# Artificial Intelligence in Medicine

## Learning (1)

Nir Friedman and Tommy Kaplan 7/11/22

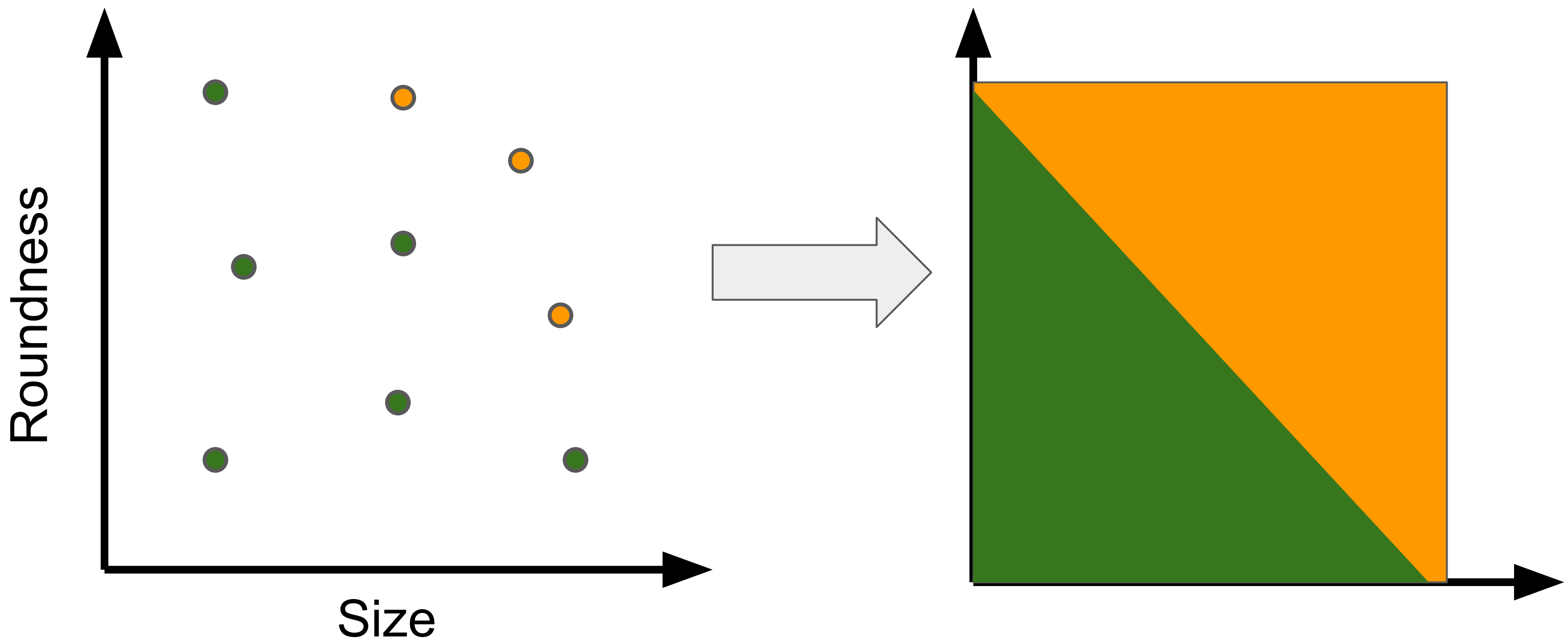
האוניברסיטה
העברית
בירושלים
THE HEBREW
UNIVERSITY
OF JERUSALEM

“It is a capital mistake to theorize before one has data.”

Sherlock Holmes

# What is learning?

Examples → Classifier → Test



# Desiderata

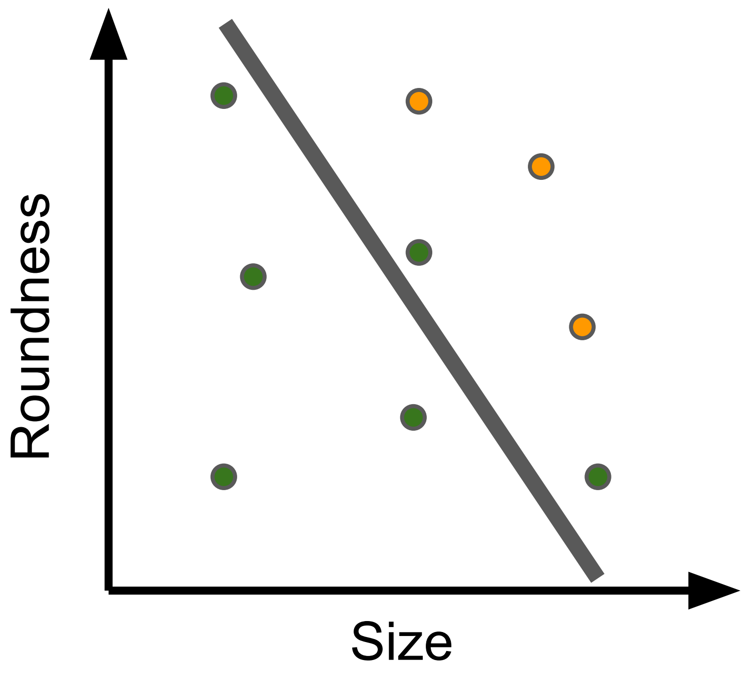
- Good performance on examples
- Good performance in future examples
- Computational costs
- Simplicity
- …

# Evaluating Performance

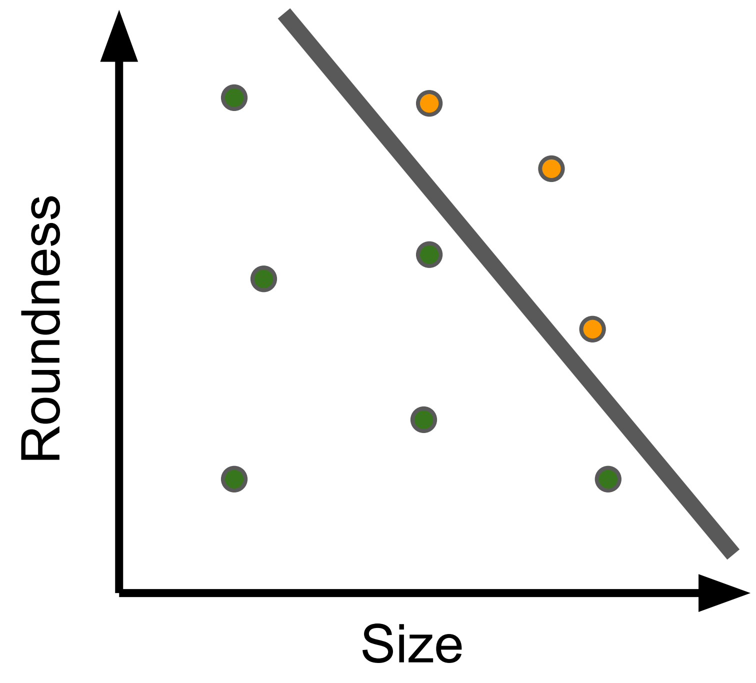
Loss function

loss(classifier,data) = …

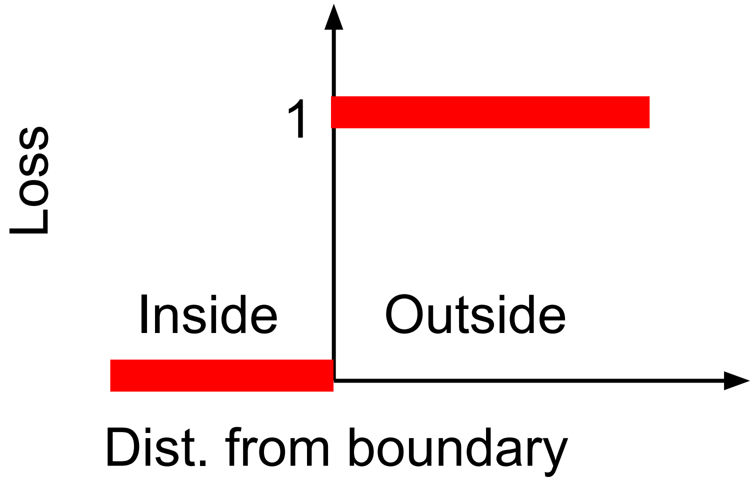
2 classification errors



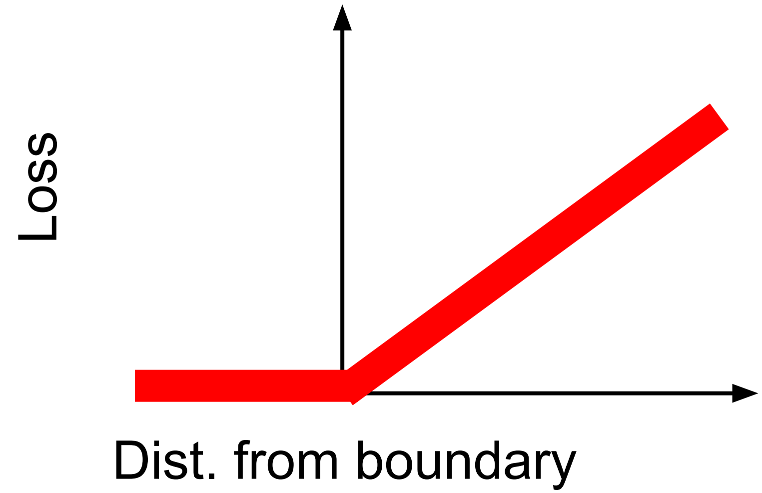

0 classification errors



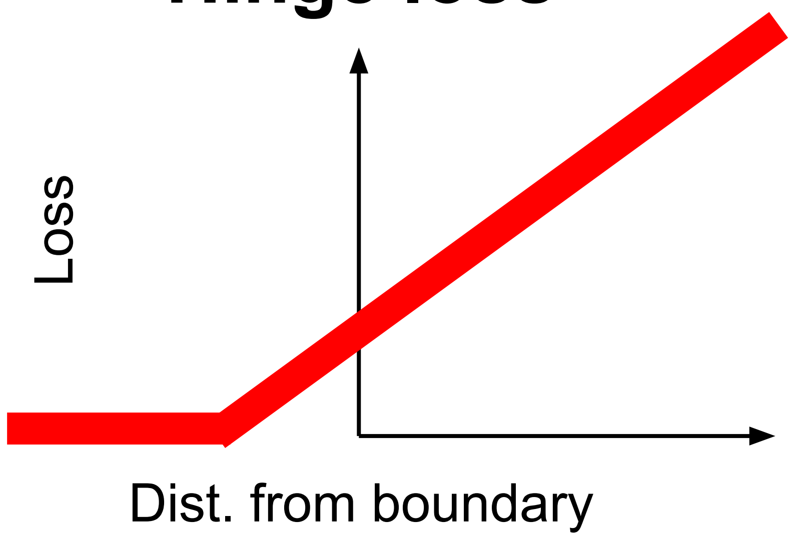
## 0/1 loss



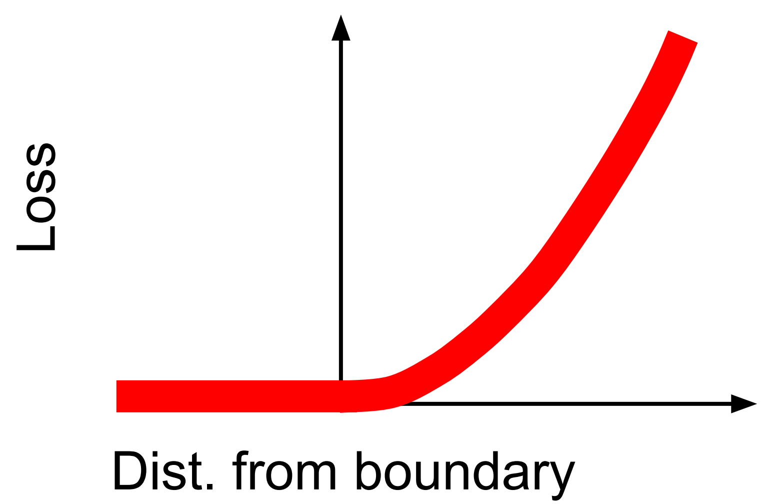

## Linear loss



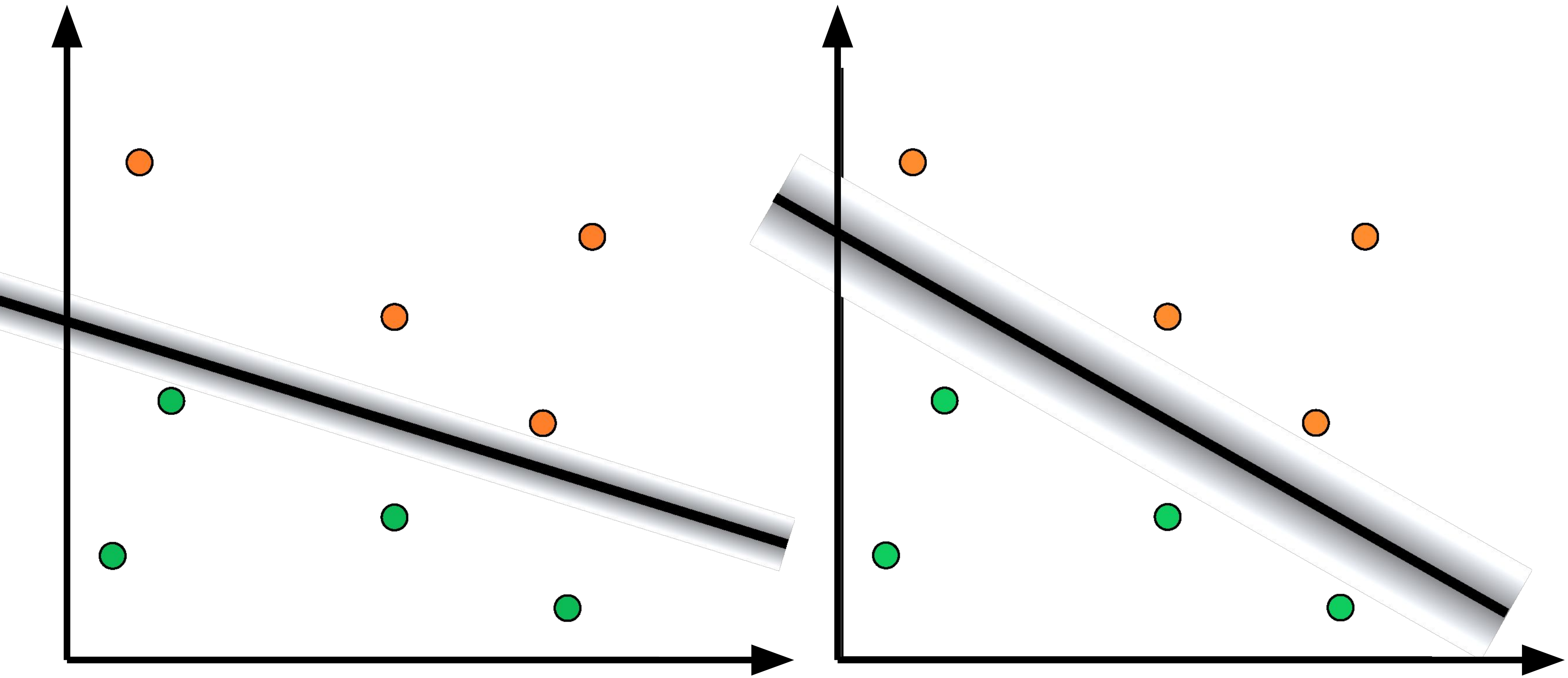

## Hinge loss




## Quadratic loss



# Beyond 0/1 loss



# Learning as optimization

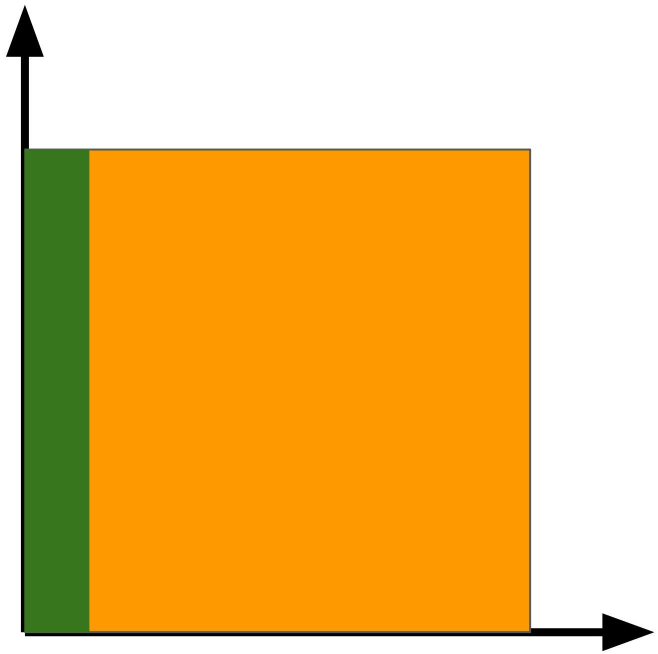
Given

- Data
- Family of classifiers
- Loss function

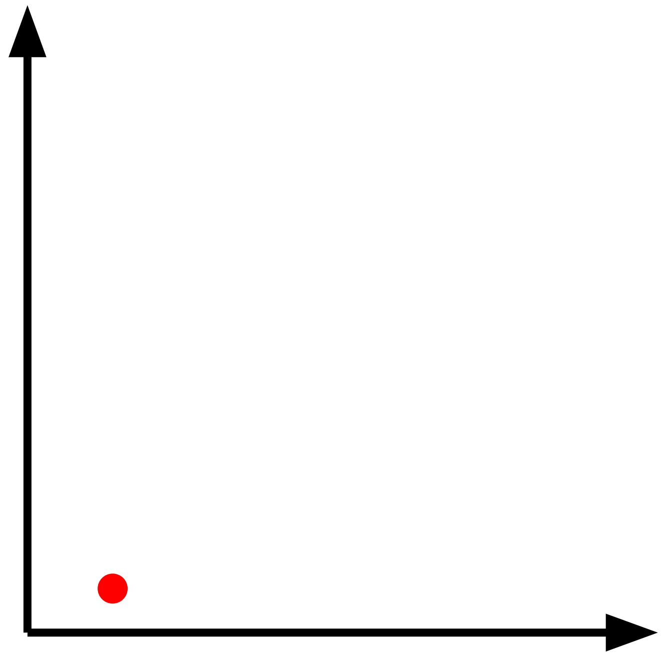
Find classifier with minimal loss

# Parameter space

Hypothesis space

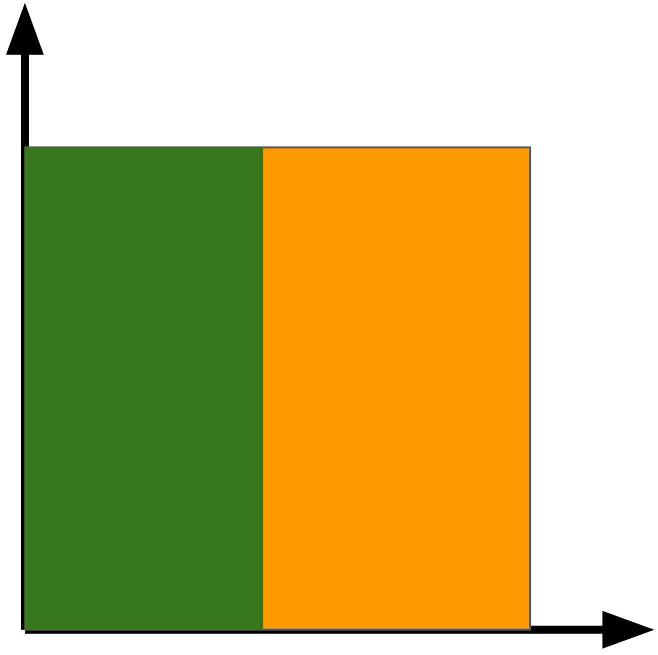


Parameter space

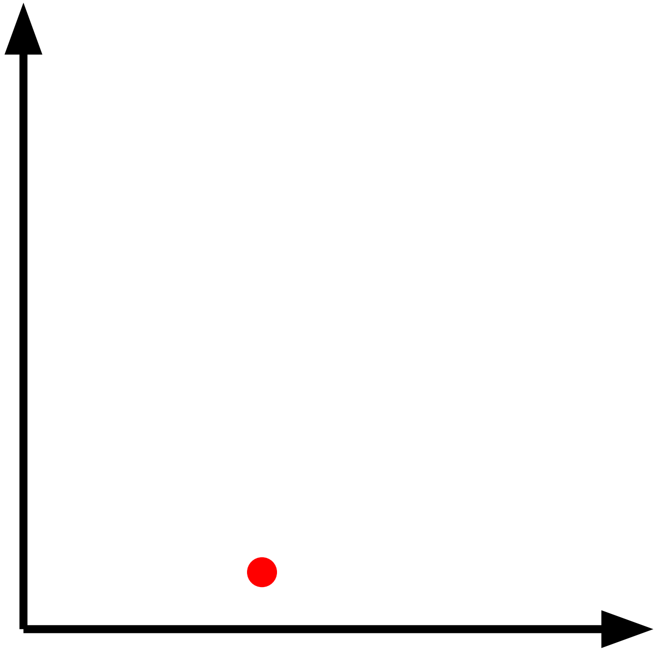


# Parameter space

Hypothesis space

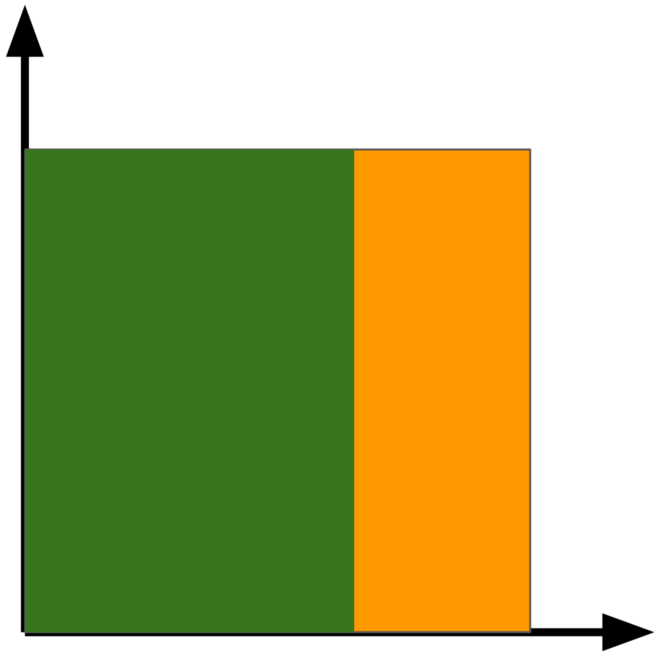


Parameter space

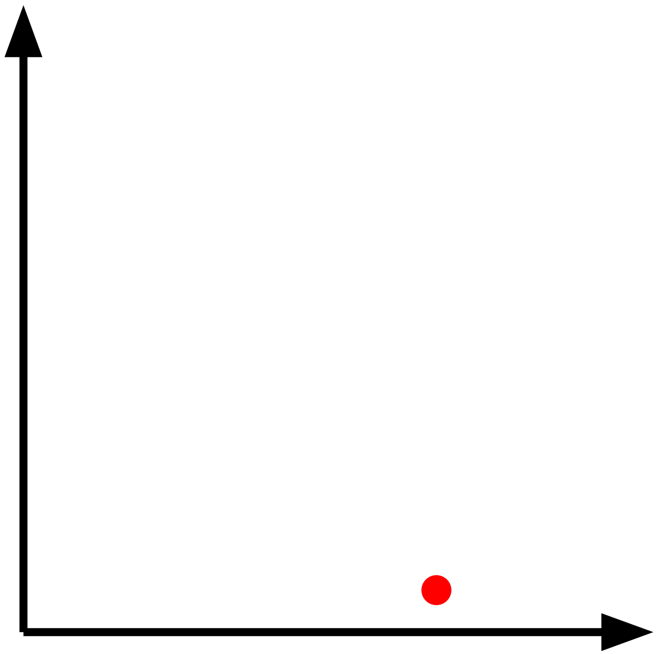


# Parameter space

Hypothesis space

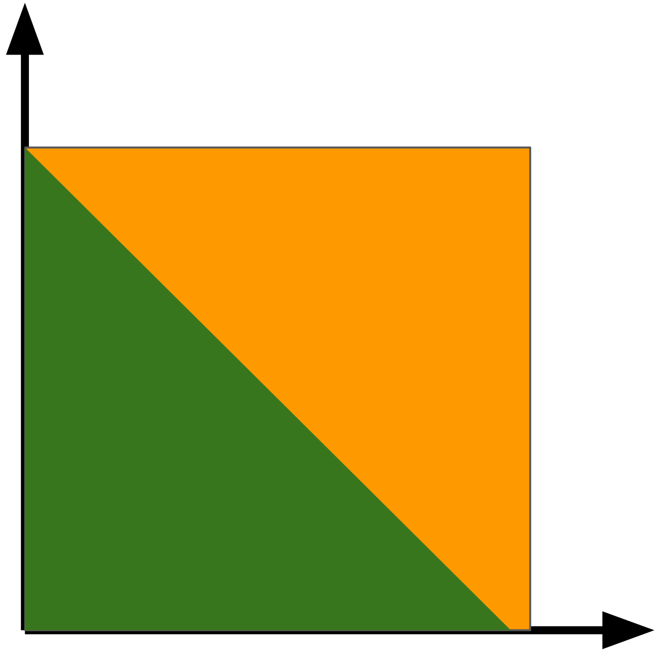


Parameter space

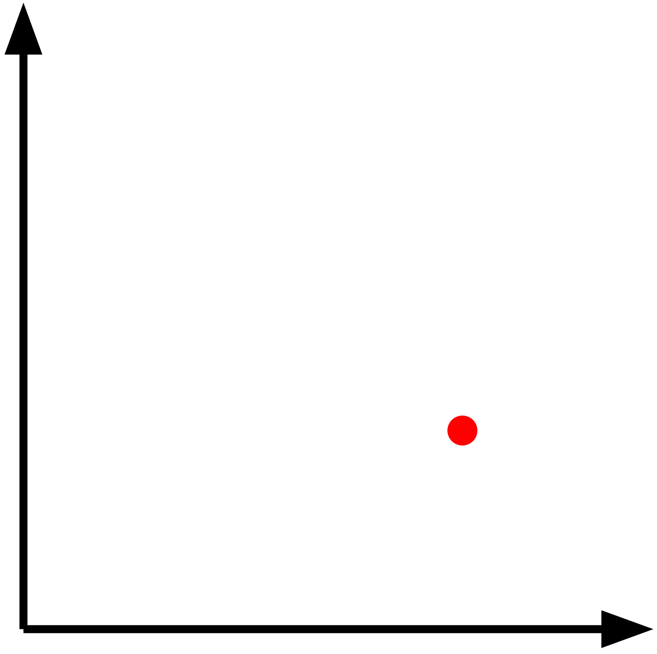


# Parameter space

Hypothesis space

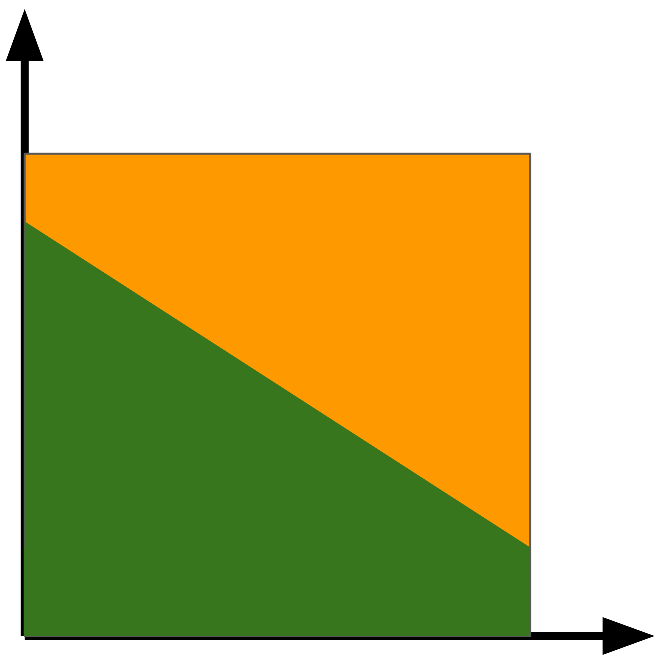


Parameter space

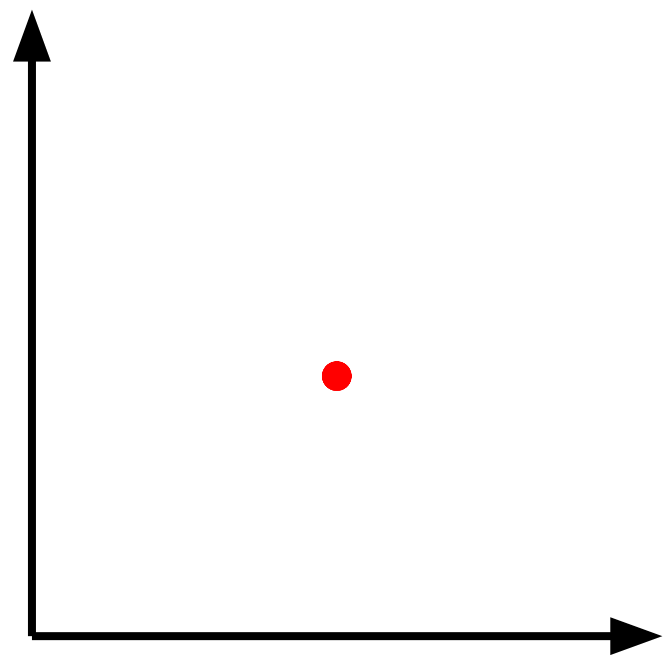


# Parameter space

Hypothesis space

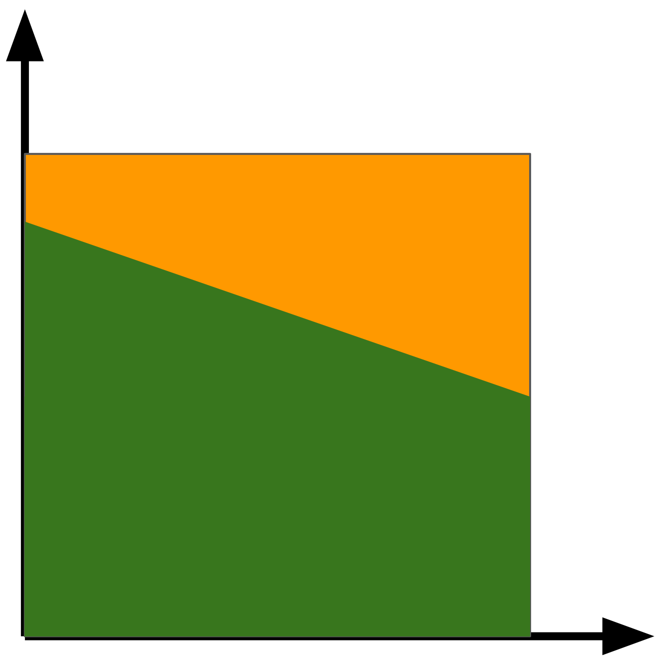


Parameter space

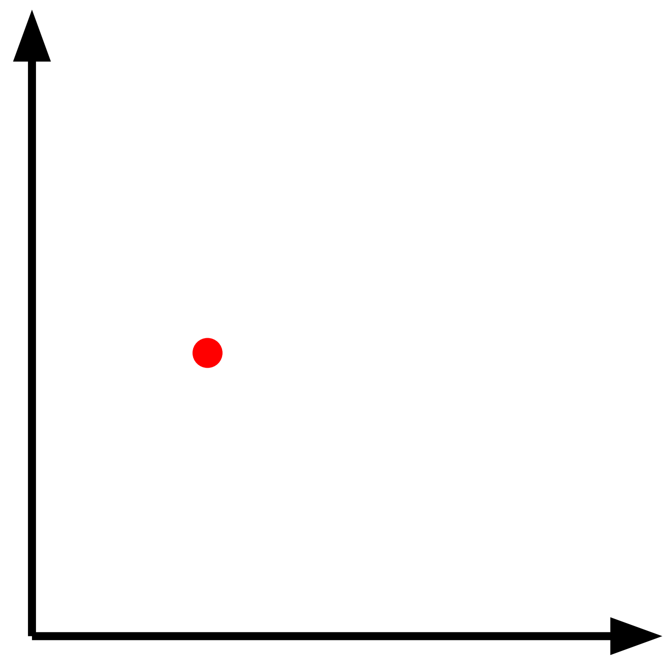


# Parameter space

Hypothesis space



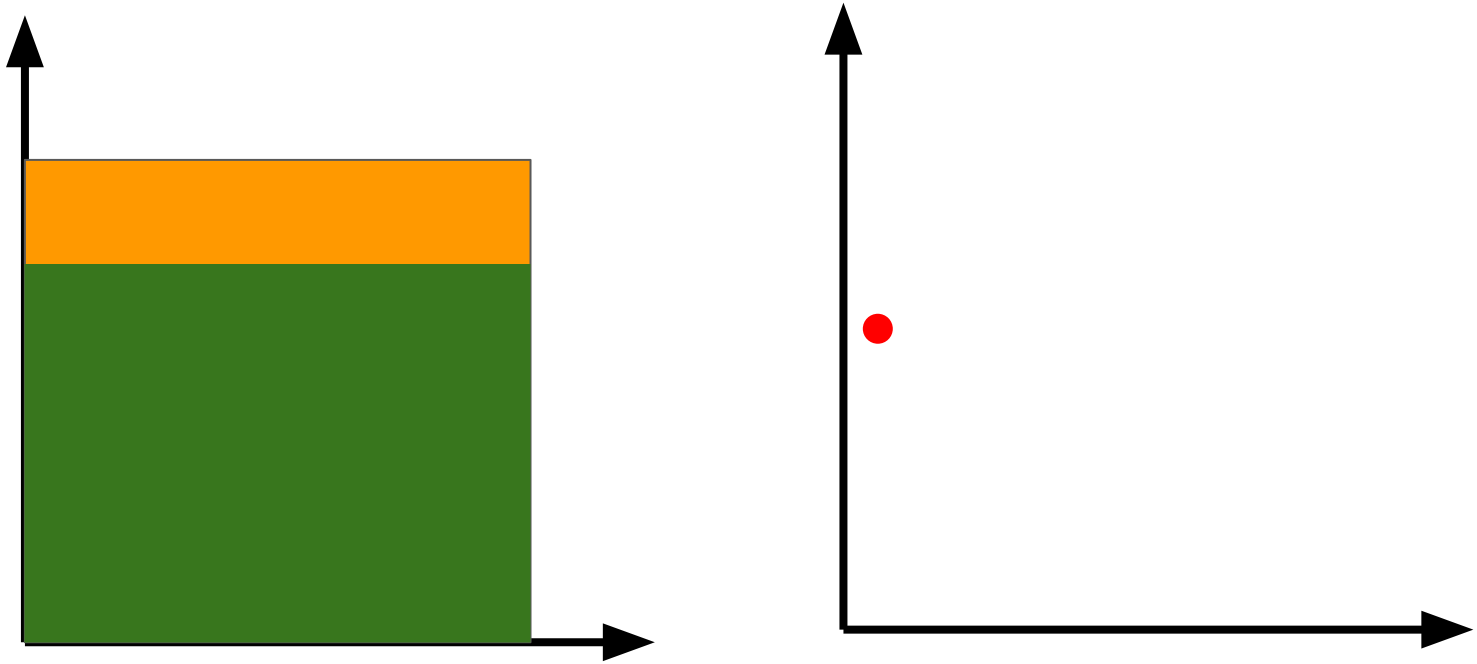
Parameter space



# Parameter space

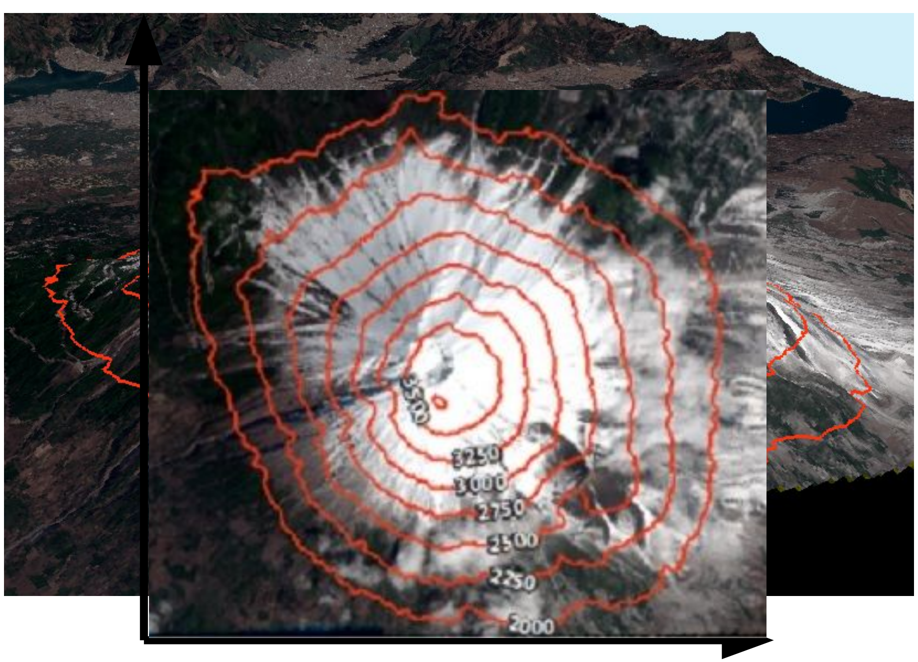
Hypothesis space

Parameter space

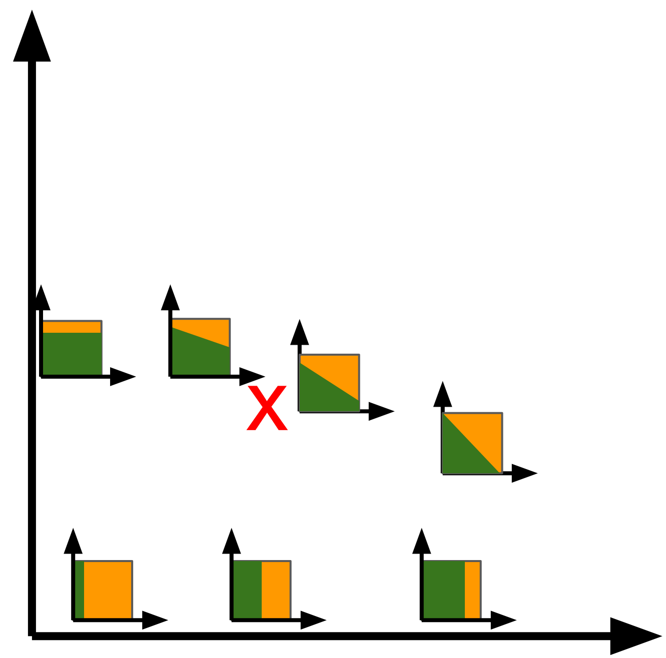


# Parameter space optimization

## Optimum search



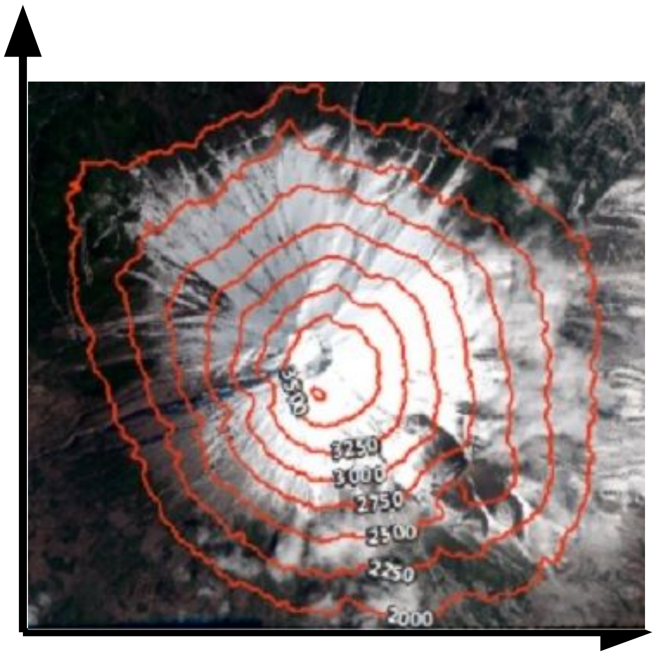

## Parameter space



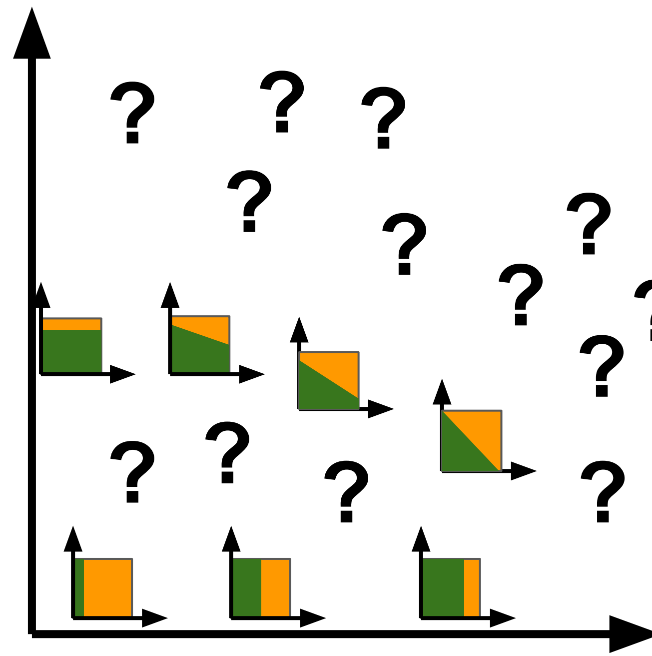
# Optimization 101

Dealing with partial knowledge

Optimum search

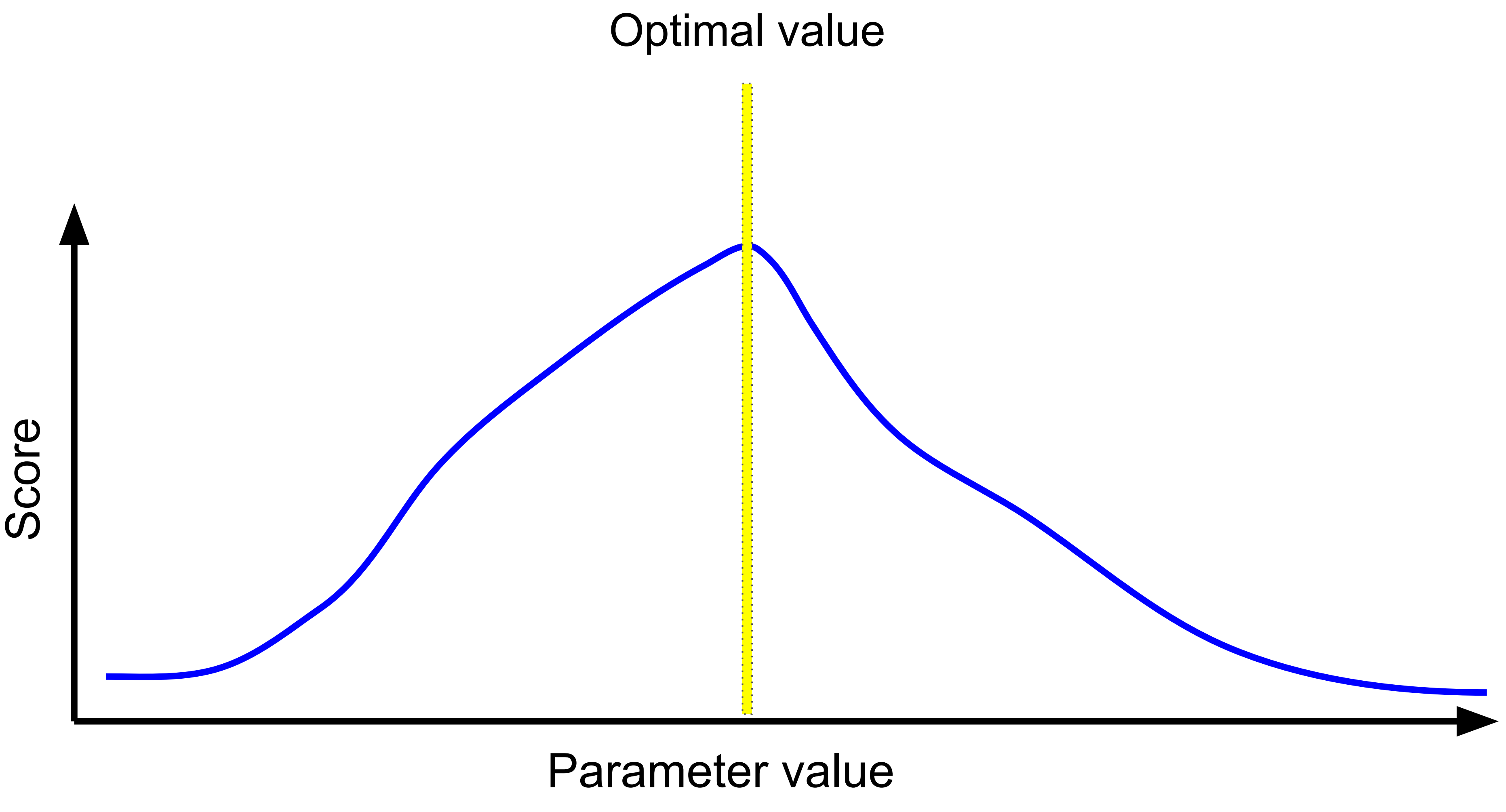



Parameter space



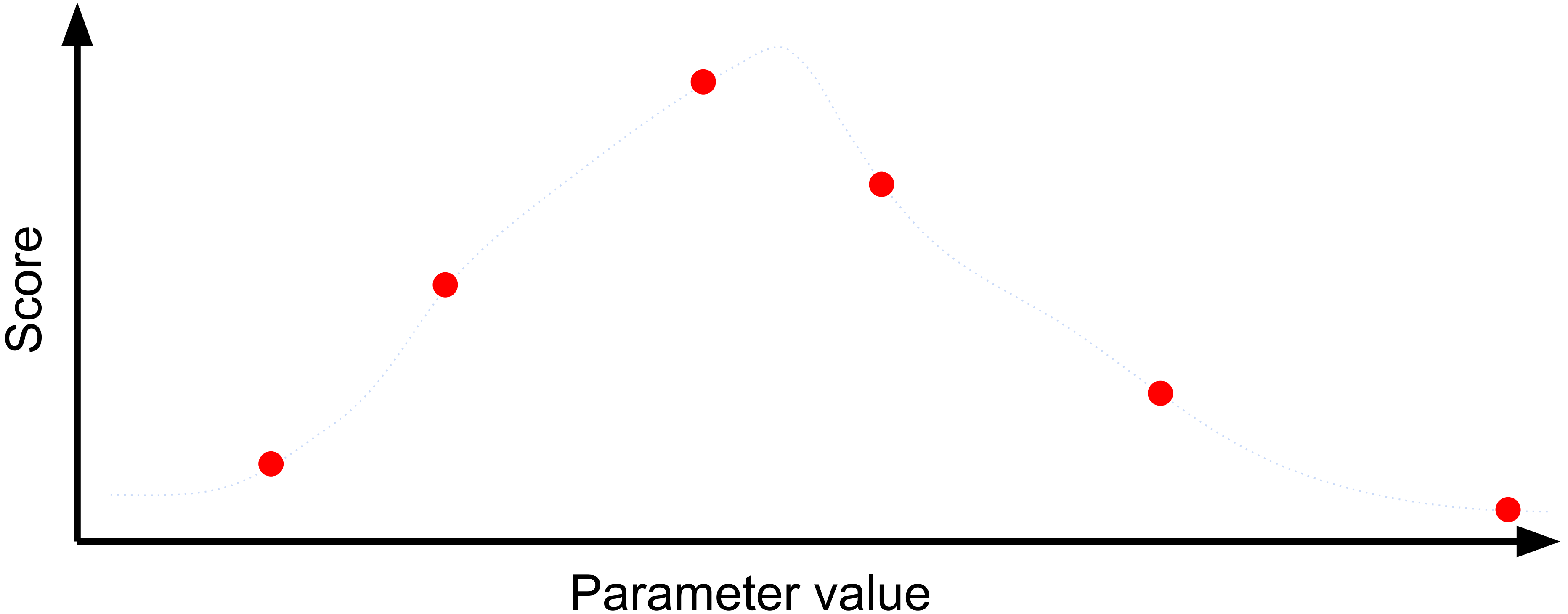
# Optimization 101

Line search



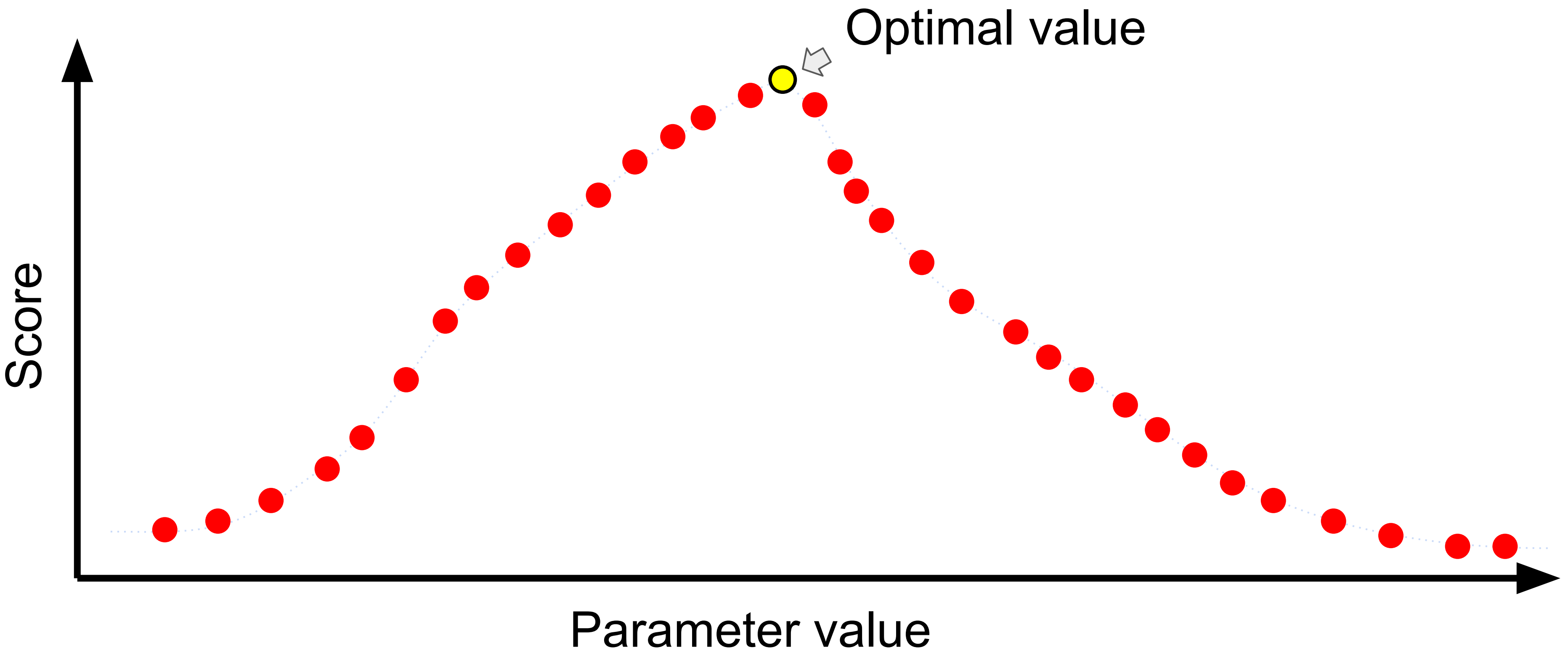
# Optimization 101

Line search



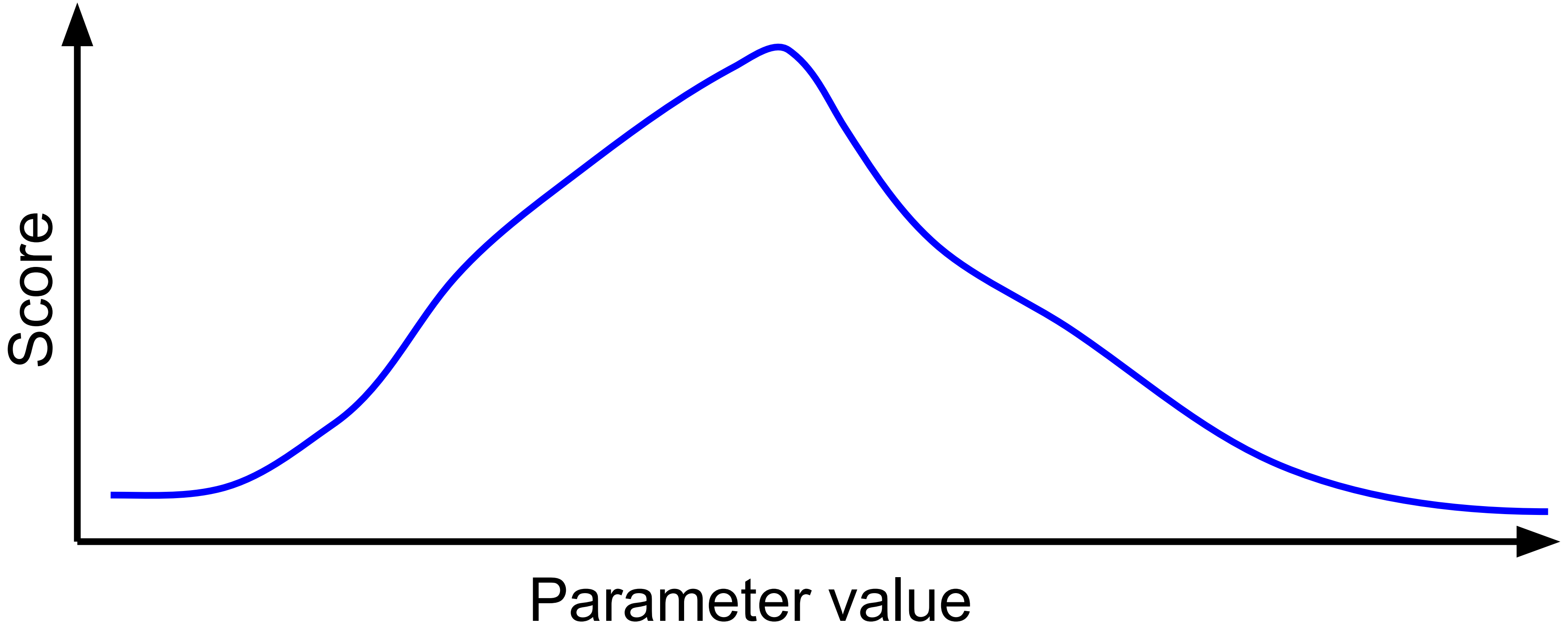
# Optimization 101

Line search via brute force enumeration



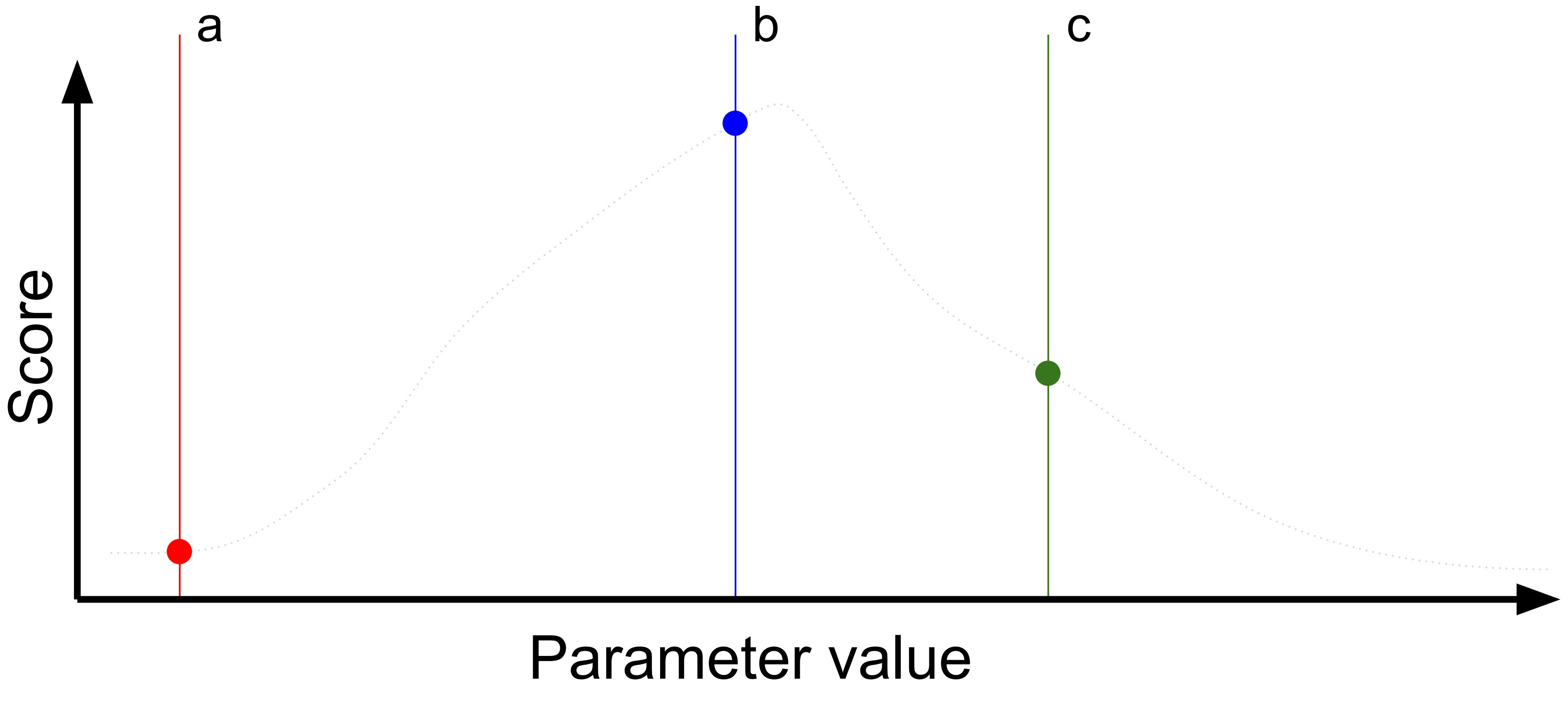
# Optimization 101

Line search



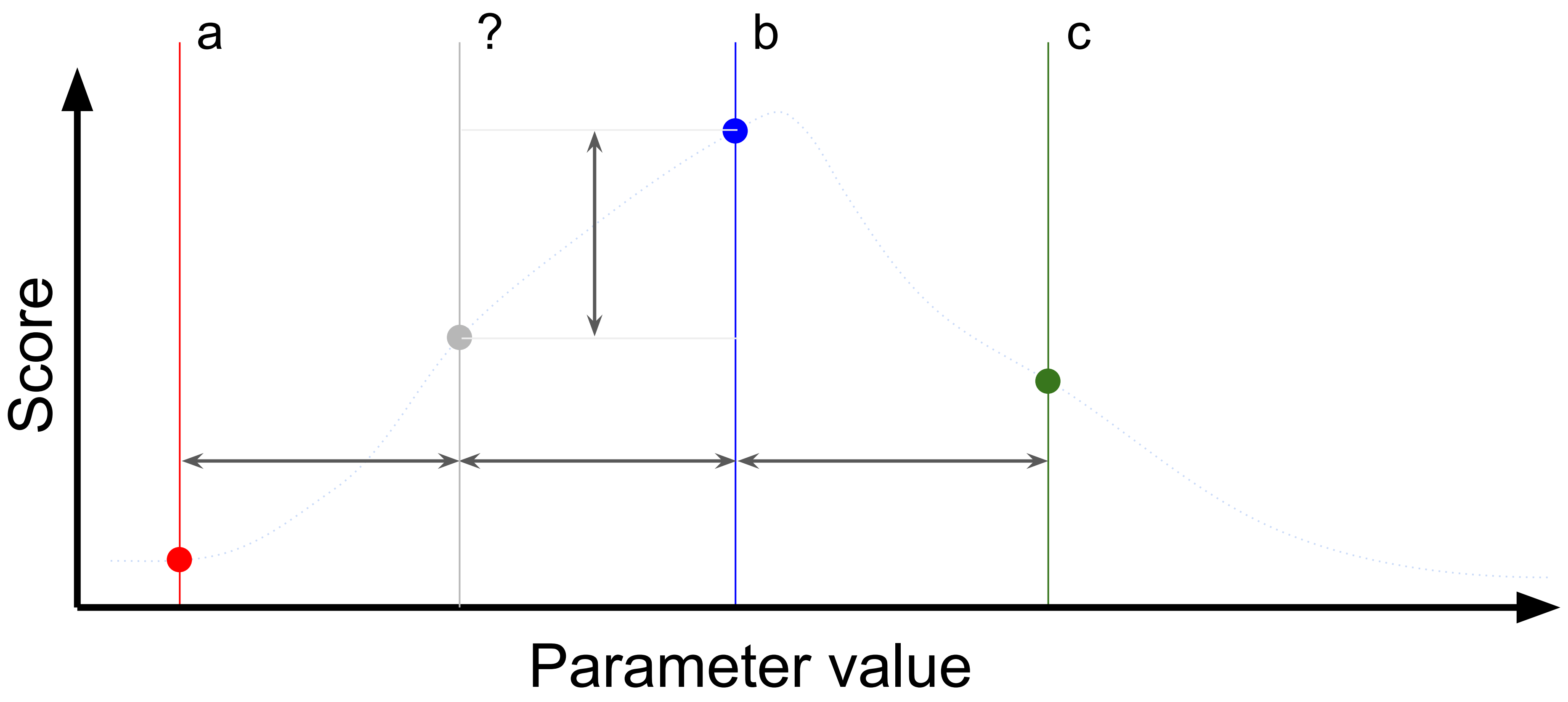
# Optimization 101

Line search using bisection



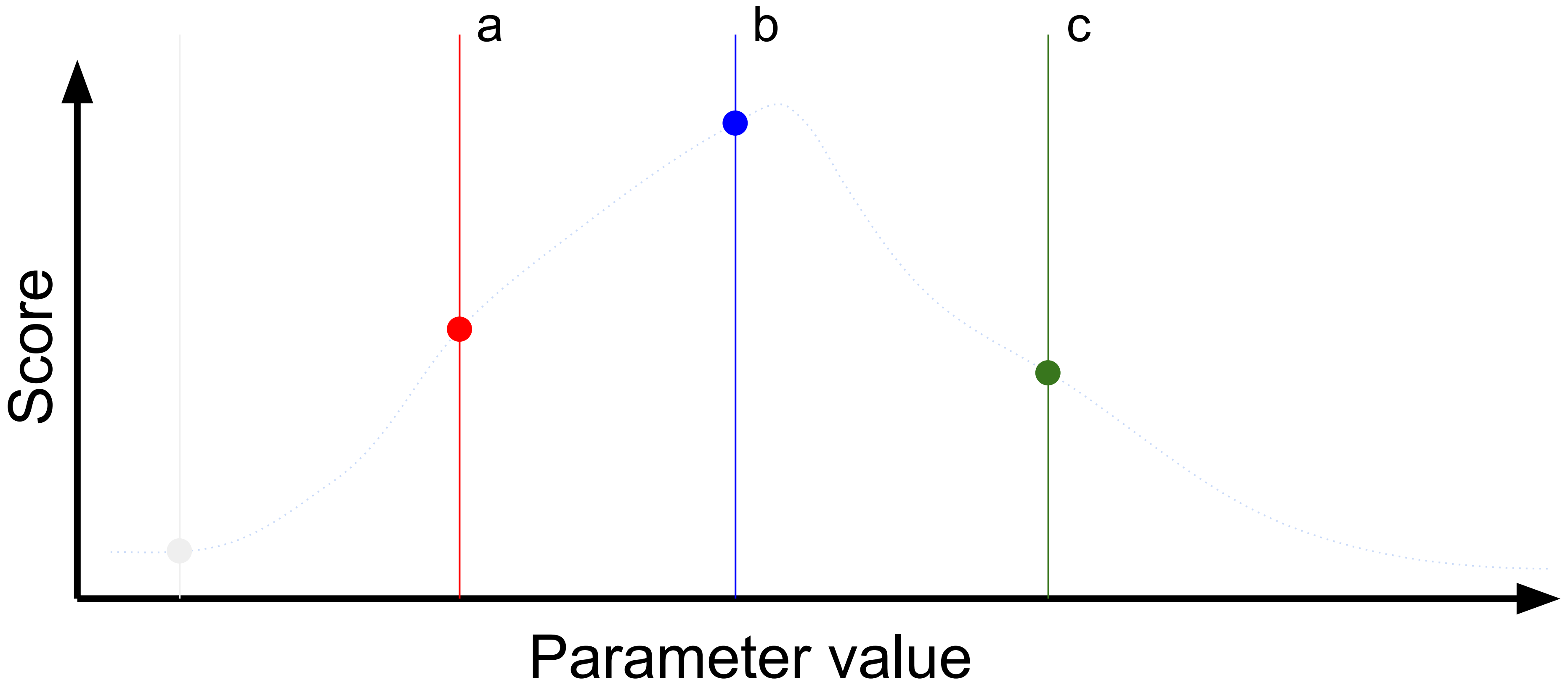
# Optimization 101

Line search using bisection



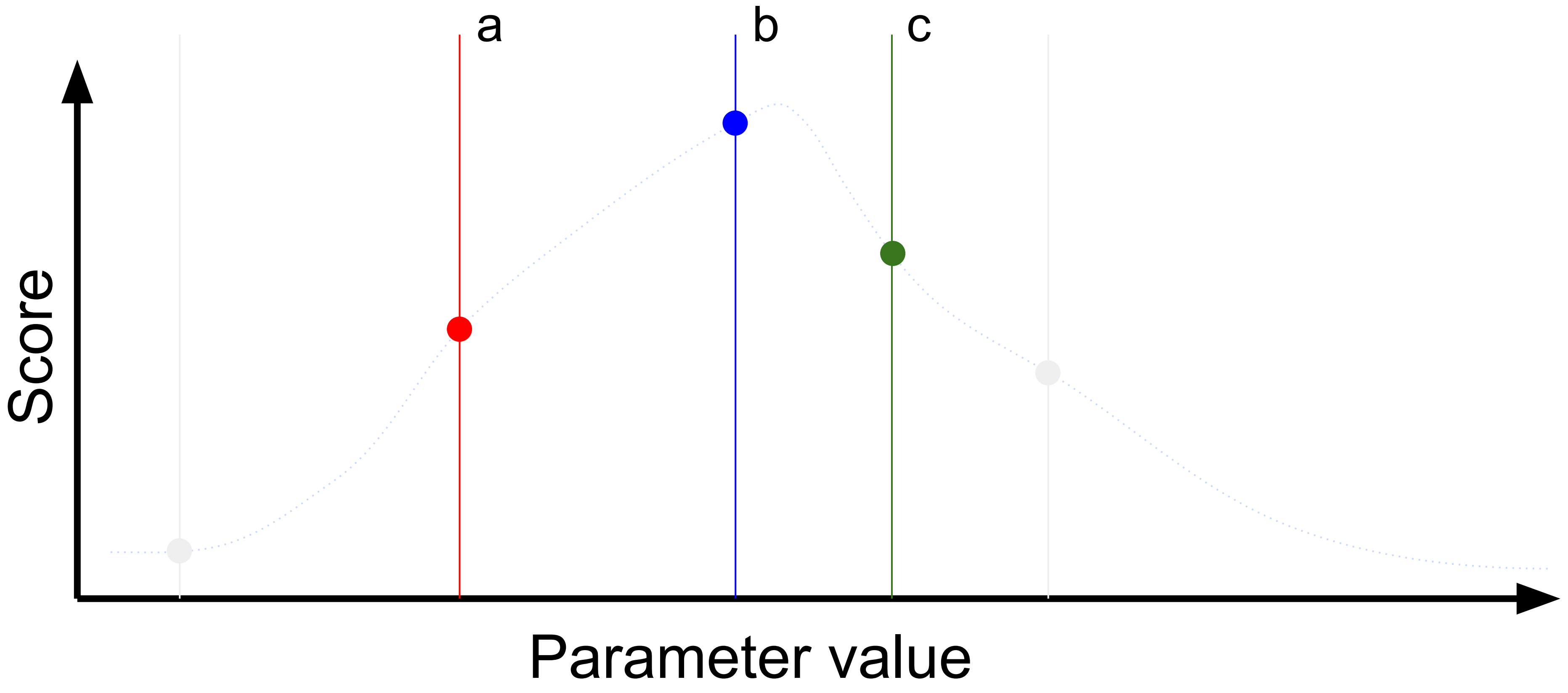
# Optimization 101

Line search using bisection



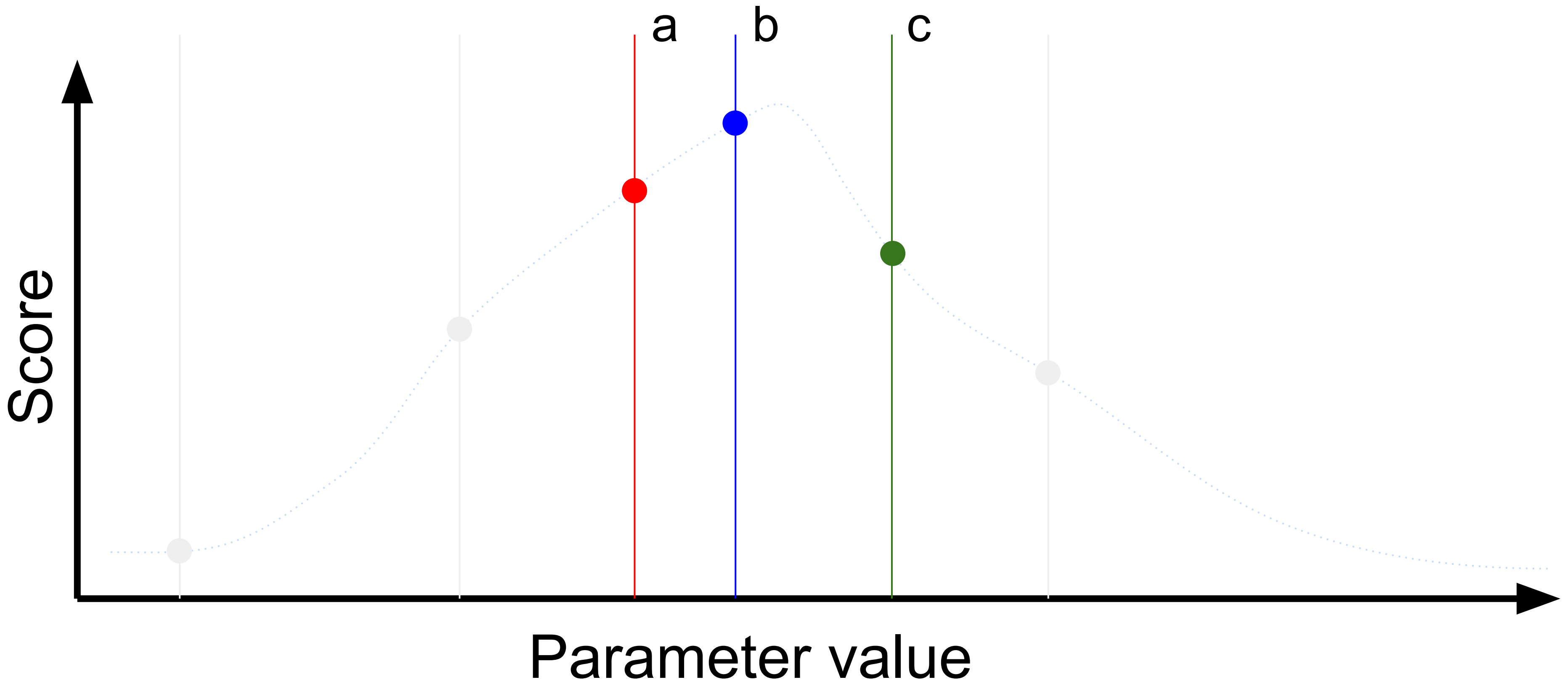
# Optimization 101

Line search using bisection



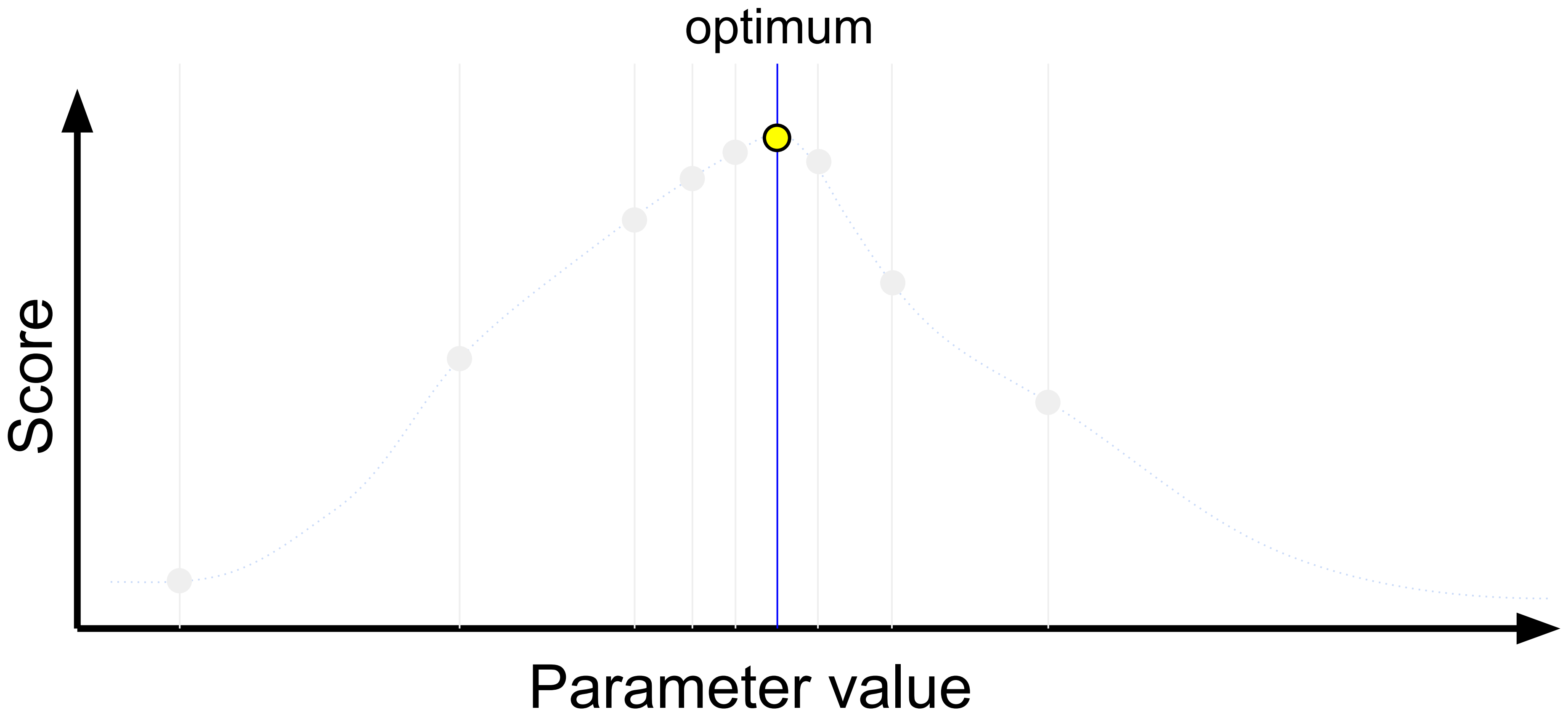
# Optimization 101

Line search using bisection



# Optimization 101

Line search using bisection



# Optimization 101

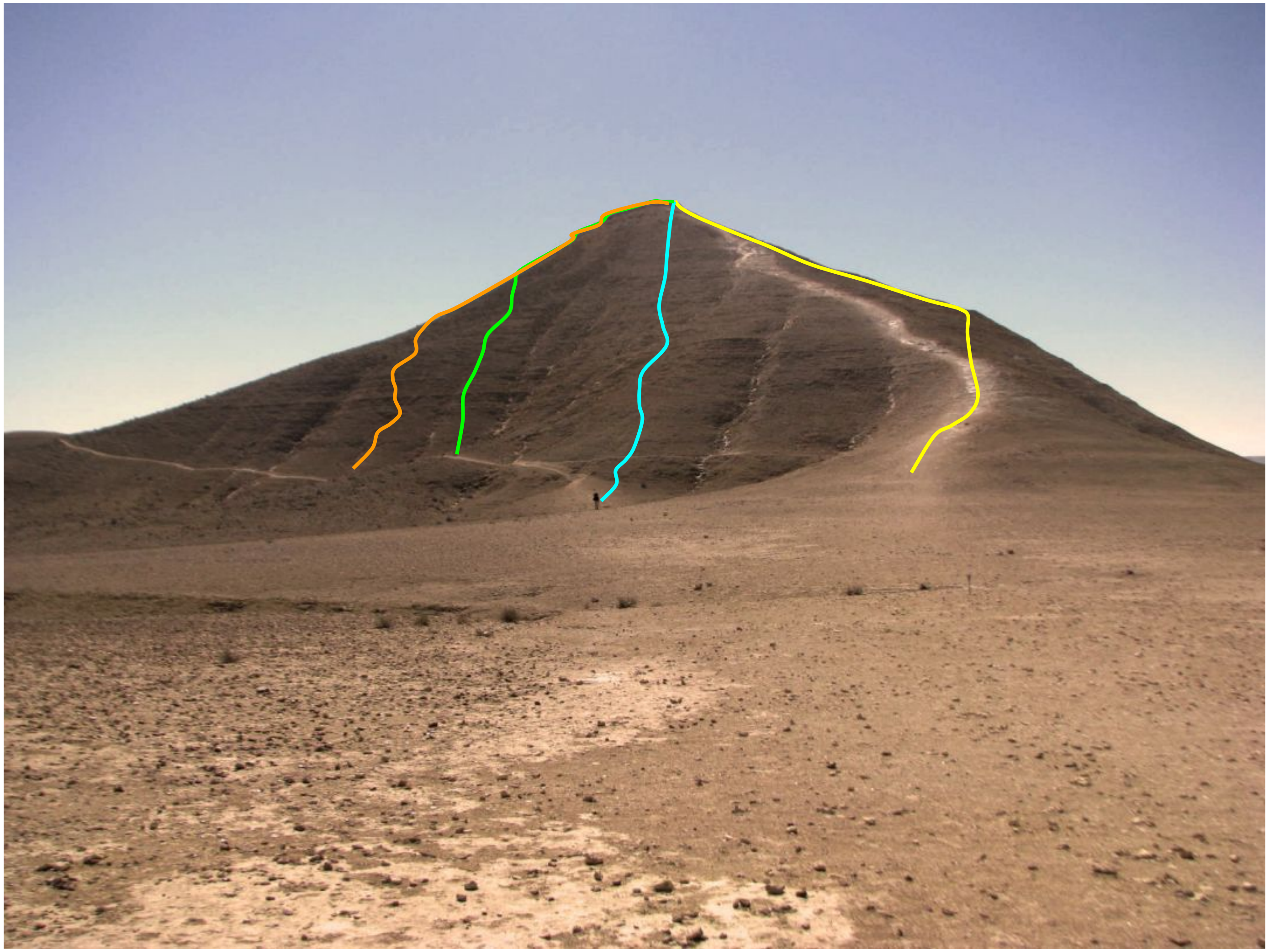
Higher order spaces

# Optimization 101

Steepest descent

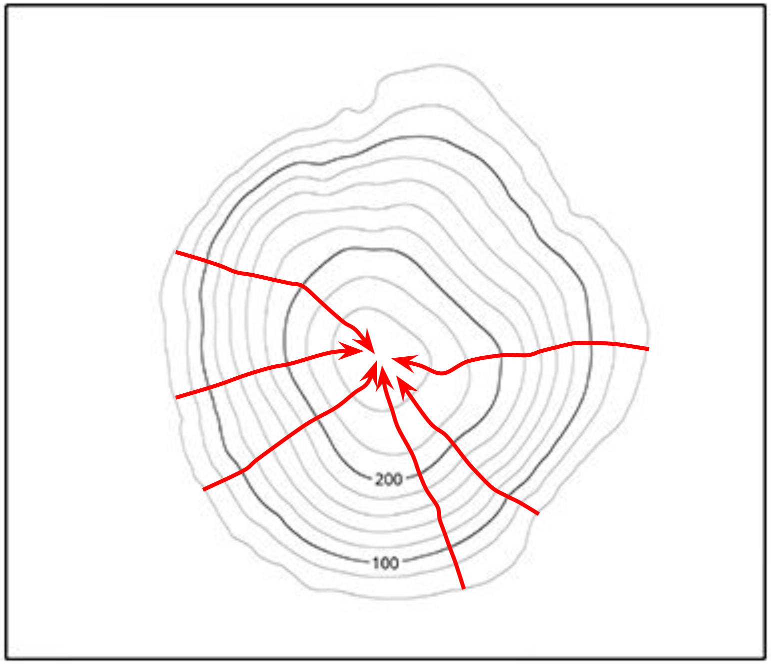
# Optimization 101

Steepest ascent



# Optimization 101

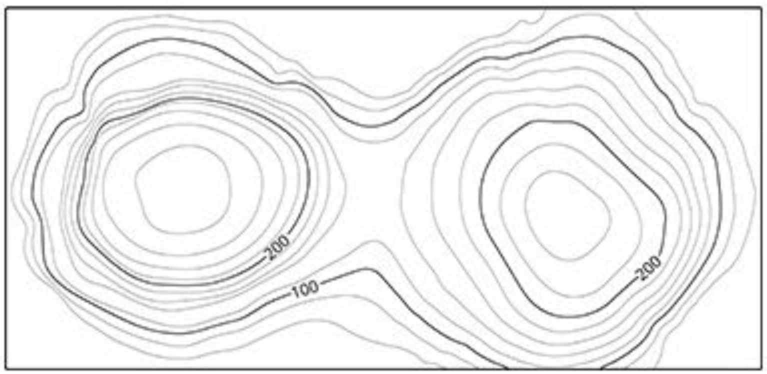
## Gradient ascent



$$\nabla(f) = \frac{\partial f}{\partial x} e_1 + \frac{\partial f}{\partial y} e_2$$

# Optimization 101

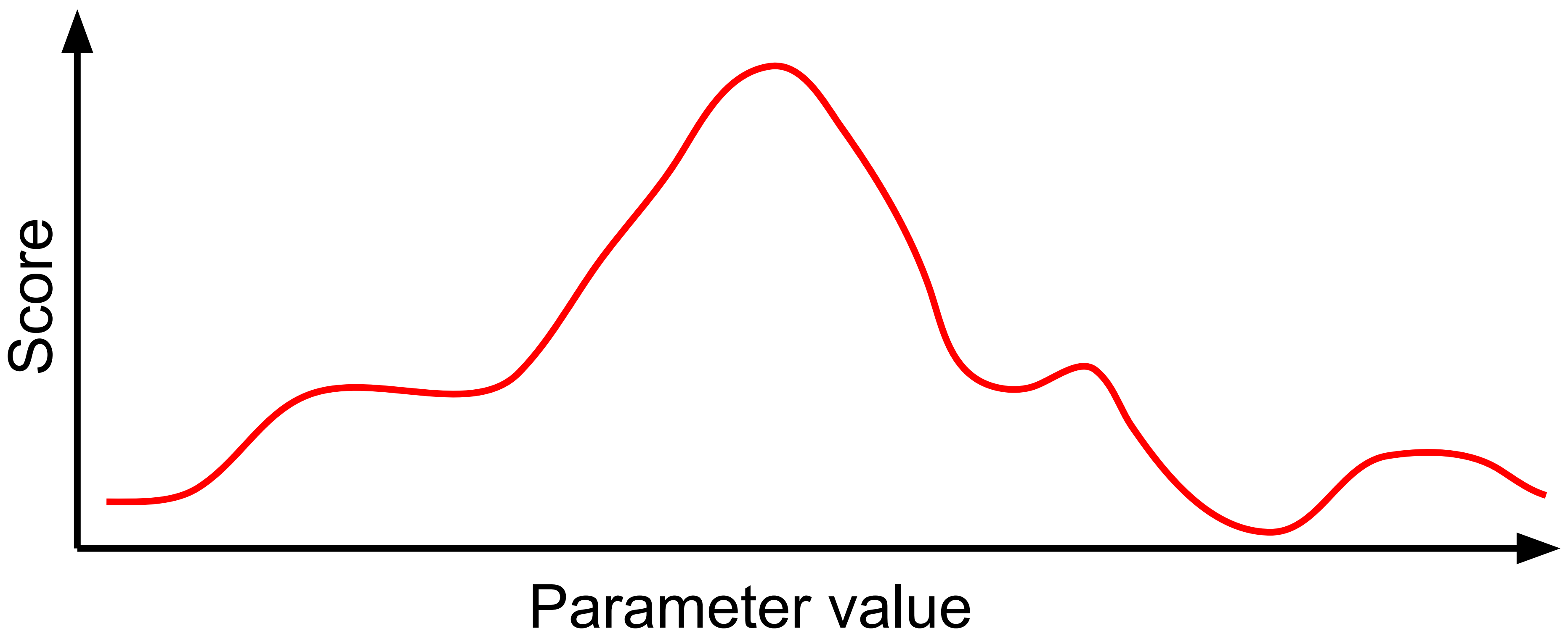
Gradient ascent



Search might fail if the search space isn’t convex

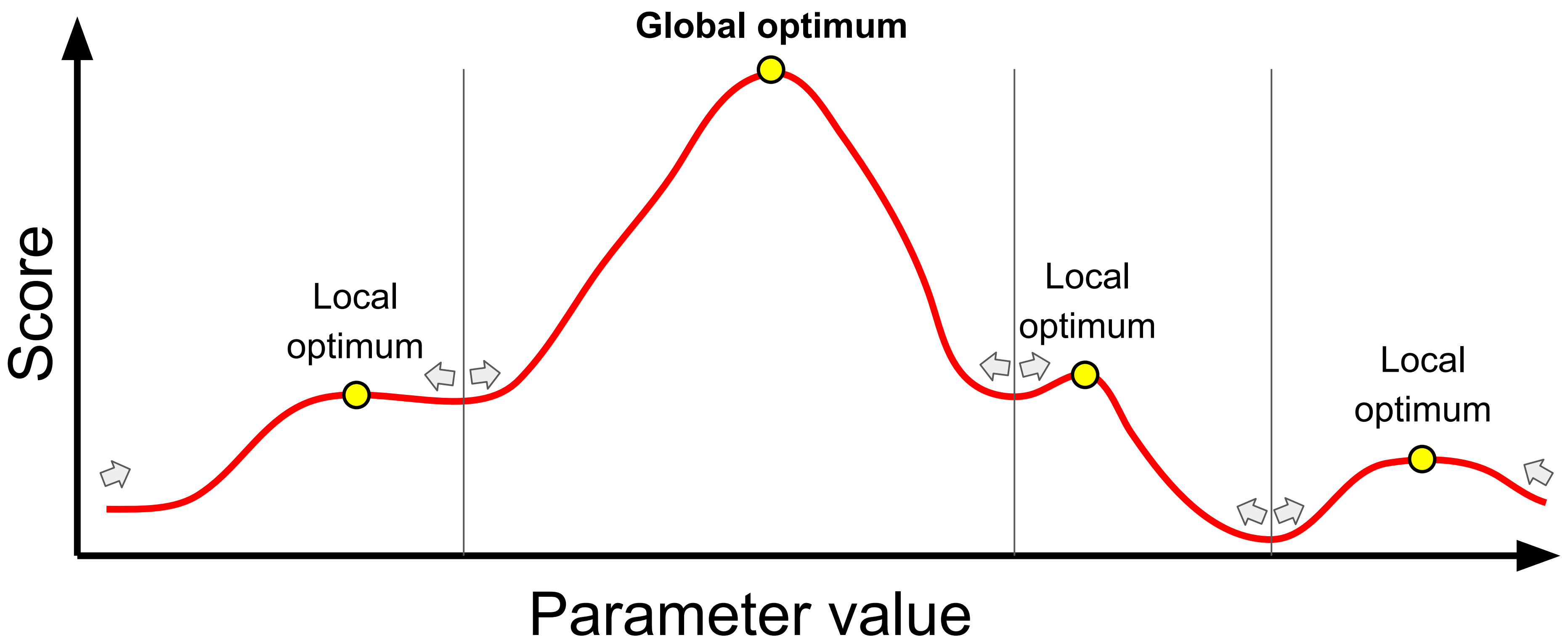
# Optimization 101

Search might fail if the search space isn't convex



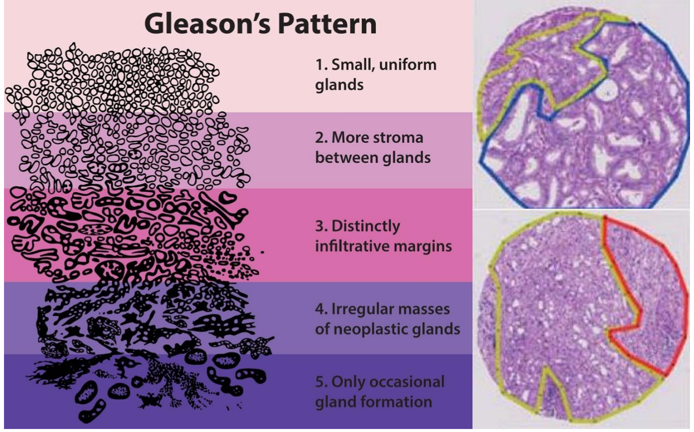
# Optimization 101

Search might fail if the search space isn't convex

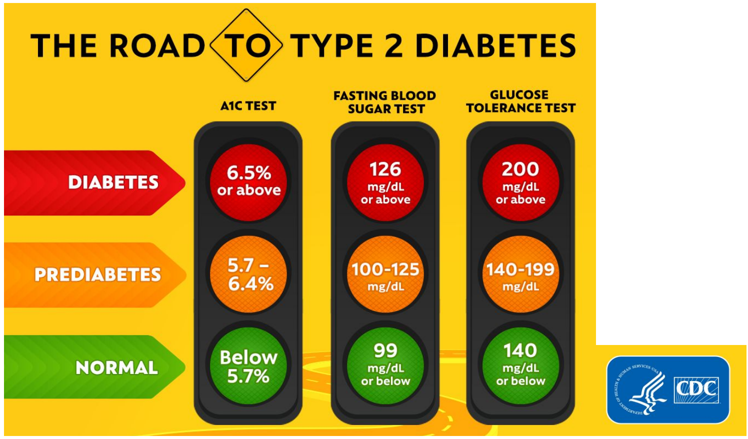


# Back to medicine - why should we care?

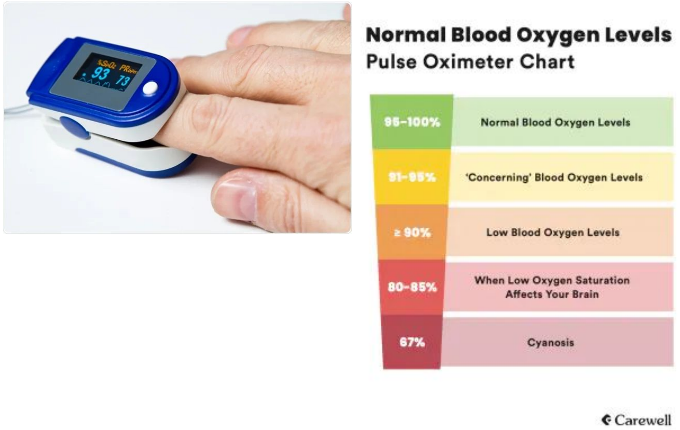
## Gleason grading system



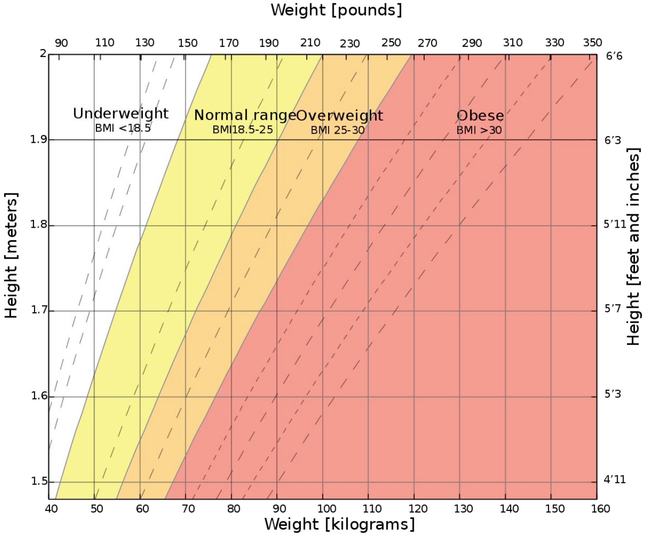

## Prediabetes diagnosis




## Blood Oxygen levels




## Body mass index



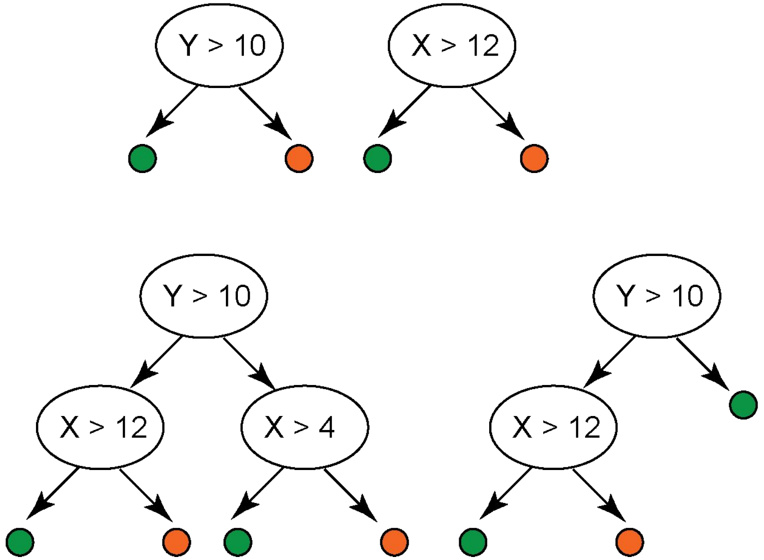

**Or cholesterol, or breast cancer, or resting heart rate, etc.**

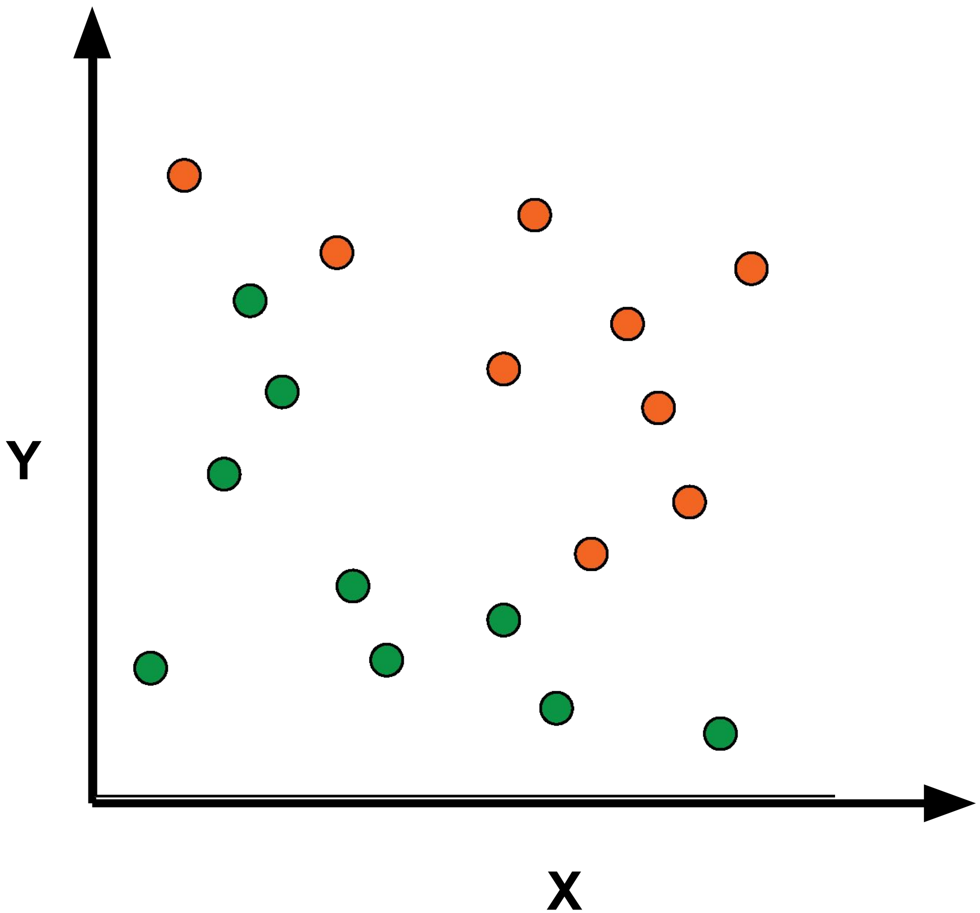
# Advanced subjects

- Optimal solution vs Heuristics
- Convexity

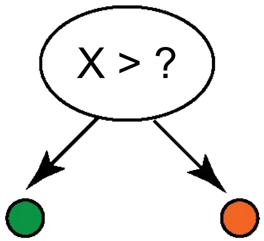
# Discrete Search

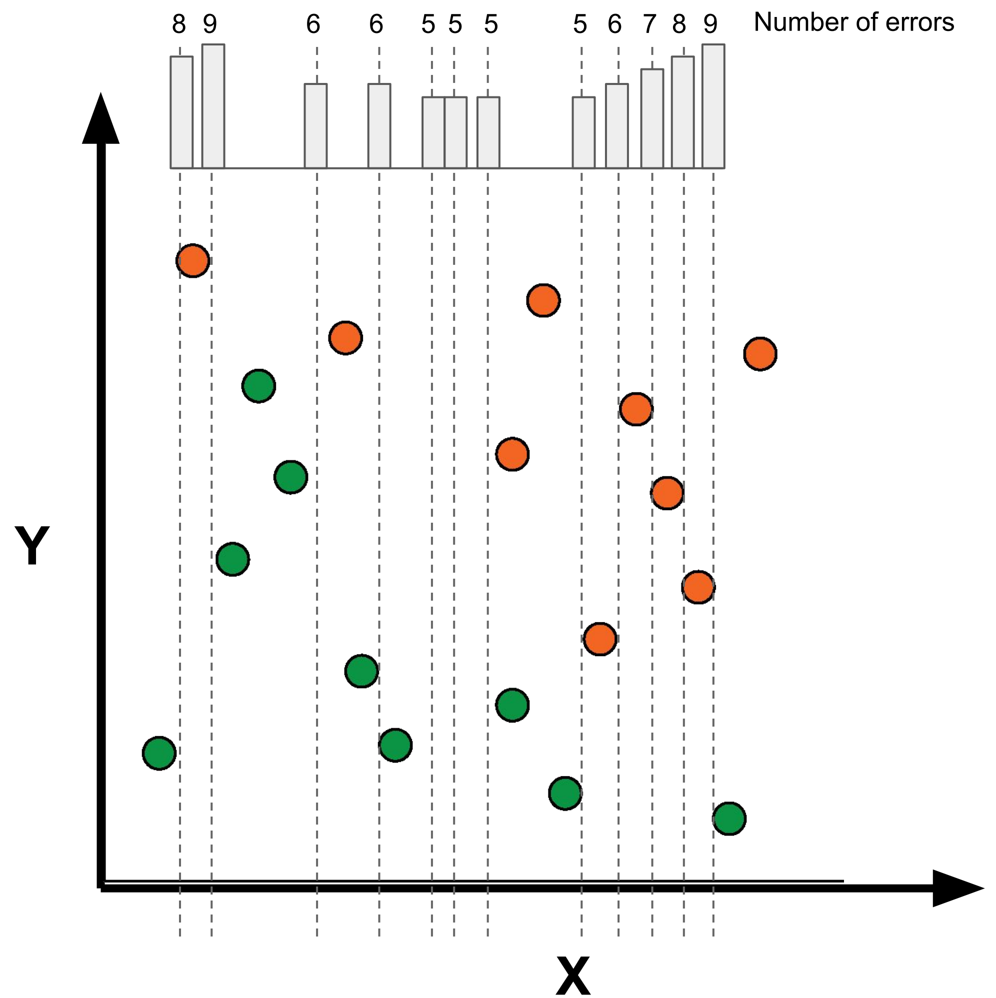
Learning decision trees?



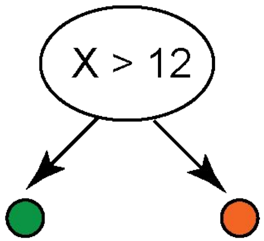



# Discrete Search

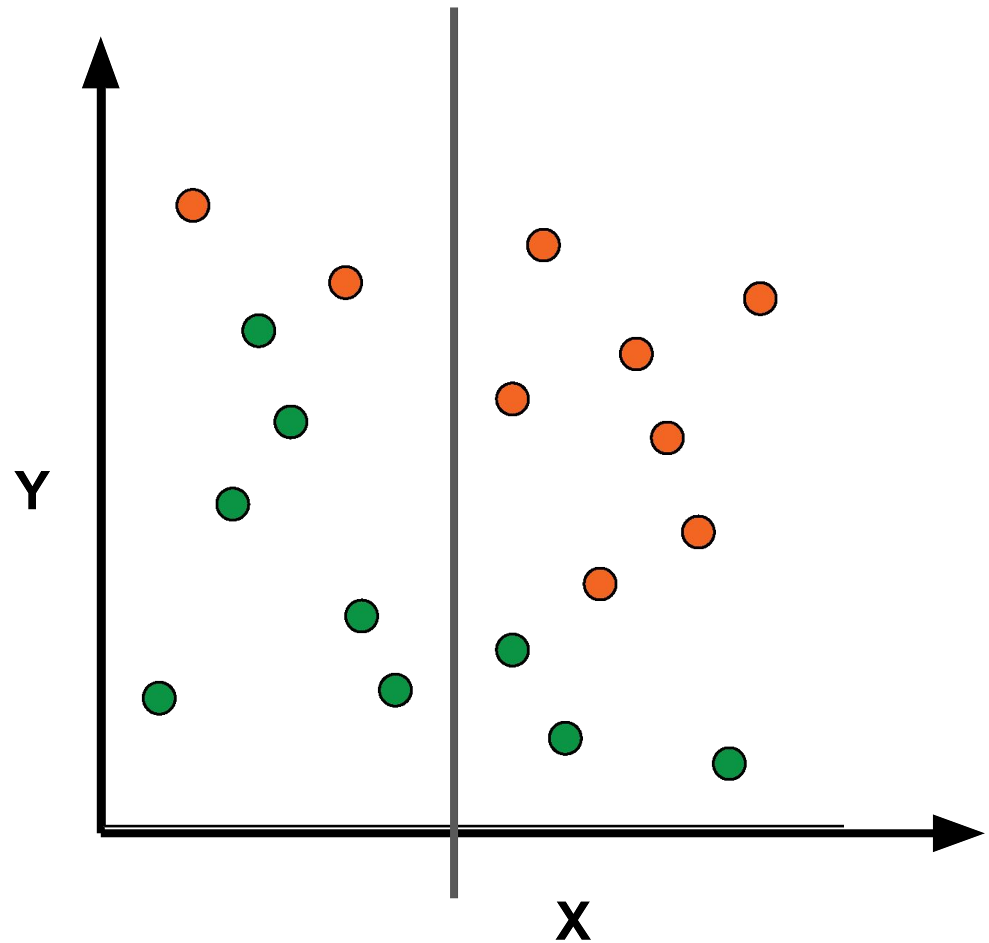


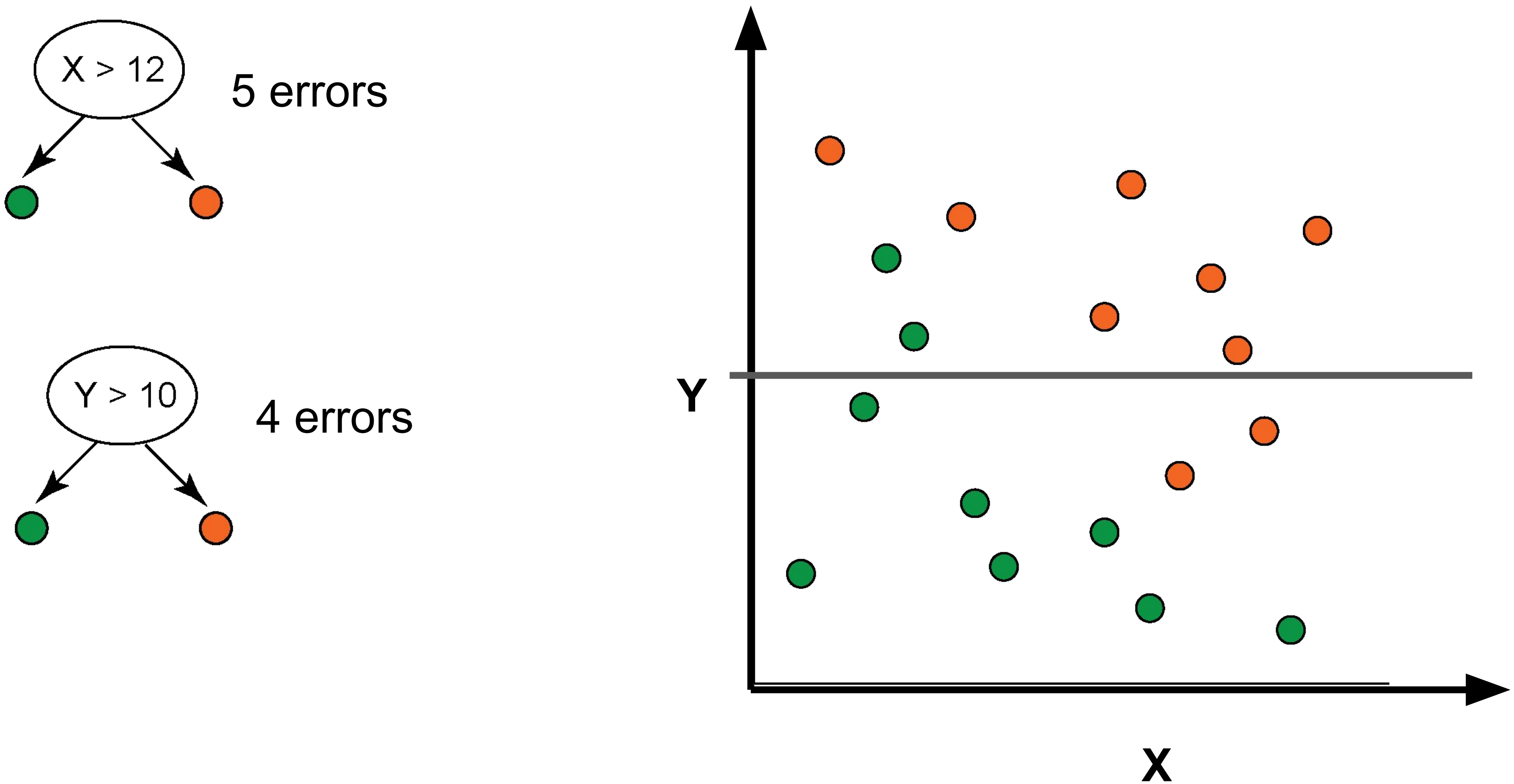
# Discrete Search



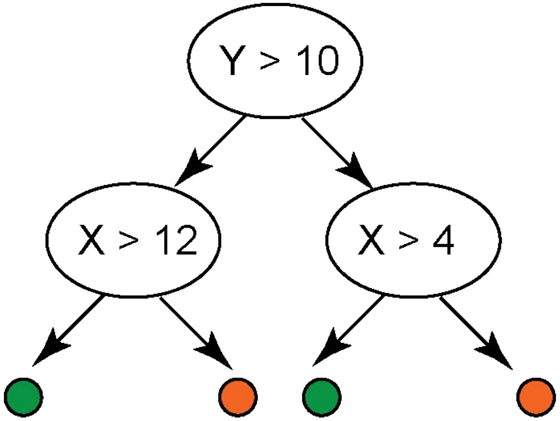

5 errors



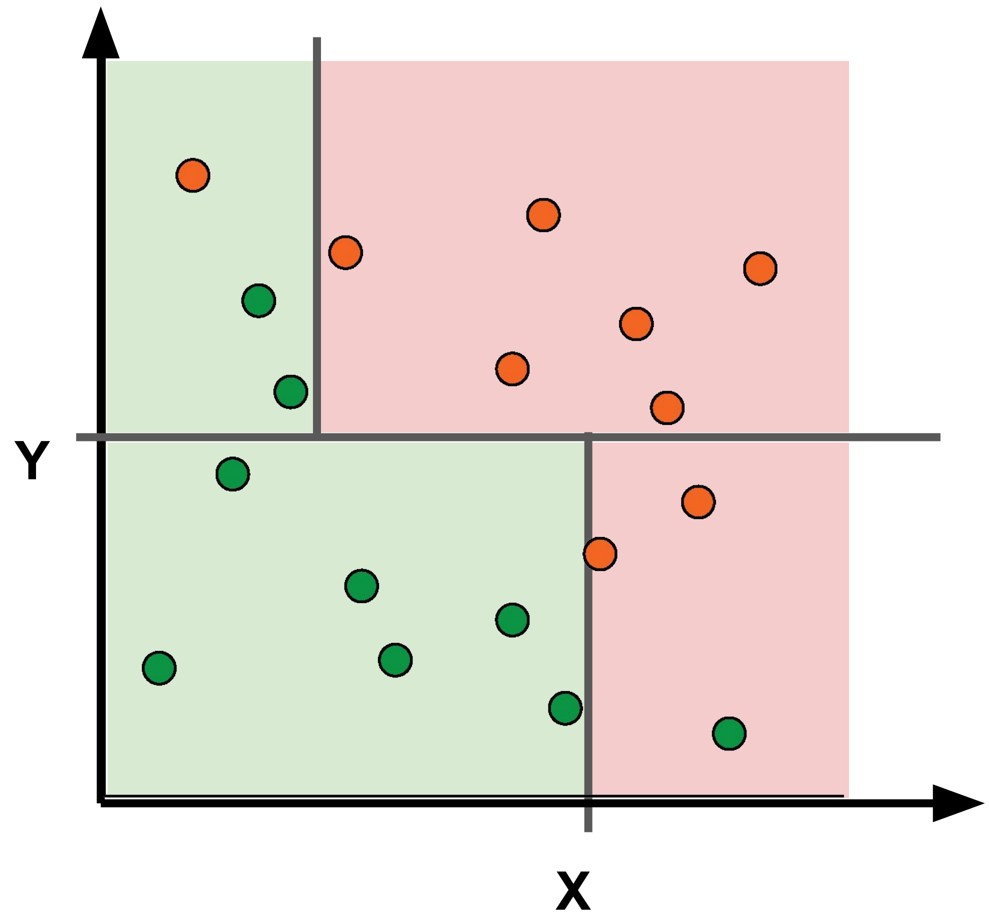
# Discrete Search



# Discrete Search




2 errors



# Greedy Algorithms

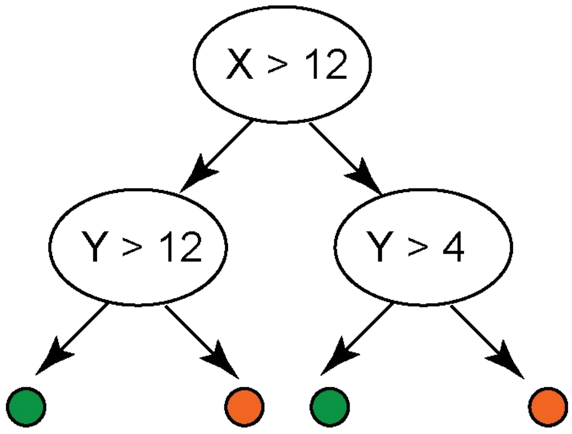
"An algorithm which always takes the best immediate, or local, solution while finding an answer"

(from http://www.nist.gov/dads/HTML/greedyalgo.html)

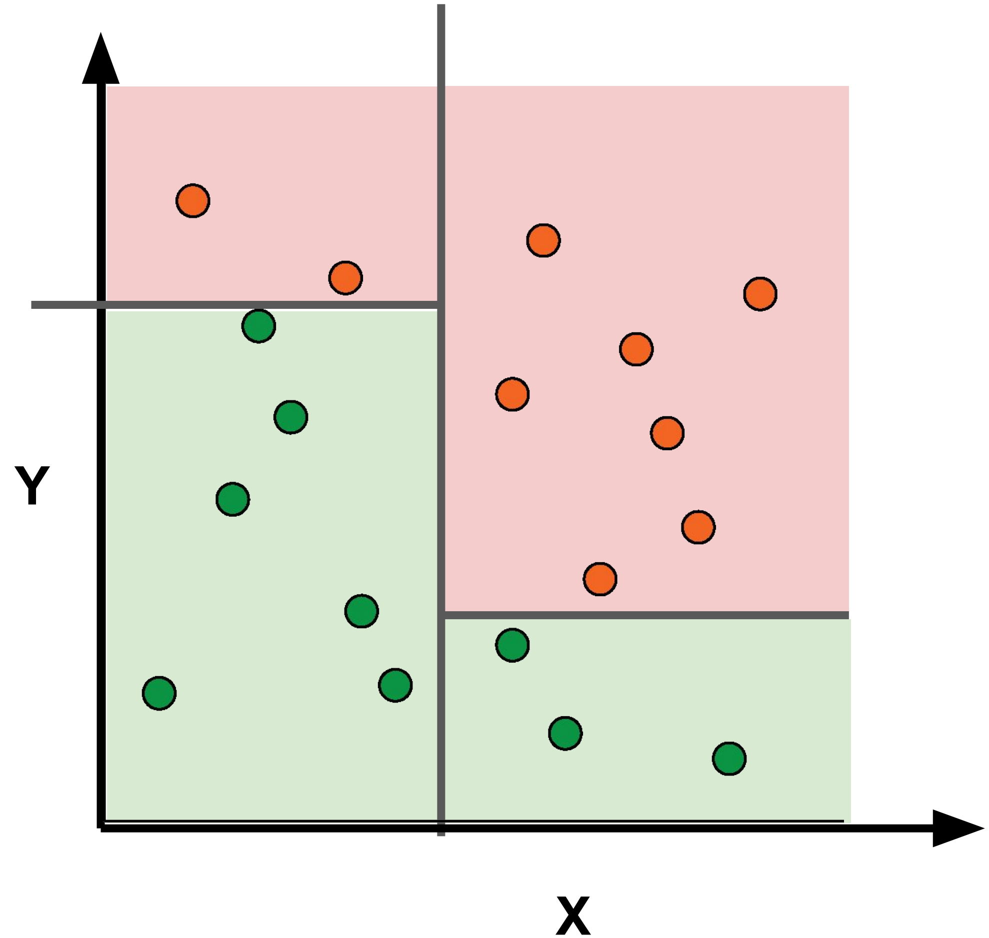
+ Simple, local decisions

- May not reach global optimum

# Discrete Search




0 errors



# Summary

- Learning (rules) from data
- Parameter space and hypothesis space
- Loss function (how much errors cost)
- Finding the parameters of the optimal classifier
- Line search optimization
- Gradient descent (and convex optimization)
- Greedy algorithms and heuristics

The basics of machine learning

# Syllabus

1. Introduction
2. Classification
3. Learning 1
4. AI in ophthalmology (Prof. Itay Chowers)
5. Learning 2
6. Regression
7. Clustering
8. Visualization (and dimensionality reduction)
9. Deep learning in image analysis (Prof. Leo Joskowicz)
10. Missing data, statistical dependencies
11. Natural language in medicine (Dr. Gabi Stanovsky)
12. Decisions (utility)
13. Longitudinal Data / Project