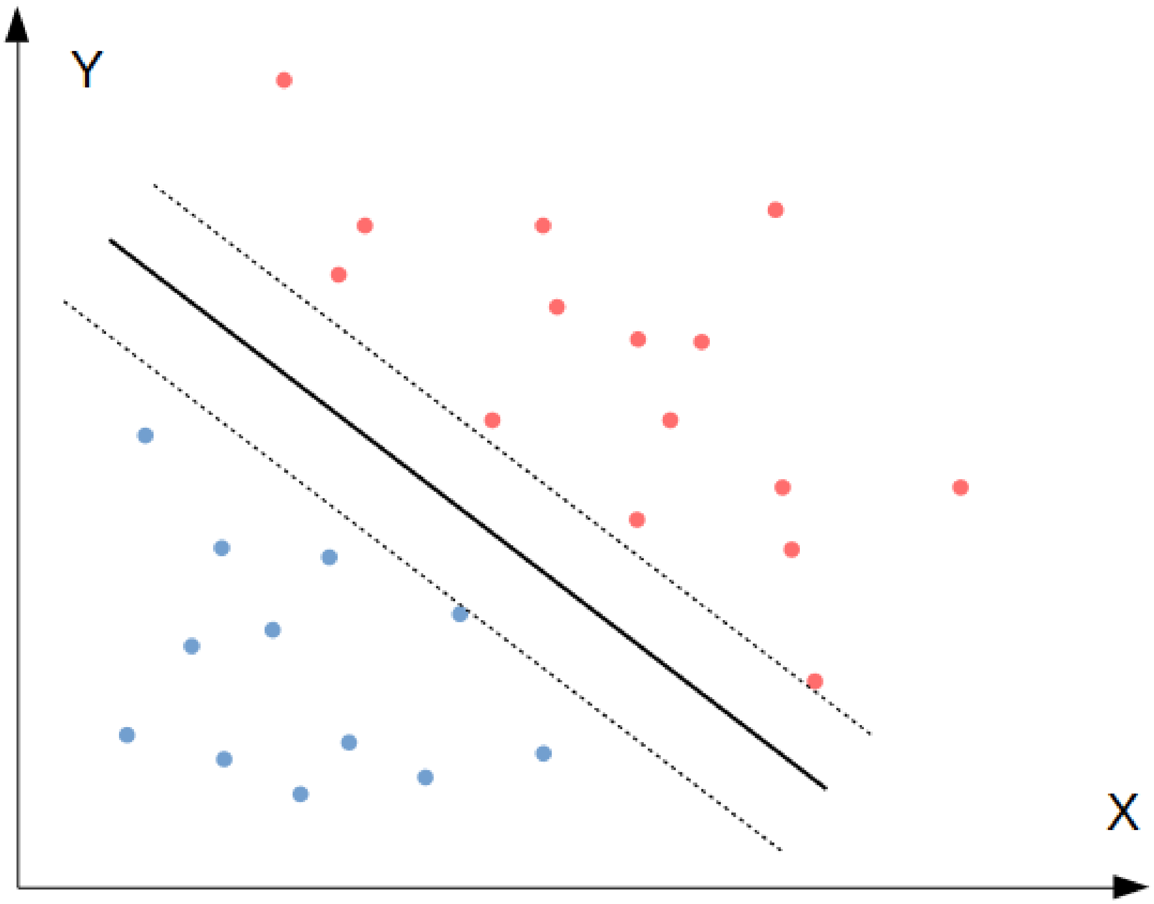
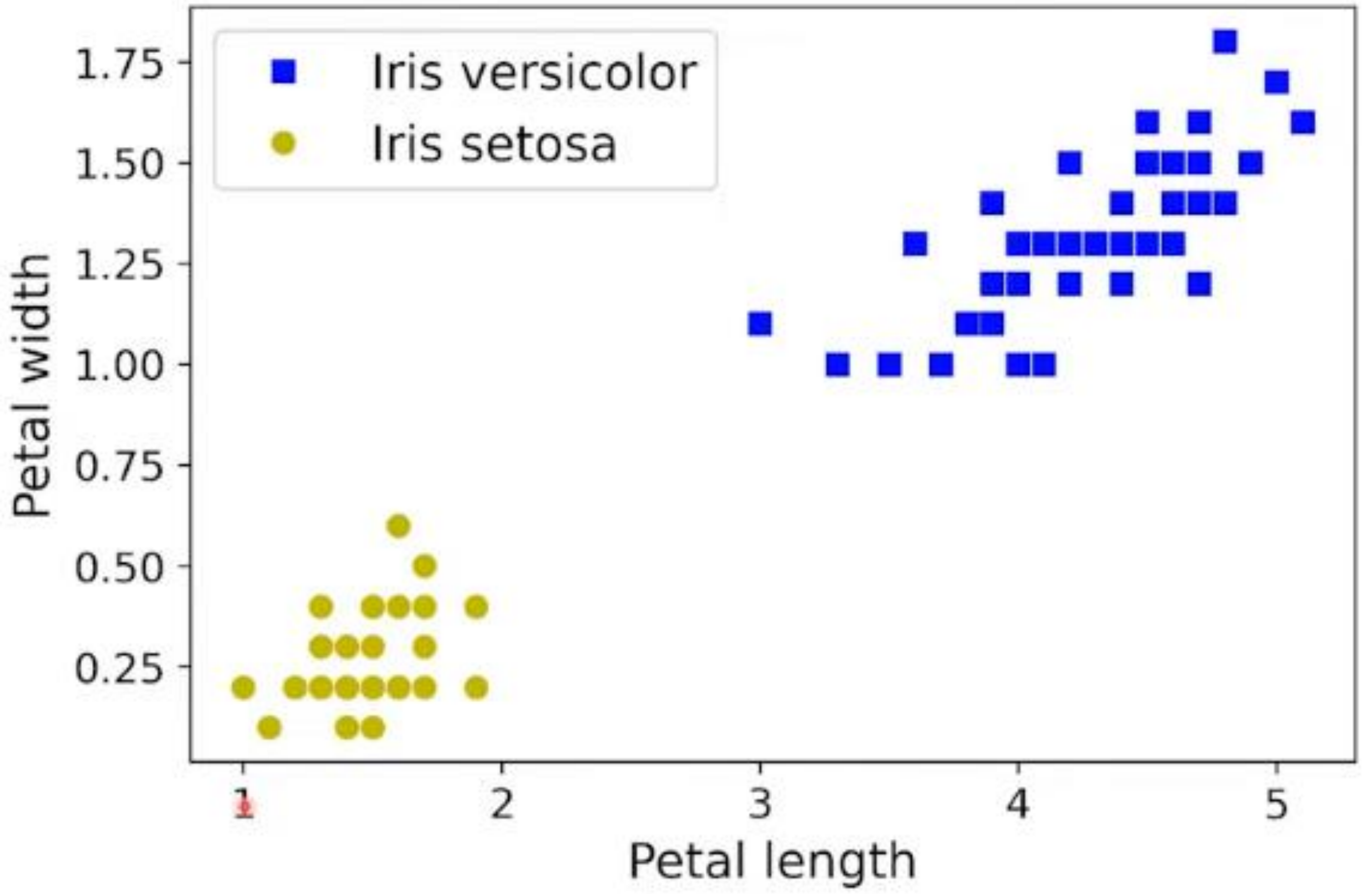
Support Vector Machine (SVM)

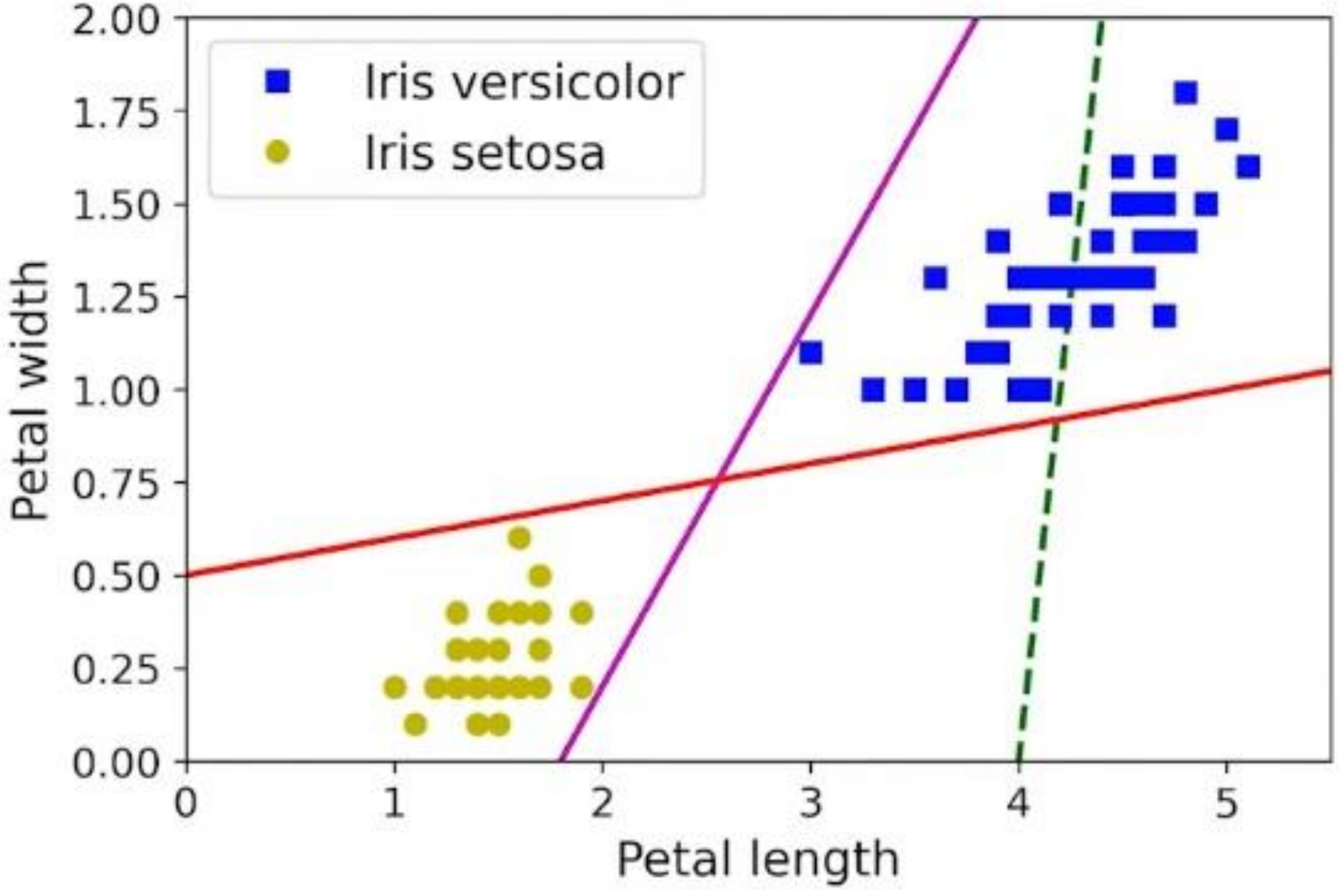
Support Vector Machine (SVM)



Linearly Separable

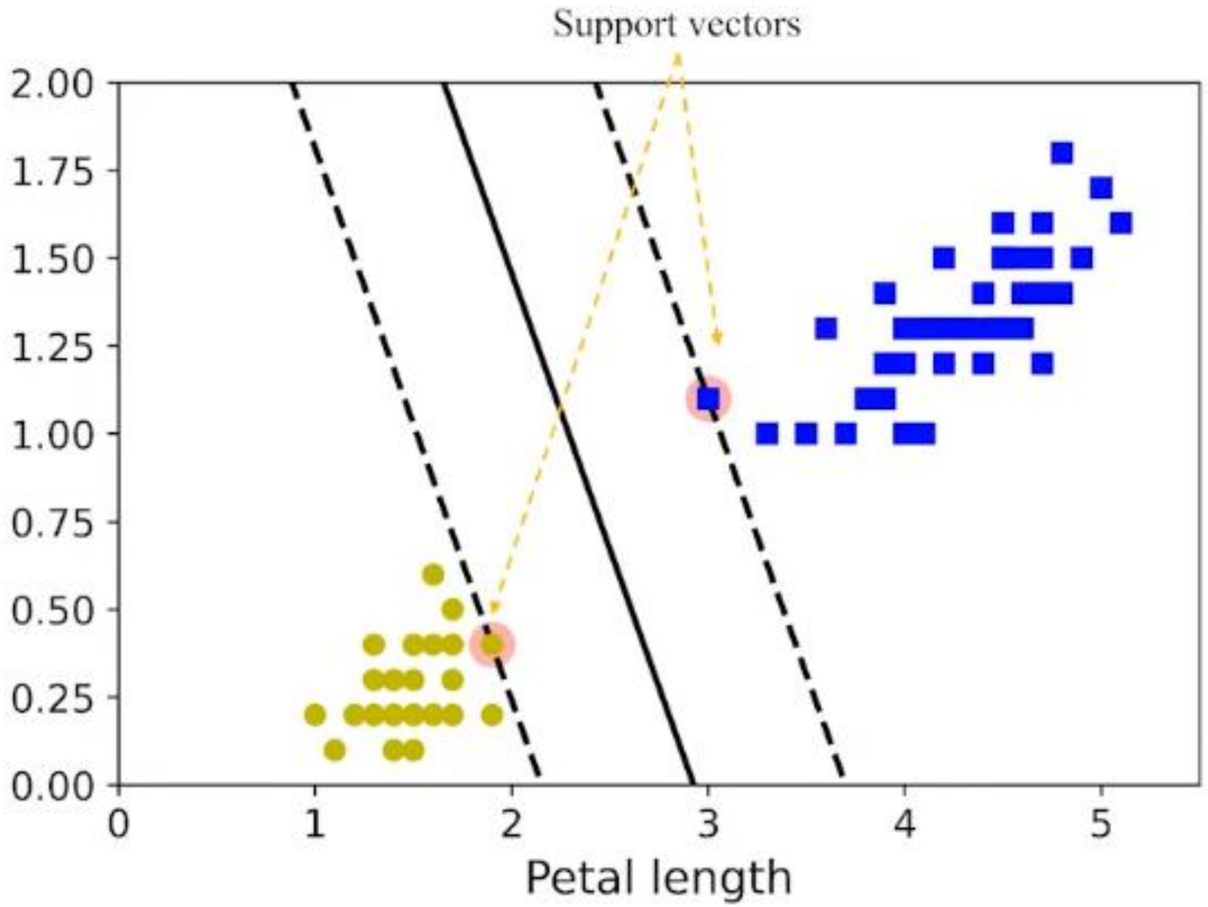


Which Line?



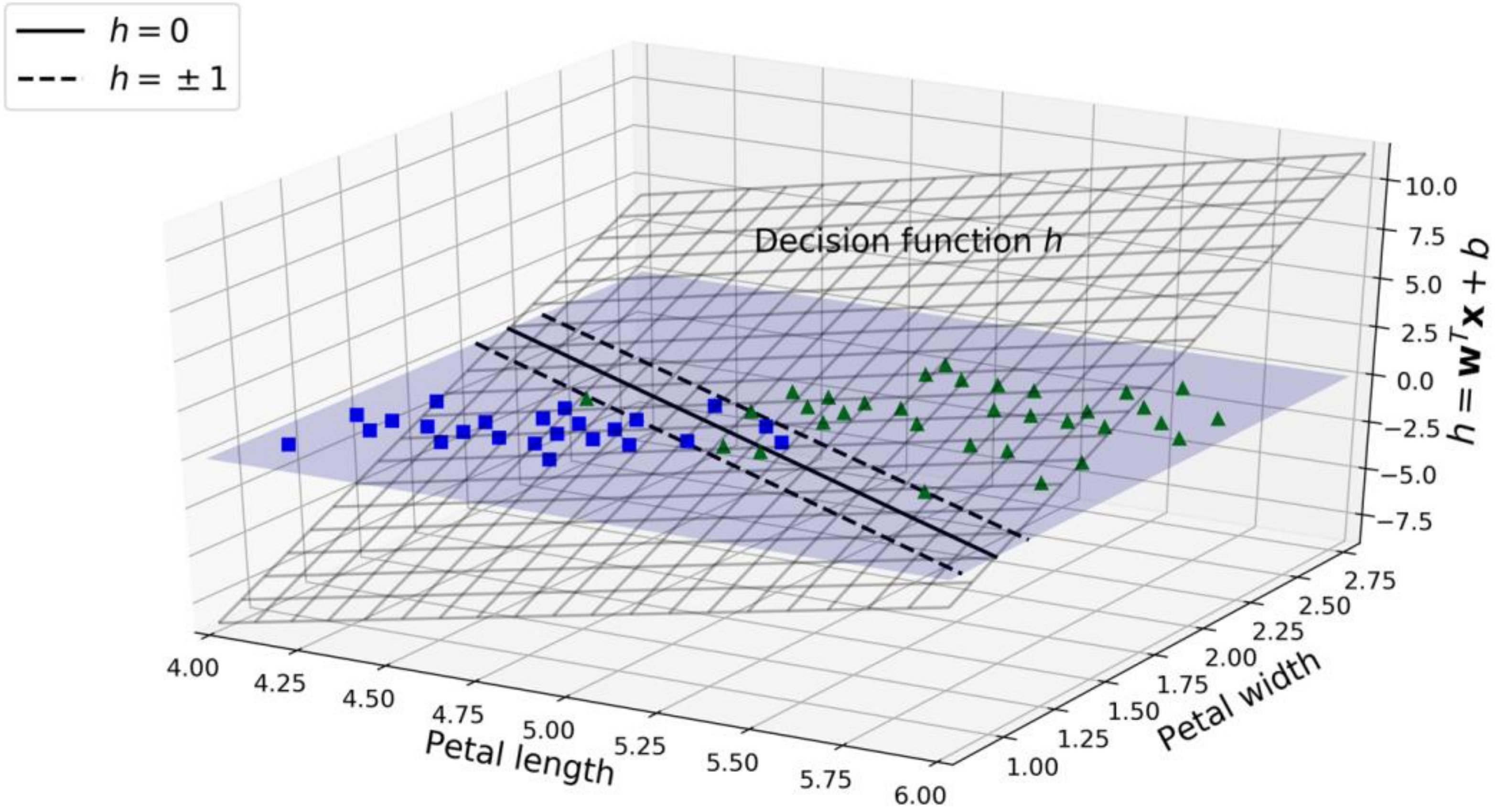
Iris versicolor
Iris setosa
Petal width
Petal length
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
1
2
3
4
5

Support Vectors



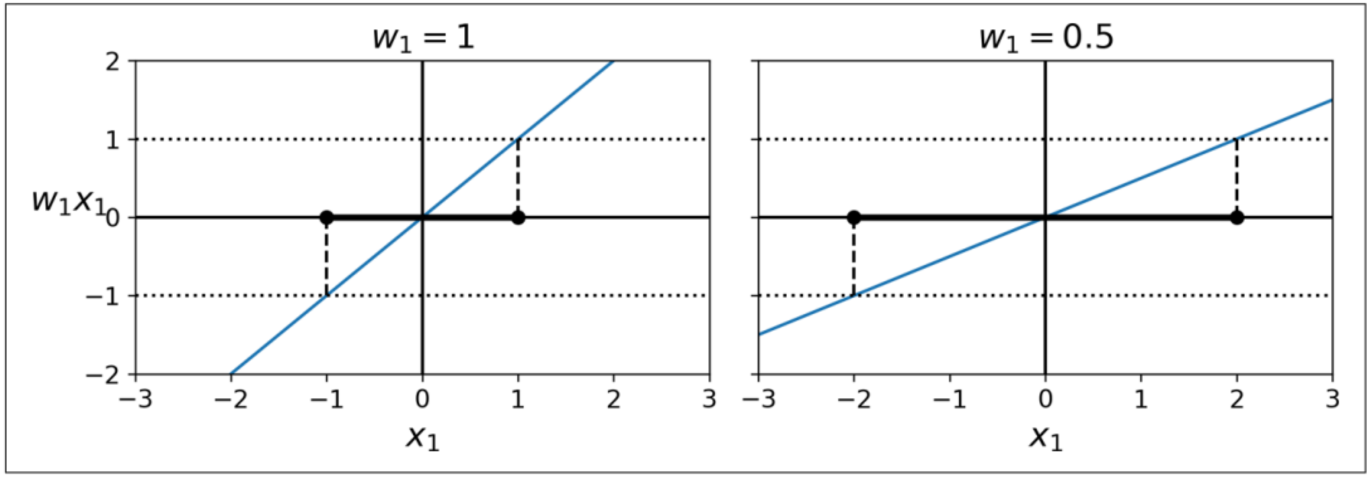
Decision Function: Linear SVM

$$\hat{y} = \begin{cases} 0, if\ W^T X + b < 0 \\ 1, if\ W^T X + b\ \geq 0 \end{cases}$$



Training Objective

- The smaller the weight vector, the larger the margin



- Margin violation. Let $t^{(i)} = -1$ for negative instance (if $y^{(i)} = 0$) and $t^{(i)} = 1$ for positive instance (if $y^{(i)} = 1$)

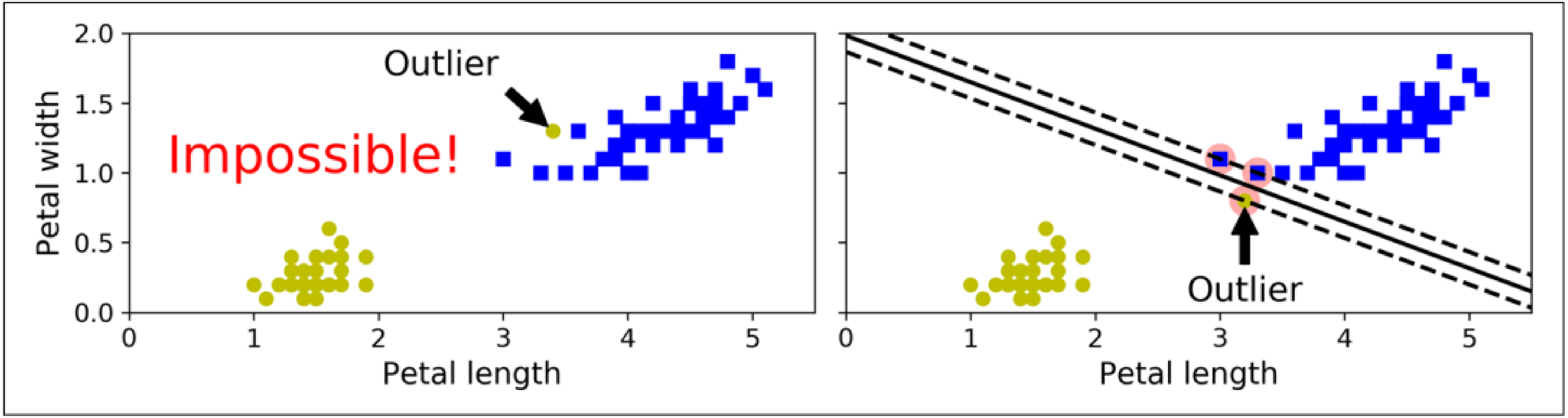
$$t^{(i)}(W^T X^{(i)} + b) \geq 1$$

Hard Margin Linear Classifier Objective Function

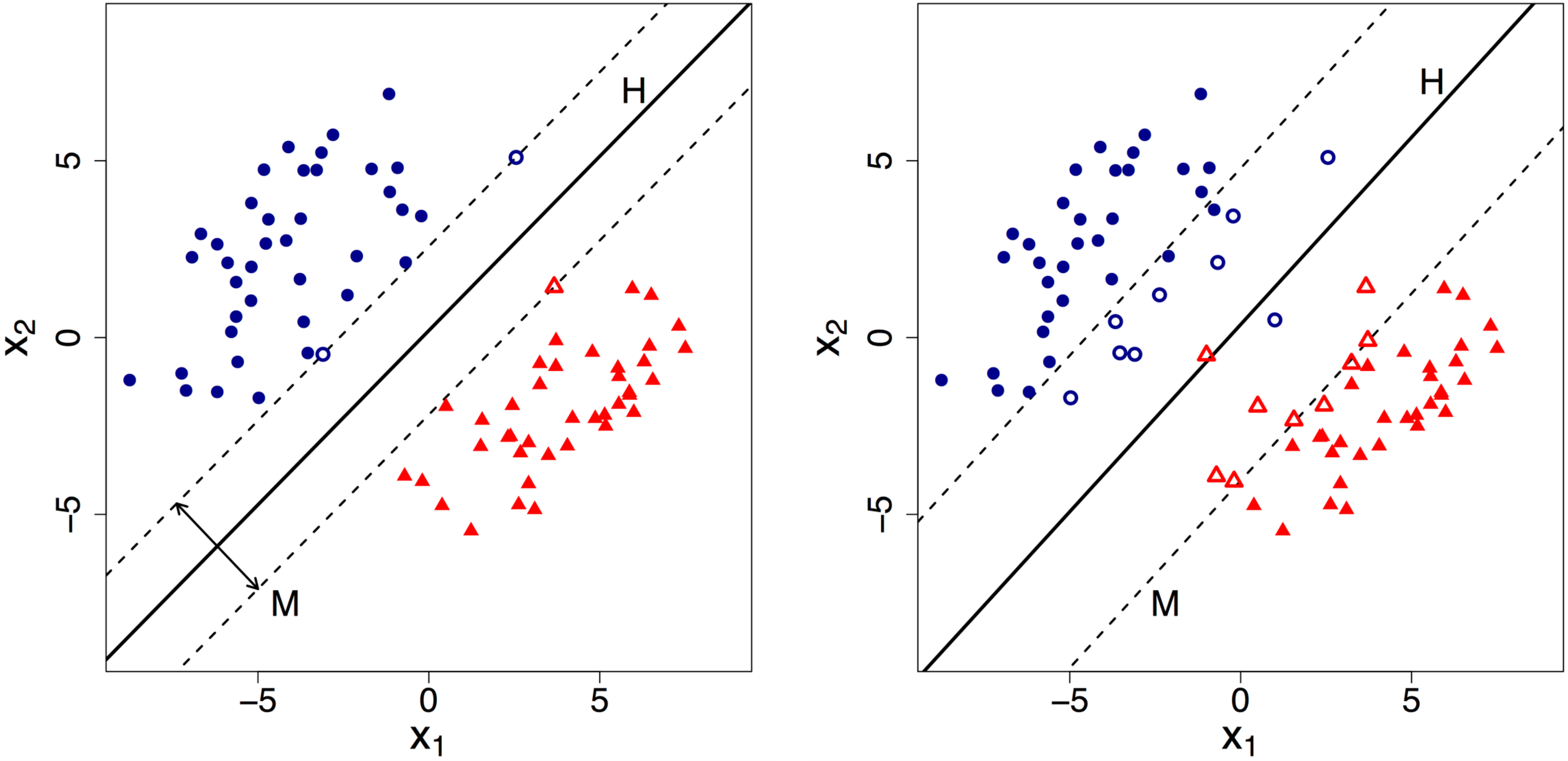
$$\min_{w,b} \frac{1}{2} W^T W$$

$$s.t.\ t^{(i)}\left(W^T X^{(i)} + b\right) \geq 1 \quad for\ i = 1,2,\ldots,m$$

Sensitivity to Outliers



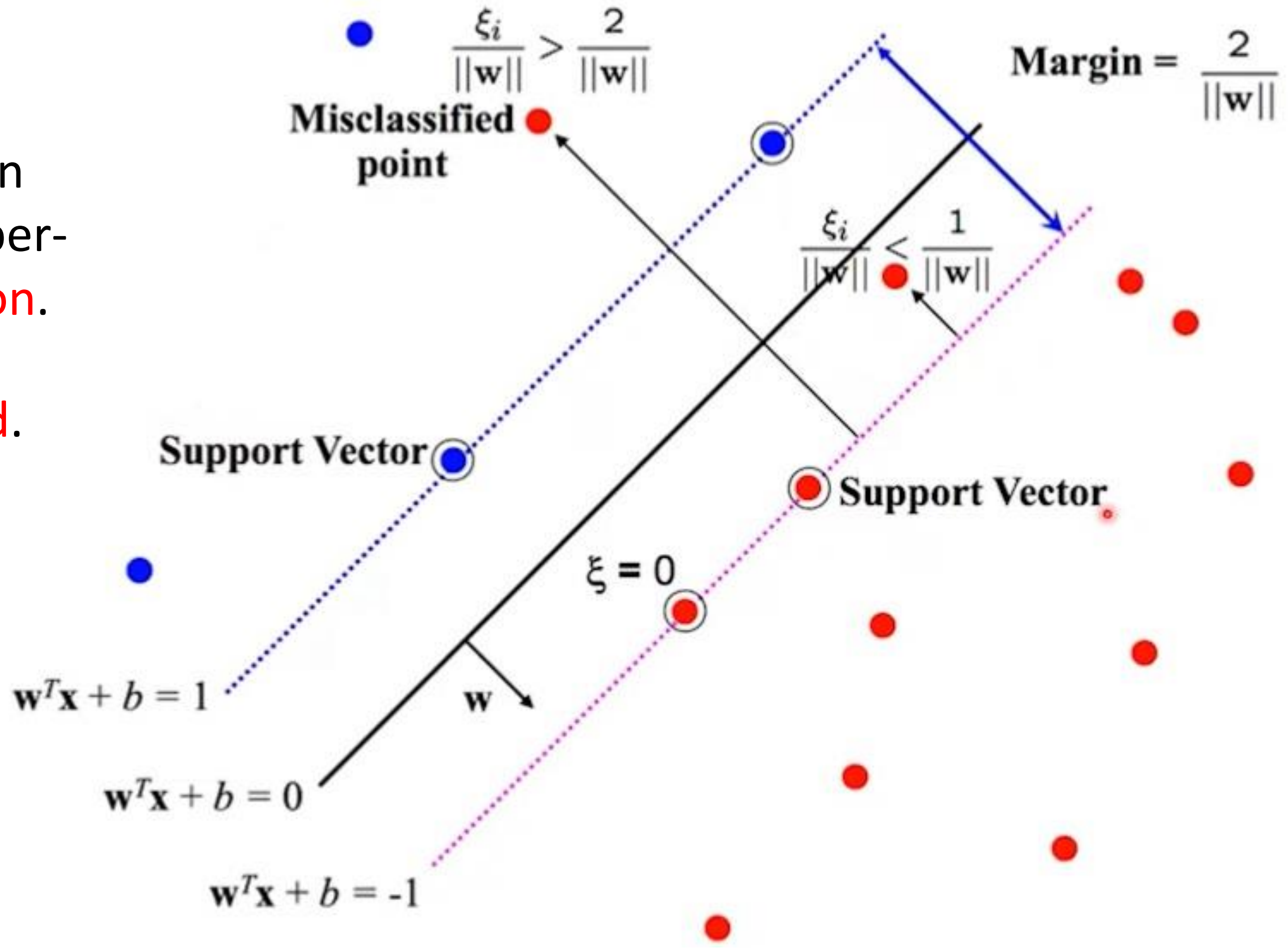
Soft Margin Classification



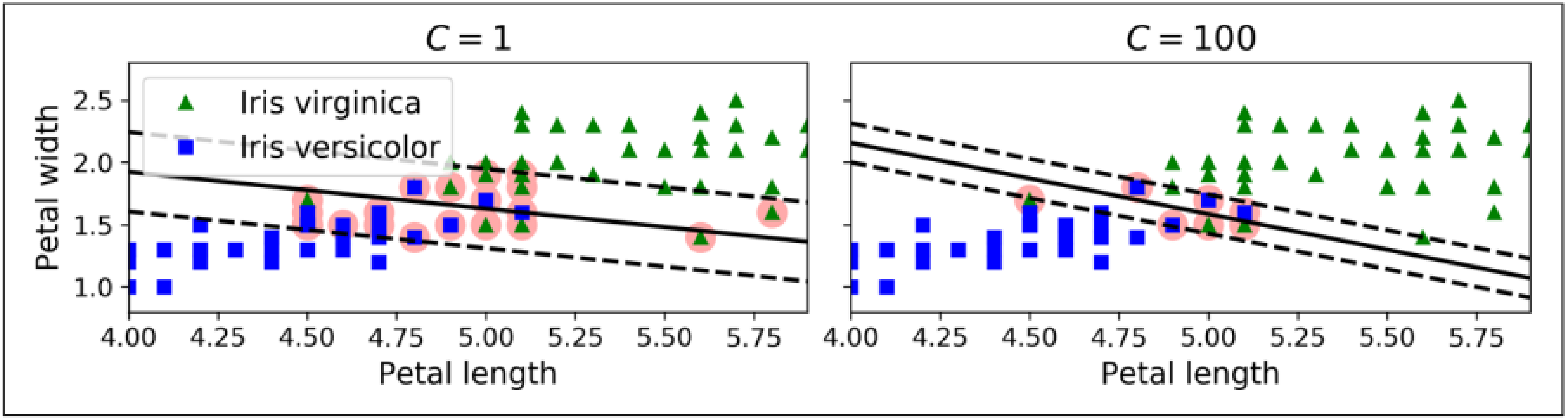
Error Variables

$$\xi_i \geq 0$$

- For $0 < \xi \leq 1$ point is between margin and correct side of hyper-plane. This is a Margin Violation.
- For $\xi > 1$ point is Misclassified.



Regularization via Margin



Soft Margin Linear Classifier Objective Function

$$\min_{w,b,\xi} \frac{1}{2} W^T W + C \sum_{i=1}^{m} \xi^{(i)}$$

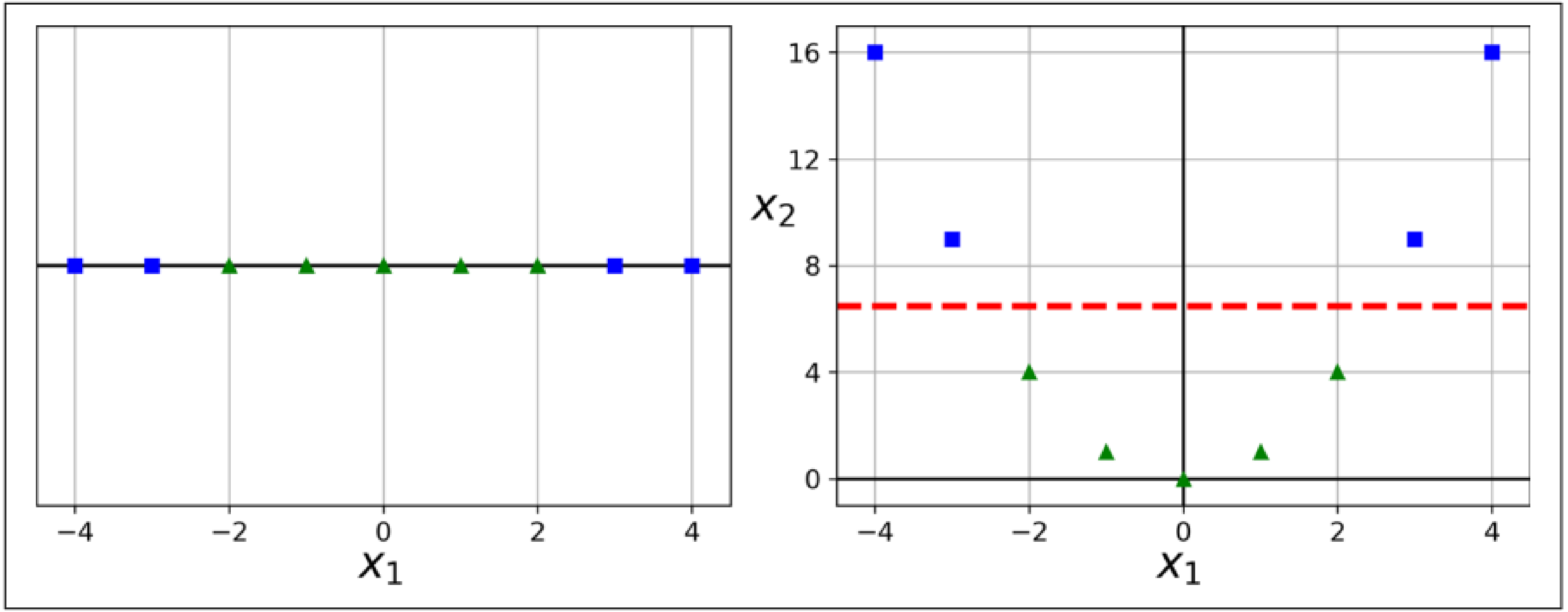
$$s.t.\ \ t^{(i)}\left(W^T X^{(i)} + b\right) \geq 1 - \xi^{(i)} \ \ and \ \ \xi^{(i)} \geq 0 \ \ for\ i = 1,2, \dots, m$$

C is a regularization parameter:

- Small C allows constraints to be easily ignored -> large margin
- Large C makes constraints hard to ignore -> narrow margin
- C = ∞ enforces all constraints: hard margin

Nonlinear SVM

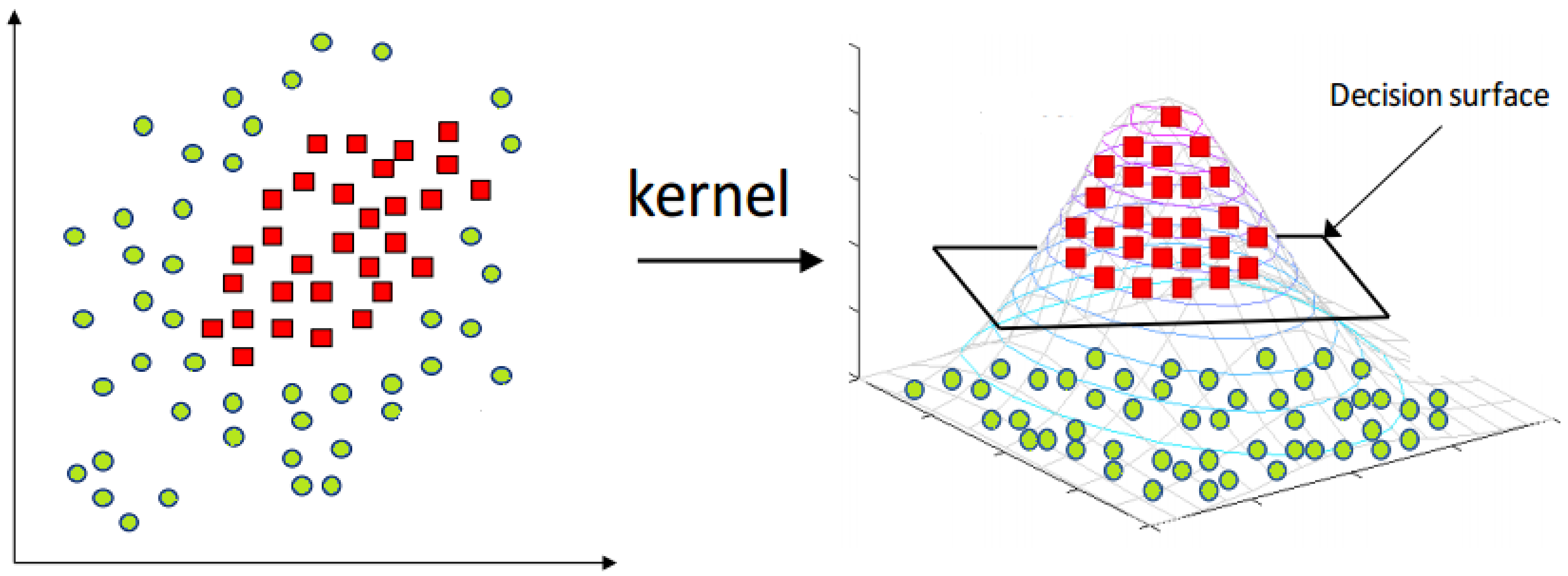
Adding features to make a dataset linearly separable



$$x_2 = (x_1)^2$$

SVM Kernel

Kernel Method



SVM Kernel

Kernel Method

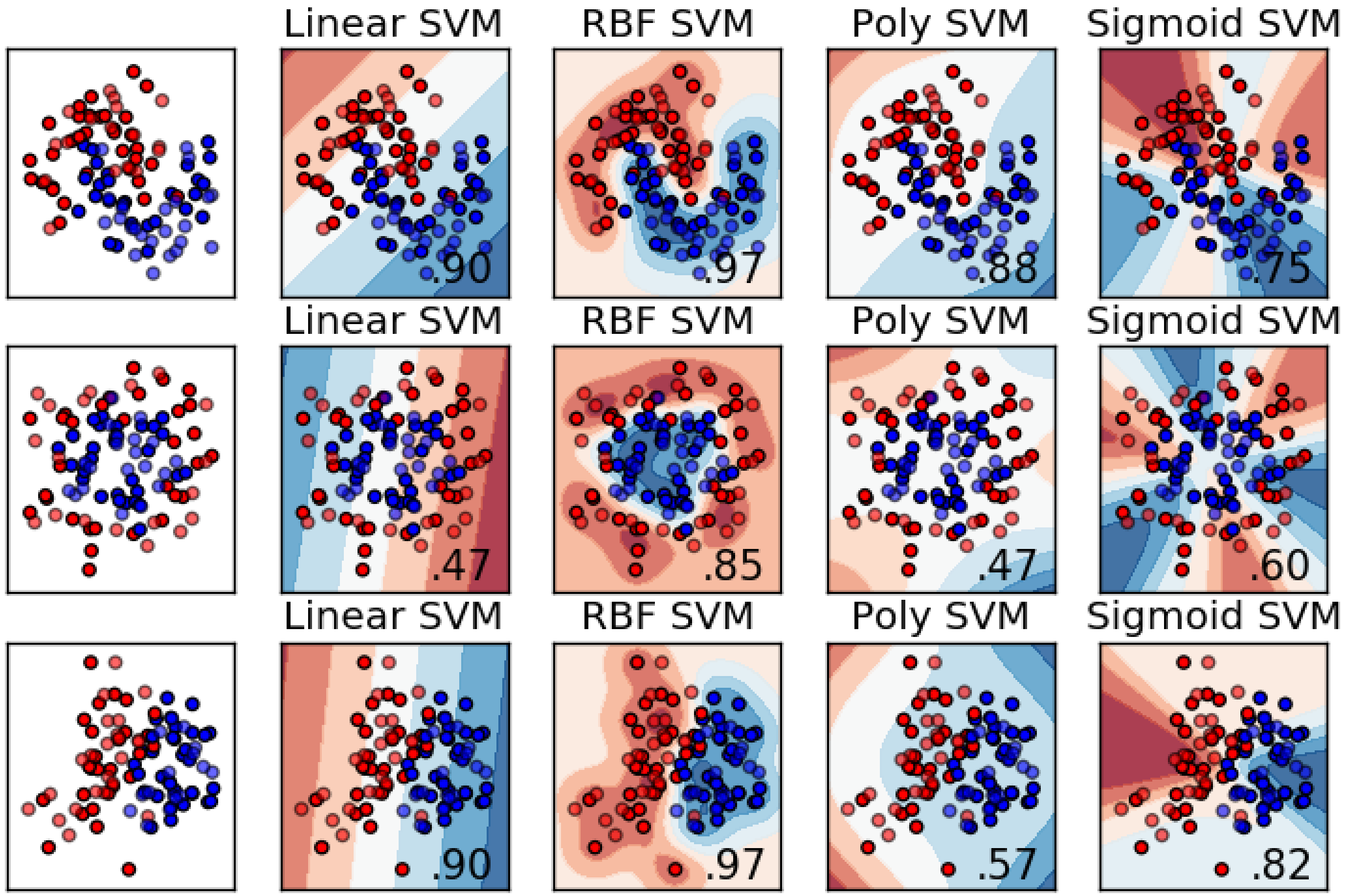
$$\text{Linear:} \quad K(a,b) = a^T b$$

$$\text{Polynomial:} \quad K(a,b) = (\gamma a^T b + r)^d$$

$$\text{Gaussian RBF:} \quad K(a,b) = e^{(-\gamma ||a-b||^2)}$$

$$\text{Sigmoid:} \quad K(a,b) = \tanh(\gamma a^T b + r)$$

SVM Kernel


Linear SVM
RBF SVM
Poly SVM
Sigmoid SVM
.90
.97
.88
.75
Linear SVM
RBF SVM
Poly SVM
Sigmoid SVM
.47
.85
.47
.60
Linear SVM
RBF SVM
Poly SVM
Sigmoid SVM
.90
.97
.57
.82