

# Numerical Computation

Lecture slides for Chapter 4 of *Deep Learning*

[www.deeplearningbook.org](http://www.deeplearningbook.org)

Ian Goodfellow

2016-09-26

# Overflow and Underflow

$$\text{softmax}(\mathbf{x})_i = \frac{\exp(x_i)}{\sum_{j=1}^n \exp(x_j)}. \quad (4.1)$$

The exponentiation can underflow when the argument is very negative, or overflow when it is very positive.

# Condition Number

$$\max_{i,j} \left| \frac{\lambda_i}{\lambda_j} \right|. \quad (4.2)$$

# Gradient Descent

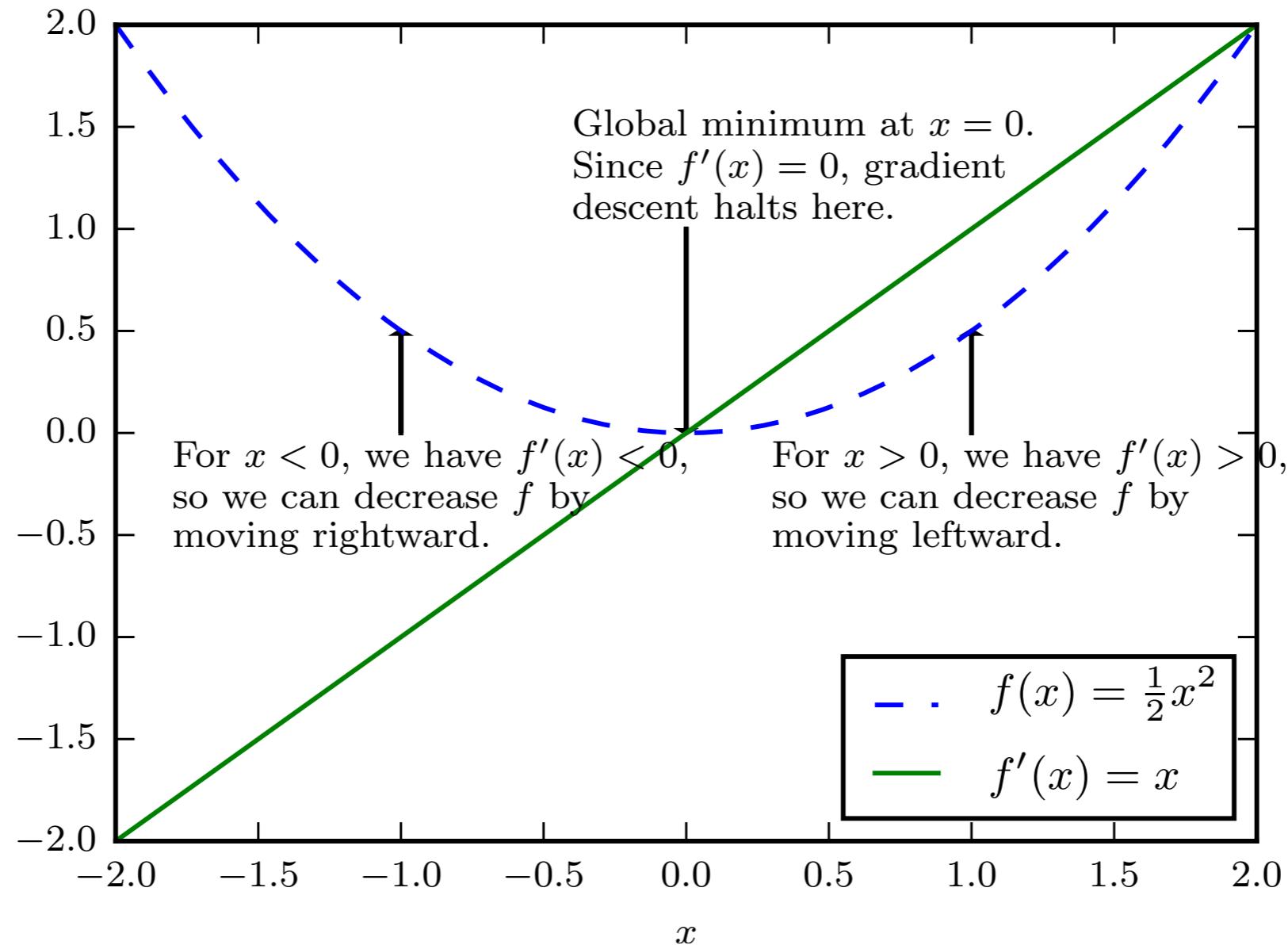
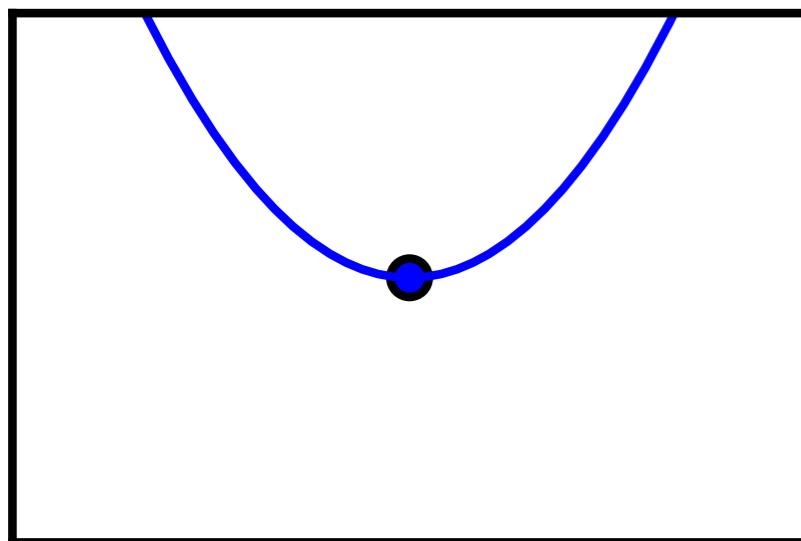


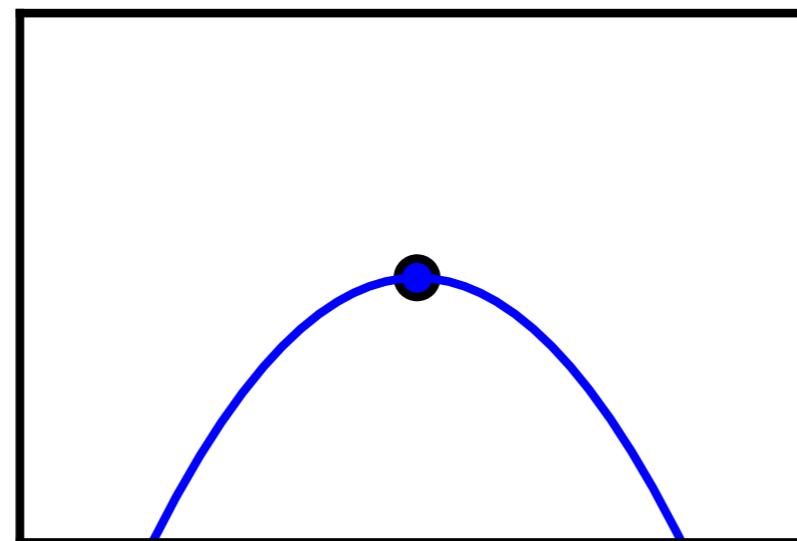
Figure 4.1

# Critical Points

Minimum



Maximum



Saddle point

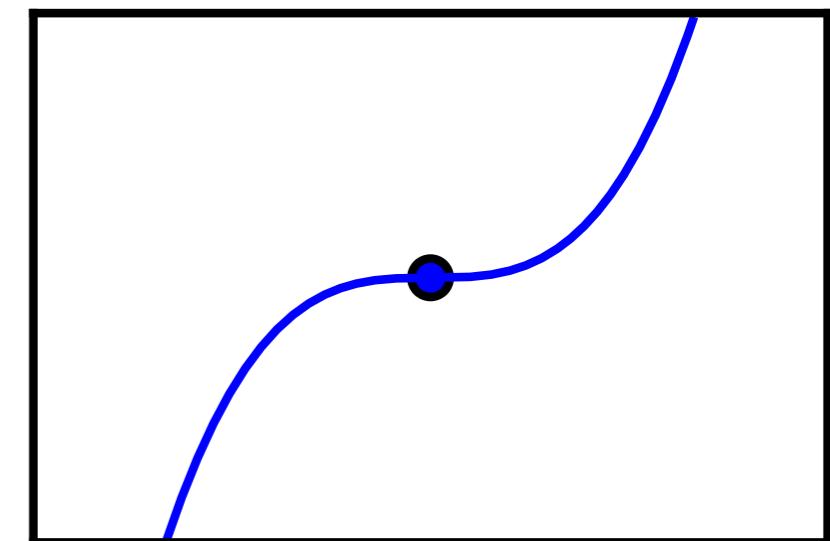


Figure 4.2

# Approximate Optimization

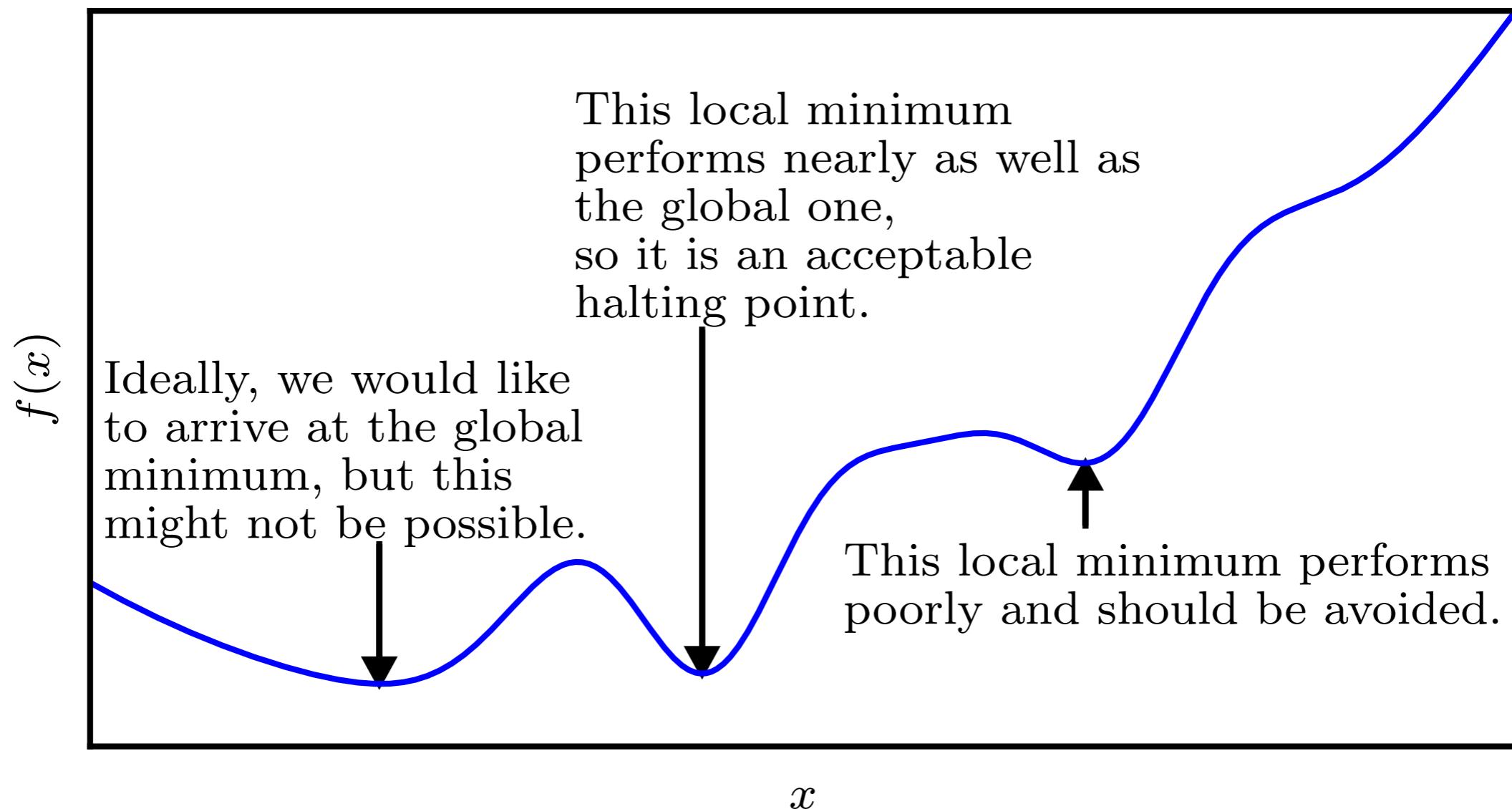


Figure 4.3

# Curvature

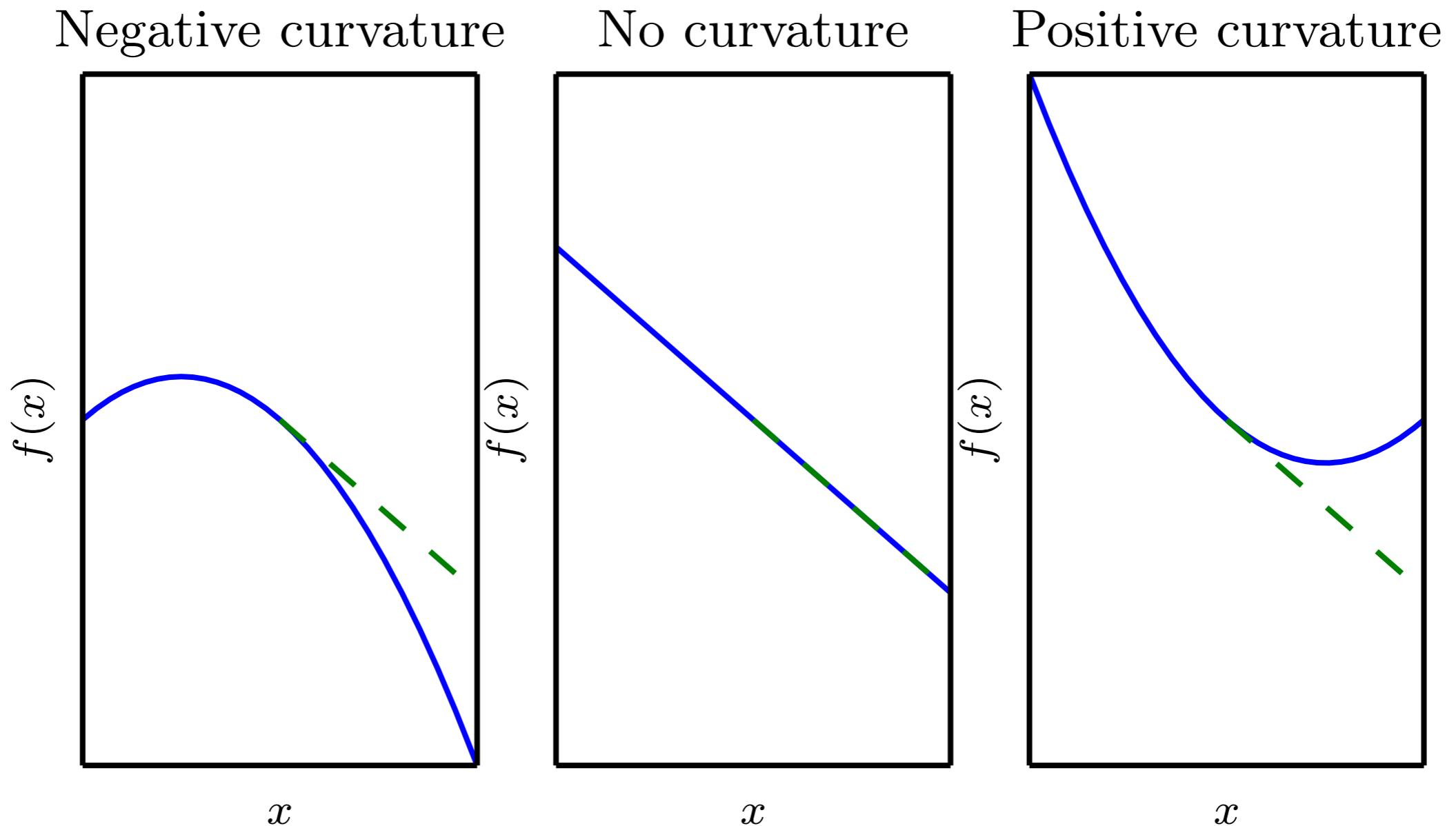


Figure 4.4

# Saddle Points

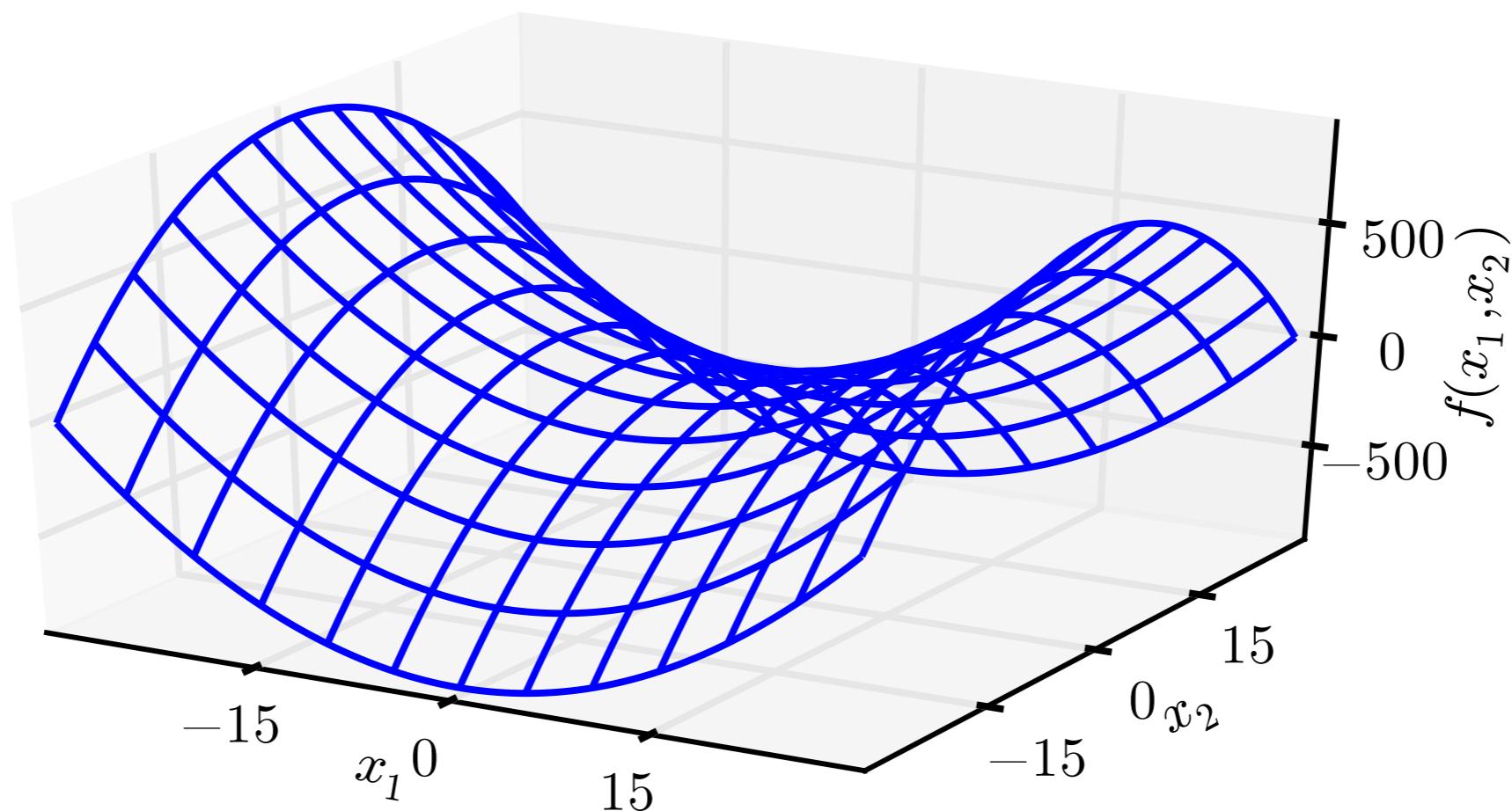


Figure 4.5

# Gradient Descent and Poor Conditioning

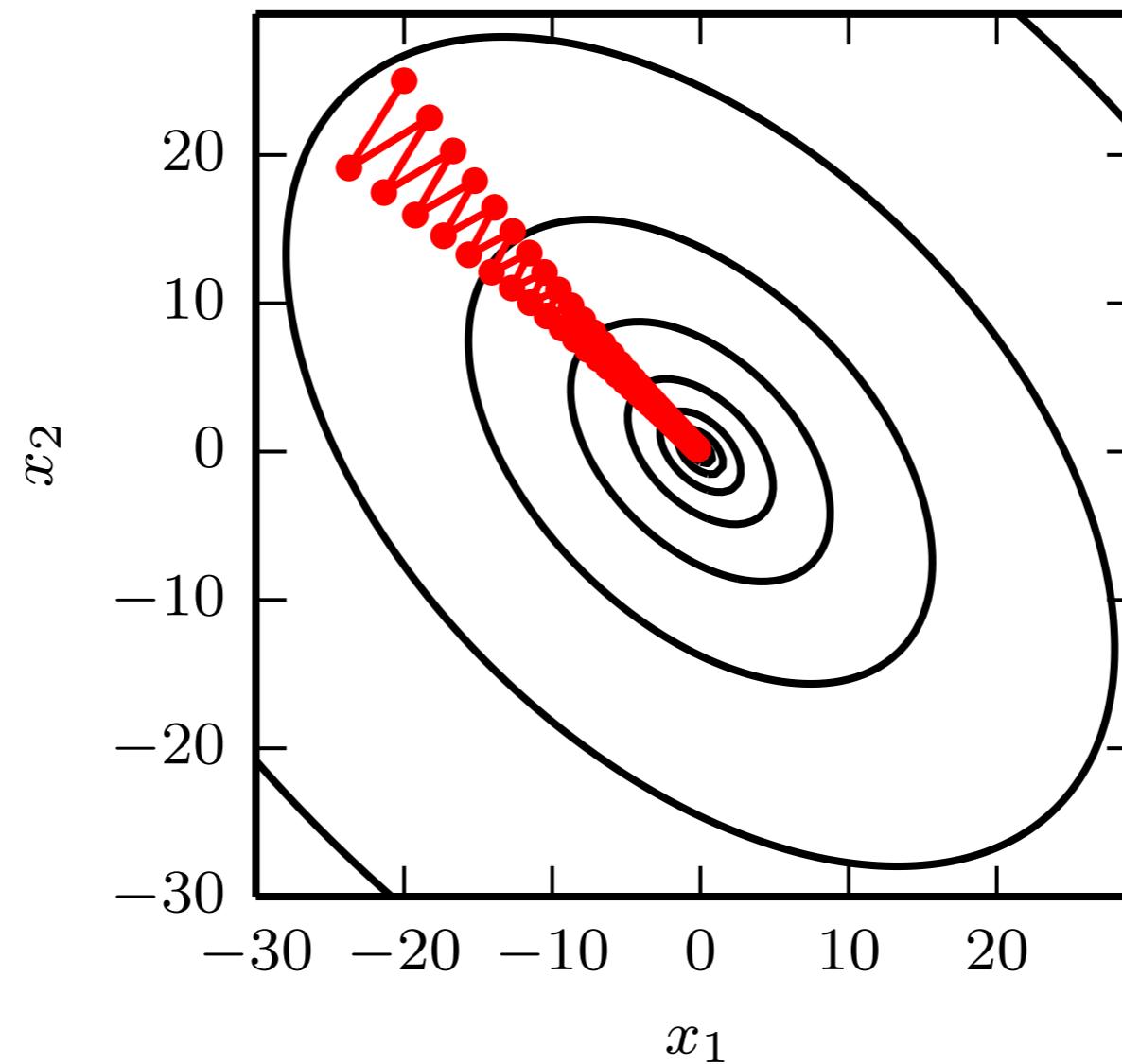


Figure 4.6

# KKT Multipliers

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{\lambda}} \max_{\boldsymbol{\alpha}, \boldsymbol{\alpha} \geq 0} -f(\boldsymbol{x}) + \sum_i \lambda_i g^{(i)}(\boldsymbol{x}) + \sum_j \alpha_j h^{(j)}(\boldsymbol{x}). \quad (4.19)$$