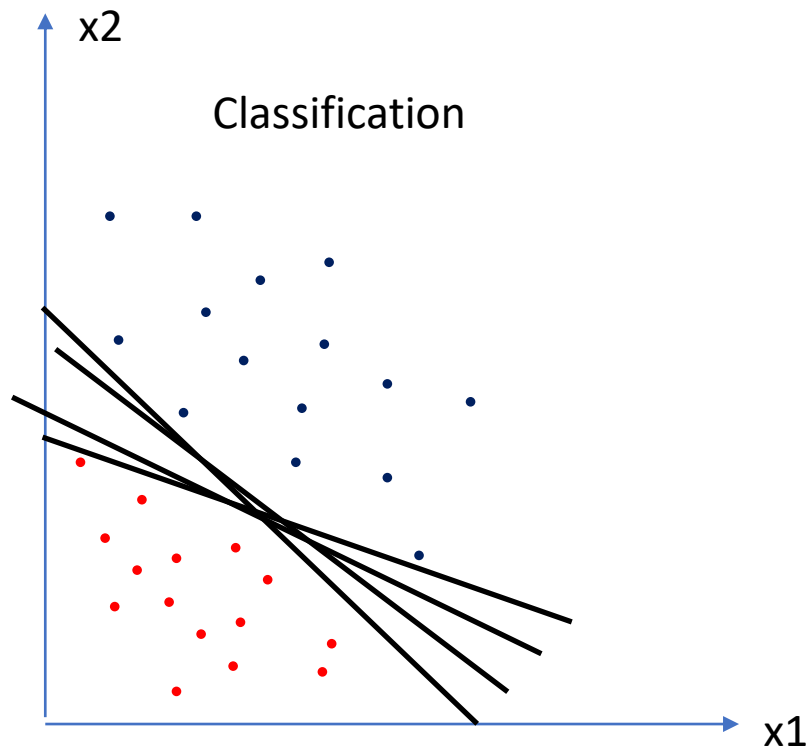


# Support Vector Machines 101

# Support Vector Machines

## Introduction

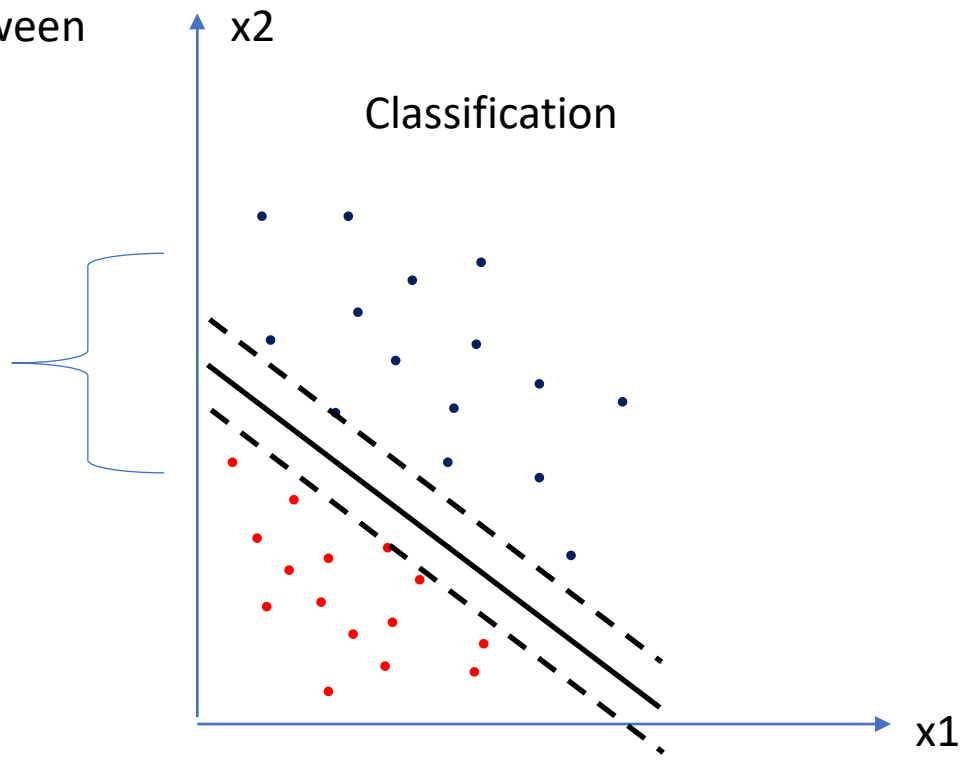
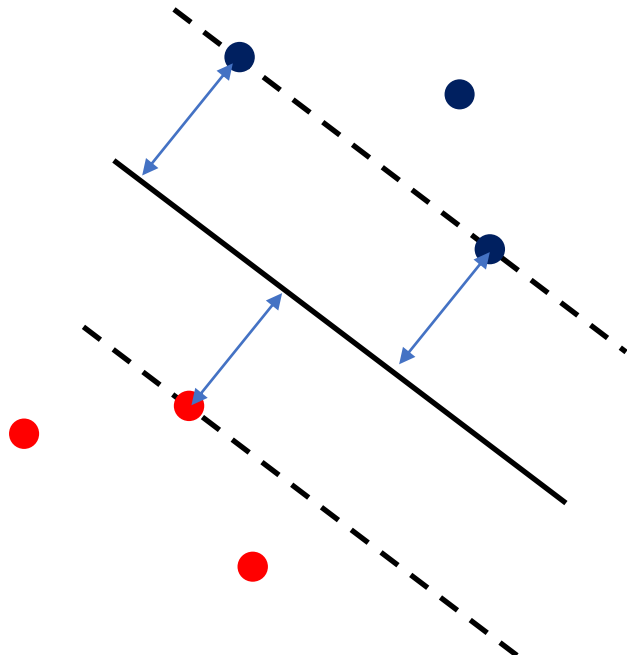
- Applicable for classification and regression tasks
- Idea: separate classes with a linear hyperplane
- Here: two-dimensional, but any-dimension possible
- Problem: how to decide on a specific hyperplane
- Which hyperplane is the best?



# Support Vector Machines

## Maximum Margin Classifier

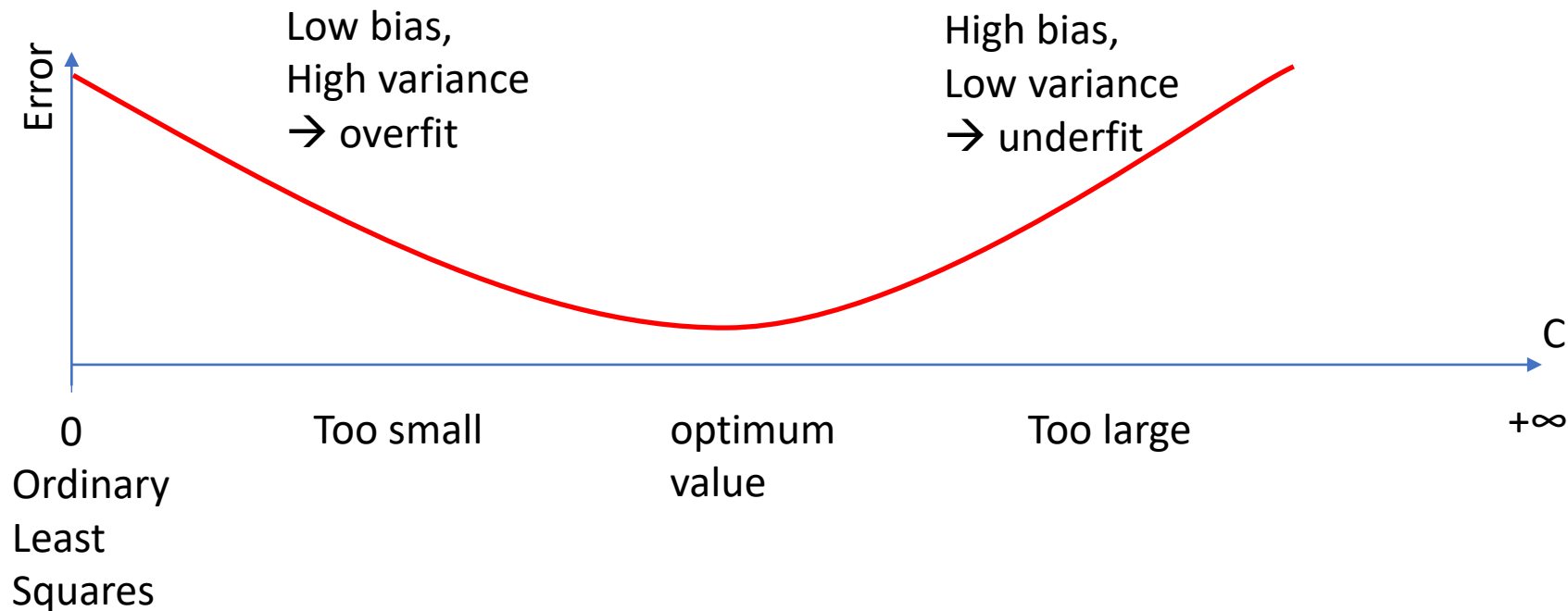
- Best hyperplane maximizes distance between hyperplane and close points
- Maximum margin classifier



# Support Vector Machines

## Soft Margin Classifier

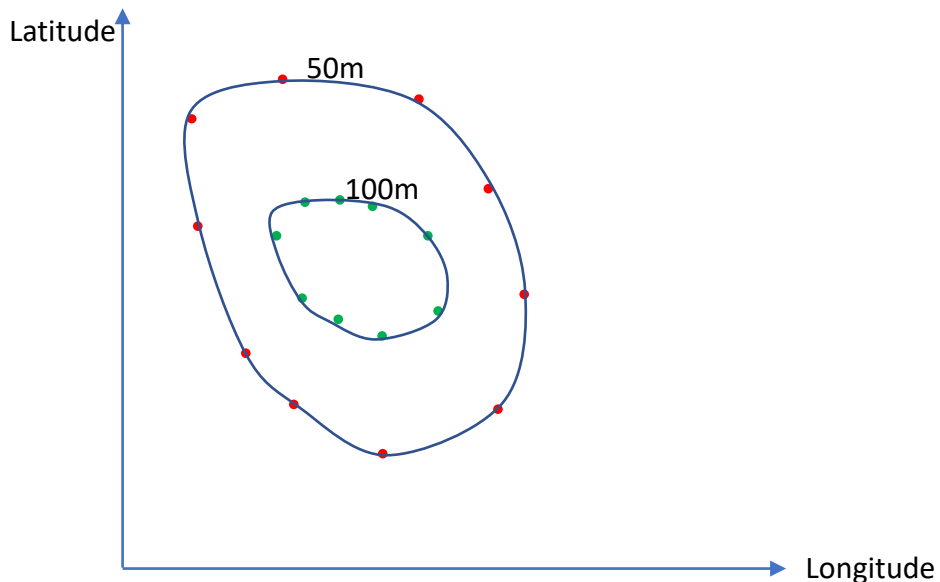
- Complete separation often impossible
- Soft margin classifier allows misclassification



# Support Vector Machines

## Non-Linearity

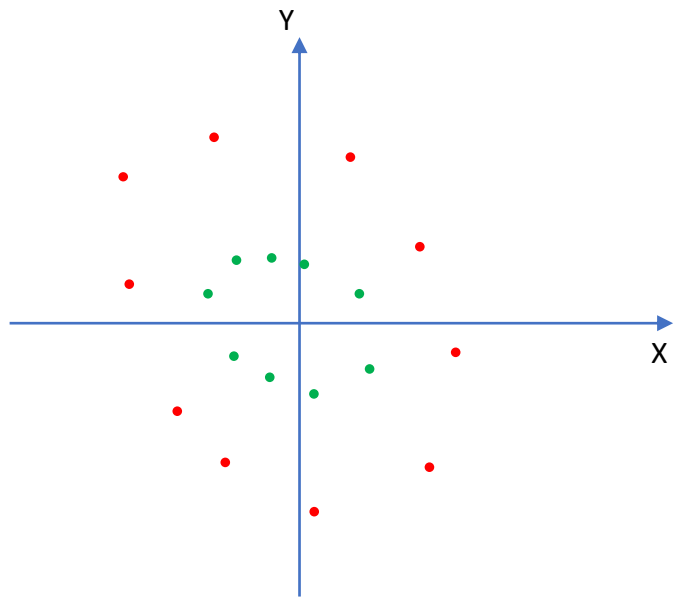
- Some data is not linearly separable  $\rightarrow$  SVMs fail
- Solution: Kernel trick
- Different kernels: polynomial, radial basis, linear.



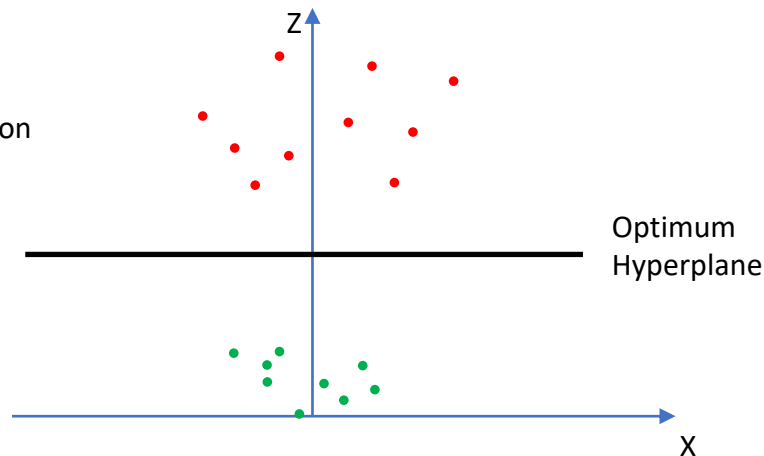
# Support Vector Machines

Higher Order Transformation

Non-linear separable data in lower space ...



Transformation  
 $Z = X^2 + Y^2$



...becomes separable in a higher dimension

# Support Vector Machines

## Hyperparameters

### Kernel

- Defines kernel: linear, polynomial, sigmoid, radial basis

### Cost

- How much misclassification is allowed?

### Gamma

- Required for kernel-functions

# Support Vector Machines

Advantages / Disadvantages



- good if groups are clearly separable
- Works well if number of variables is greater than number of observations
- good if separable non-linearly
- good model results



- Black-box
- High computational effort
- not suitable for large noise
- Very sensitive to kernel parameters