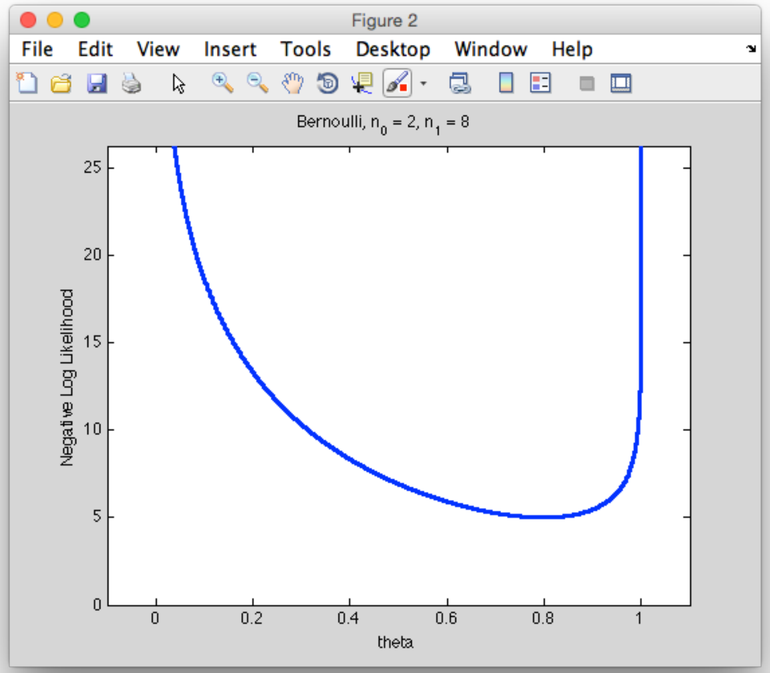
$$\alpha \in [0, 1]$$

$$x_1, x_2 \in \mathbb{R}^d$$

Convex Functions

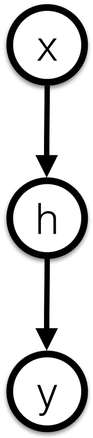


Logistic Regression Neg. Log Likelihood

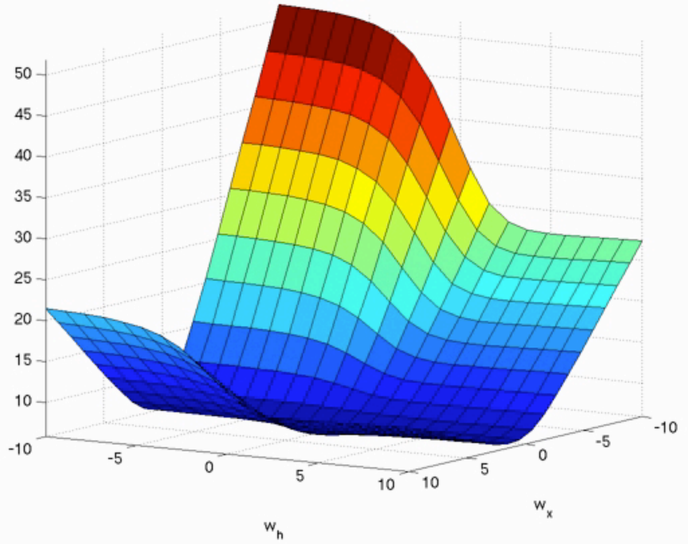


Bernoulli Negative Log Likelihood

Non-Convex Optimization

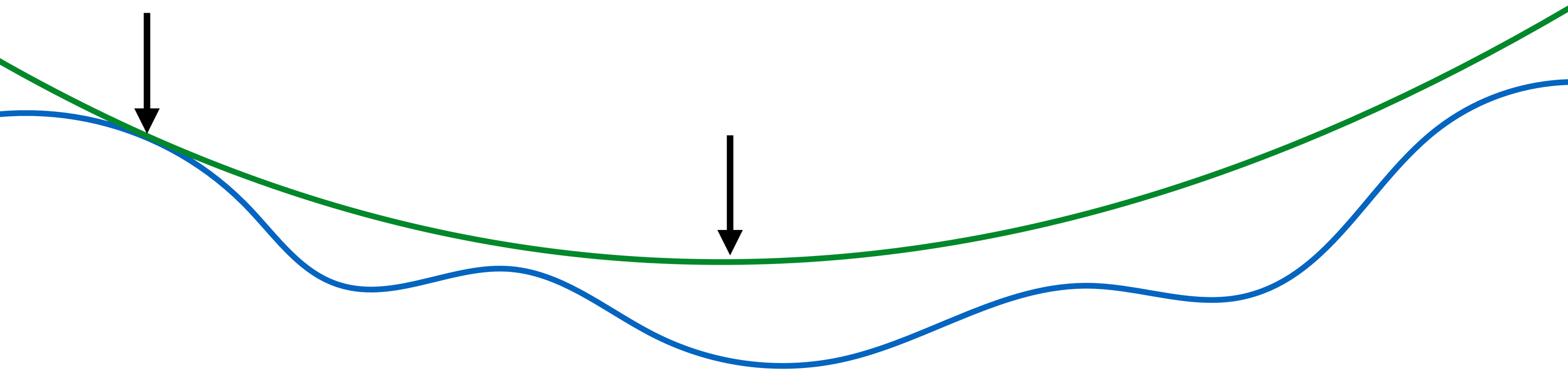


$$f(x) = \sigma(w_h \sigma(w_x x))$$

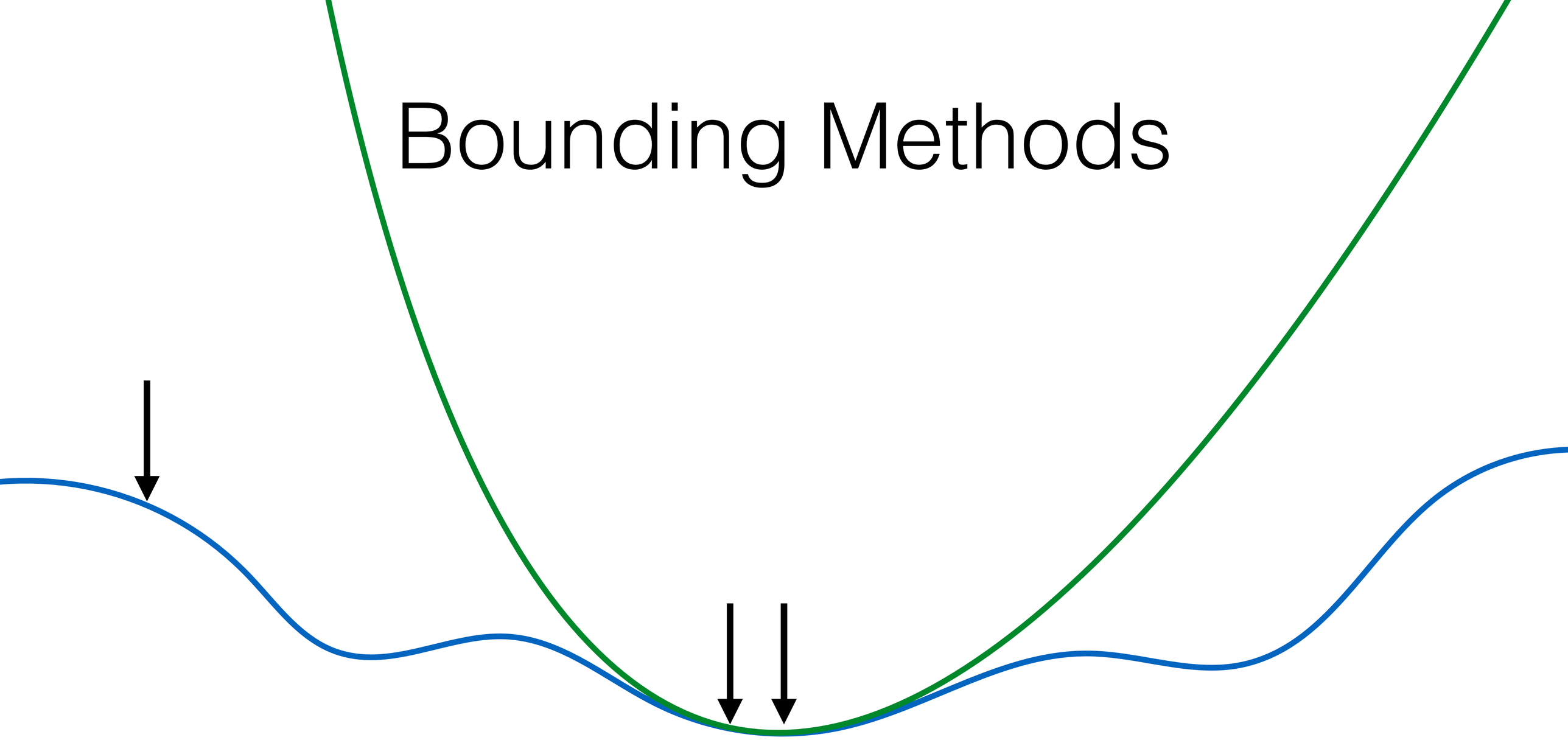


Local Optimization

Bounding Methods



Bounding Methods



Discrete Optimization

- Feature selection, decision tree learning
- Often intractable
- Relaxation to continuous optimization

Categorizing Machine Learning Techniques

Convex

Nonconvex

Constrained

support vector machines

Unconstrained

logistic regression

perceptron

naive Bayes

support vector machines

multilayered perceptron

expectation maximization

Discrete

decision tree learning

Saddle Point

support vector machines

expectation maximization