

Машинное обучение

Семинар 5

13 марта 2014 г.

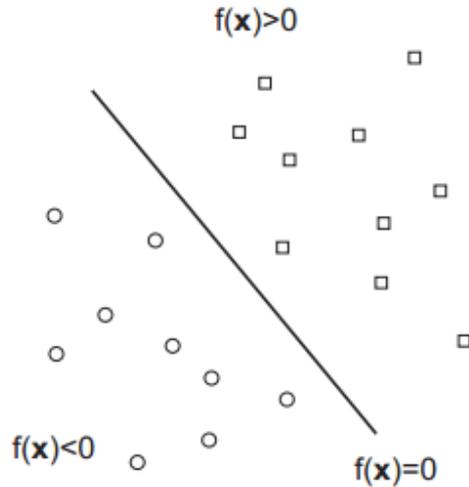
Машины опорных векторов

Линейная модель

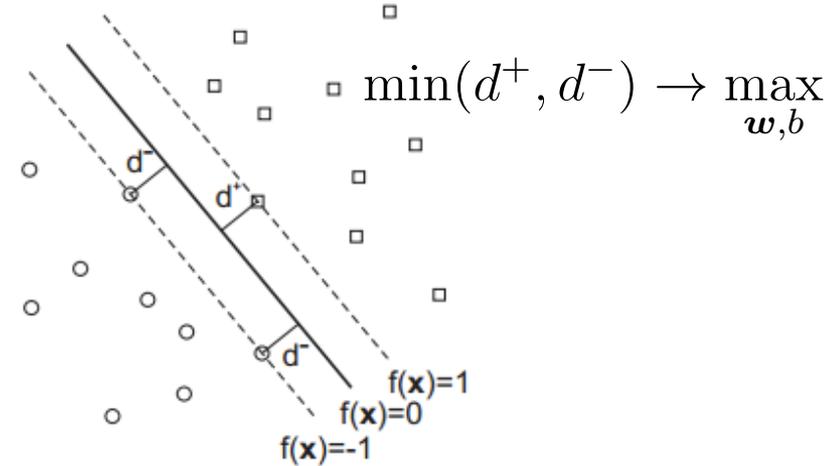
$$f(\mathbf{x}) = \sum_{j=1}^d w_j x(j) + b = \mathbf{w}^T \mathbf{x} + b$$

$$f(\mathbf{x}) = \sum_{j=1}^d w_j \phi_j(\mathbf{x}) + b = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}) + b$$

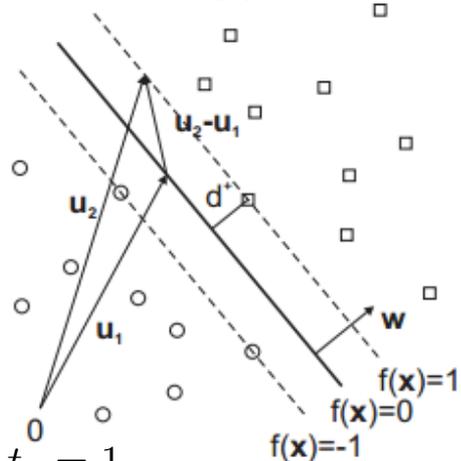
SVM: Разделяющая гиперплоскость



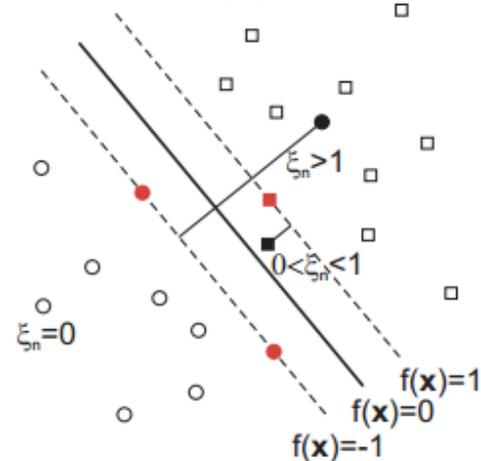
(a)



(b)



(c)



(d)

$$\begin{cases} \frac{2}{\|w\|} \rightarrow \max_{w,b}, \\ w^T x_n + b \geq 1, \text{ если } t_n = 1, \\ w^T x_n + b \leq -1, \text{ если } t_n = -1 \end{cases} \quad (c)$$

SVM: Задача QR (квадратичного программирования)

$$\text{minimize } f(x) := \frac{1}{2} x^T B x - x^T b$$

$$\text{over } x \in \mathbb{R}^n$$

$$\text{subject to } A_1 x = c$$

$$A_2 x \leq d ,$$

$B \in \mathbb{R}^{n \times n}$ is symmetric,

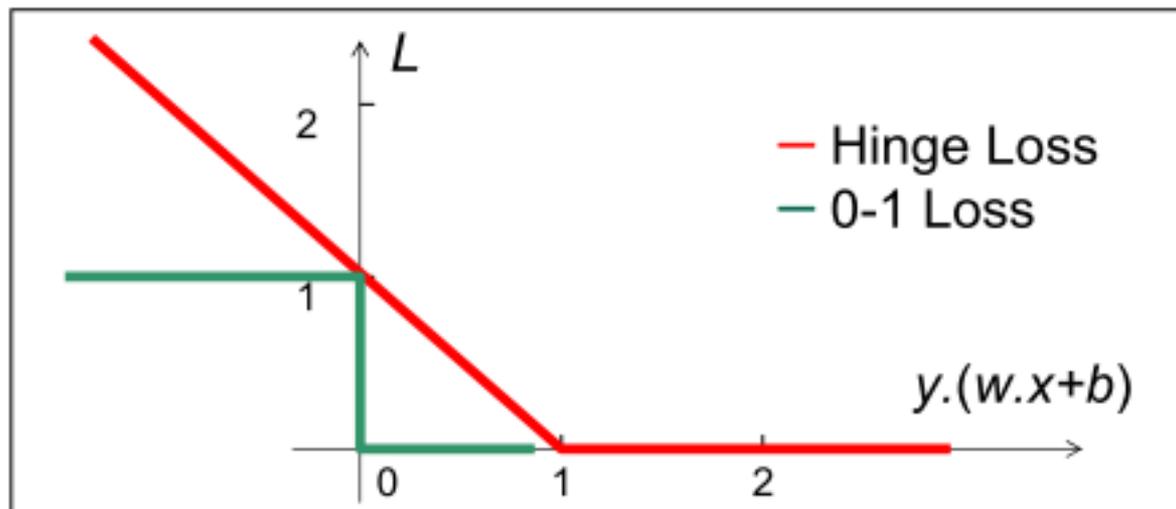
$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{n=1}^N \xi_n \rightarrow \min_{\mathbf{w}, b, \xi},$$

$$t_n(\mathbf{w}^T \mathbf{x}_n + b) \geq 1 - \xi_n,$$

$$\xi_n \geq 0.$$

SVM: Минимизация риска (ошибки)

$$\sum_{n=1}^N l_{hinge}(t_n, f(\mathbf{x}_n)) + \frac{1}{2C} \|\mathbf{w}\|^2 \rightarrow \min_{\mathbf{w}, b}$$



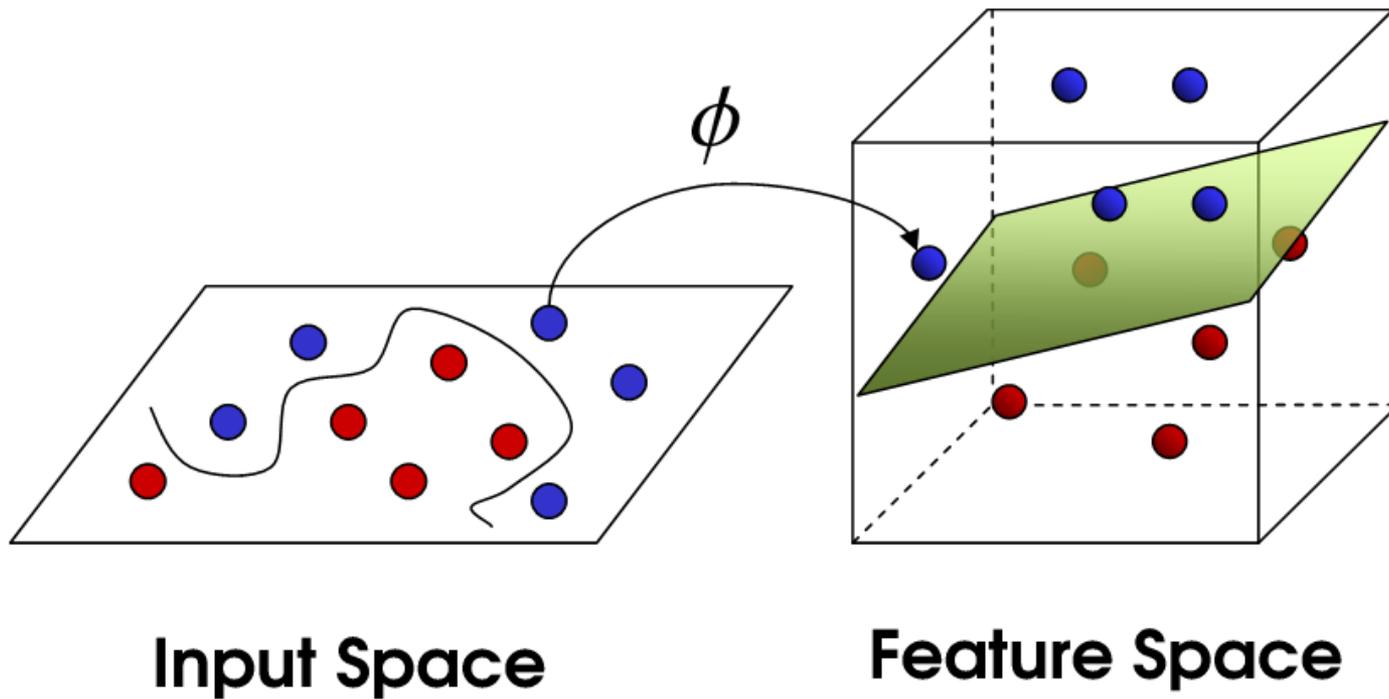
Двойственная задача для SVM

$$\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j K(x_i, x_j) \alpha_i \alpha_j \rightarrow \min_{\alpha}$$

$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, N$$

$$\sum_{i=1}^N y_i \alpha_i = 0$$

Нелинейный переход



Ядра

$$k(x, y) = (x^T y + 1)^n$$

$$k(x, y) = \exp(-\gamma \|x - y\|^2)$$

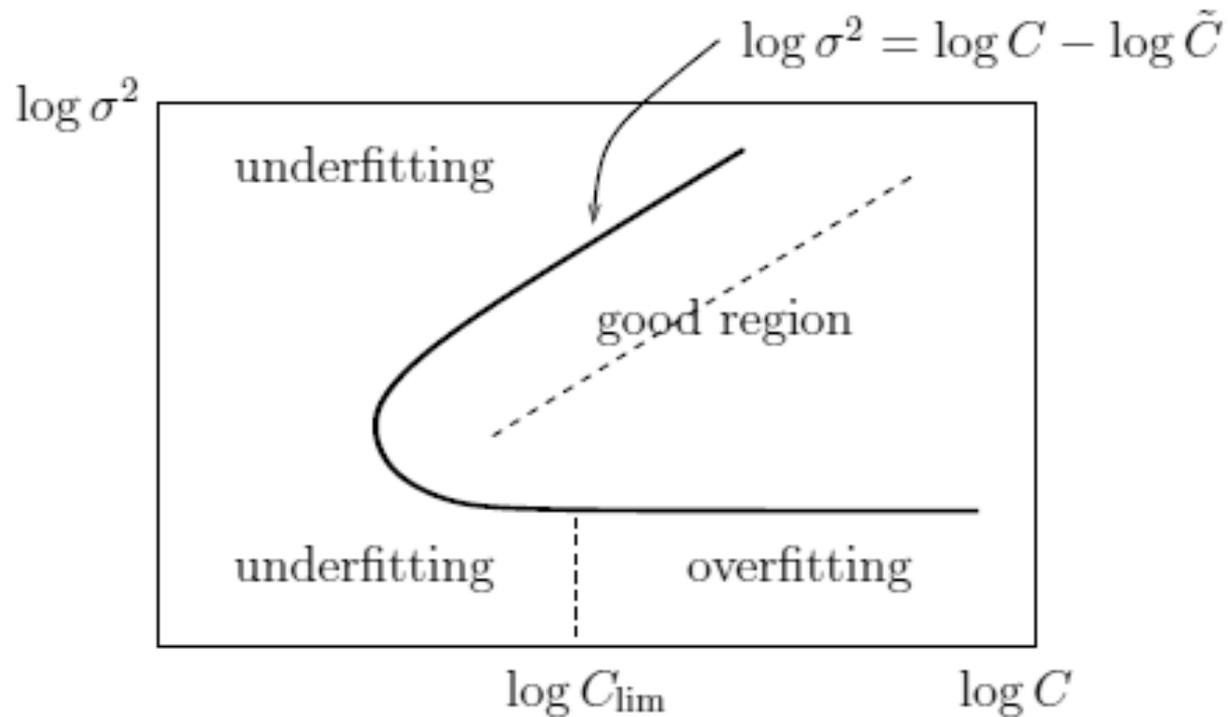
$$k(x, y) = \sum_{i=1}^d \sqrt{x_i y_i}$$

$$k(x, y) = 2 \sum_{i=1}^d \frac{x_i y_i}{x_i + y_i}$$

$$k(x, y) = \sum_{i=1}^N \min(x_i, y_i)$$

RBF-SVM

$$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma^2}\right), \sigma > 0$$



Линейный SVM

Решение **прямой** задачи
Квадратичного программирования
Для SVM

$$f(\mathbf{x}) = \sum_{j=1}^d w_j x(j) + b = \mathbf{w}^T \mathbf{x} + b$$

$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{n=1}^N \xi_n \rightarrow \min_{\mathbf{w}, b, \xi}$$

$$t_n(\mathbf{w}^T \mathbf{x}_n + b) \geq 1 - \xi_n,$$

$$\xi_n \geq 0.$$

Библиотеки и алгоритмы:

- LibLinear
- Pegasos
- Vowpal Wabbit
- SciKit-Learn: LinearSVC

- Легко масштабируется
- Большое число объектов в обучении
- Грубая линейная модель
 - Расширяется путем добавления искусственных признаков

Ядровой SVM

Решение **двойственной** задачи
Квадратичного программирования
Для SVM

$$f(x) = \sum_{i=1}^N \alpha_i k(x, x_i) = \boldsymbol{\alpha}^T \mathbf{k}_i$$
$$\sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j K(x_i, x_j) \alpha_i \alpha_j \rightarrow \min_{\alpha}$$

Библиотеки и алгоритмы:

- LibSVM
- SVMLight
- Плохо масштабируется
- Требуется работа с плотной матрицей (объекты x объекты)
- Нелинейные, гибкие модели для различных предметных областей

Лабораторная работа

- Реализовать
 - Линейный SVM
 - Ядровой SVM
- Имея в распоряжении QP-солвер