

# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)



8.1 对偶支撑向量机动机 (*Motivation of Dual SVM*)

8.2 对偶支撑向量机的拉格朗日分析 (*Lagrange Dual SVM*)

8.3 求解对偶支撑向量机最佳值 (*Solving Dual SVM*)

8.4 对偶支撑向量机讨论 (*Messages behind Dual SVM*)

8.5 核函数支撑向量机 (*Kernel SVM*)

## 8.1 对偶支撑向量机

### 线性支撑向量机模型

最佳的 $(w, b) = ?$

$$\min_w \quad \frac{1}{2} w^T w$$

$$\text{Subject to} \quad y_n(w^T x_n + b) \geq 1$$

$$\text{for } n = 1, 2, \dots, N$$

$$\textcircled{1} \quad Q = \begin{bmatrix} 0 & \mathbf{0}_d^T \\ \mathbf{0}_d & I_d \end{bmatrix}, \quad \mathbf{p} = \mathbf{0}_{d+1},$$

$$\mathbf{a}_n^T = y_n [1 \quad \mathbf{x}_n^T], \quad c_n = 1,$$

$$\textcircled{2} \quad \begin{bmatrix} b \\ w \end{bmatrix} \leftarrow \text{QP}(Q, \mathbf{p}, \mathbf{A}, \mathbf{c})$$

③ 返回最终的 $b$ 和 $w$ 作为学到的 $g_{\text{SVM}}$

## 8.1 对偶支撑向量机

### 非线性支撑向量机模型

最佳的  $(\mathbf{w}, b) = ?$

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{Subject to} \quad & y_n (\mathbf{w}^T \mathbf{z}_n + b) \geq 1 \\ & \mathbf{z}_n = \phi(\mathbf{x}_n) \\ & \text{for } n = 1, 2, \dots, N \end{aligned}$$

如果  $\tilde{d}$  很大, 甚至无穷大, 挑战巨大

$$\textcircled{1} \quad \mathbf{Q} = \begin{bmatrix} 0 & \mathbf{0}_{\tilde{d}}^T \\ \mathbf{0}_{\tilde{d}} & I_{\tilde{d}} \end{bmatrix}, \quad \mathbf{p} = \mathbf{0}_{\tilde{d}+1},$$

$$\mathbf{a}_n^T = y_n [1 \quad \mathbf{x}_n^T], \quad c_n = 1,$$

$$\textcircled{2} \quad \begin{bmatrix} b \\ \mathbf{w} \end{bmatrix} \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c})$$

③ 返回最终的  $b$  和  $\mathbf{w}$  作为学到的  $g_{\text{SVM}}$

QP 针对  $(\tilde{d} + 1)$  个变量和  $N$  个约束条件求解

目的: SVM 算法求解可以不依赖于  $\tilde{d}$  吗?

## 8.1 对偶支撑向量机

### 将有约束条件下的寻优转变为无约束条件下的寻优问题

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{Subject to} \quad & y_n(\mathbf{w}^T \mathbf{z}_n + b) \geq 1 \\ & \text{for } n = 1, 2, \dots, N \end{aligned}$$

用Lagrange乘子  $\alpha_n$  构造Lagrange函数

$$\mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n(\mathbf{w}^T \mathbf{z}_n + b))$$

约束项隐含在 $\max$ 中

$$\text{SVM} \equiv \min_{b, \mathbf{w}} \left( \max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha}) \right) = \min_{b, \mathbf{w}} \left( \infty \text{ if } \textit{violating}, \frac{1}{2} \mathbf{w}^T \mathbf{w} \text{ if } \textit{feasible} \right) = \frac{1}{2} \mathbf{w}^T \mathbf{w}$$

- 任何  $(b, \mathbf{w})$  不在可行域内:  $\max_{\text{all } \alpha_n \geq 0} \left( \blacksquare + \sum_n \alpha_n (\text{一些正数}) \right) \rightarrow \infty$
- 任何  $(b, \mathbf{w})$  在可行域内:  $\max_{\text{all } \alpha_n \geq 0} \left( \blacksquare + \sum_n \alpha_n (\text{所有非正数}) \right) = \blacksquare$

# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)



- 8.1 对偶支撑向量机动机 (*Motivation of Dual SVM*)
- 8.2 对偶支撑向量机的拉格朗日分析 (*Lagrange Dual SVM*)
- 8.3 求解对偶支撑向量机最佳值 (*Solving Dual SVM*)
- 8.4 对偶支撑向量机讨论 (*Messages behind Dual SVM*)
- 8.5 核函数支撑向量机 (*Kernel SVM*)

## 8.2 对偶支撑向量机的拉格朗日分析

在所有的  $\alpha_n \geq 0$  的中挑选任意一个  $\alpha'$  ( $\because \max \geq \text{any}$ )

$$\min_{b, w} \left( \max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, w, \alpha) \right) \geq \min_{b, w} (\mathcal{L}(b, w, \alpha'))$$

如果  $\alpha' \geq 0$  是上式右边  $\max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, w, \alpha)$  中的最佳值 ( $\because \text{best is one of any}$ )

$$\min_{b, w} \left( \max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, w, \alpha) \right) \geq \underbrace{\max_{\text{all } \alpha'_n \geq 0} \left( \min_{b, w} (\mathcal{L}(b, w, \alpha')) \right)}_{\text{Lagrange Dual Problem}}$$

原问题(求解  $b, w$ )与拉格朗日对偶问题(求解  $\alpha$ )的关系

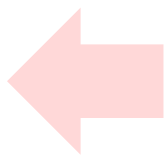
## 8.2 对偶支撑向量机的拉格朗日分析

### 二次规划(QP)满足强对偶特性

$$\underbrace{\min_{b,w} \left( \max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, w, \alpha) \right)}_{\substack{\text{equiv.to original SVM} \\ \text{Primal Problem}}} \geq \underbrace{\max_{\text{all } \alpha_n \geq 0} \left( \min_{b,w} (\mathcal{L}(b, w, \alpha)) \right)}_{\text{Lagrange Dual Problem}}$$

- “ $\geq$ ”是一种弱对偶关系(weak duality)
- “ $=$ ”是一种强对偶关系(strong duality), 如果满足:

二次规划(QP)问题



- 原问题是凸函数
- 原问题存在可行解
- 约束条件为线性的

非线性变换  $\Phi(\mathbf{x}_n)$

## 8.2 对偶支撑向量机的拉格朗日分析

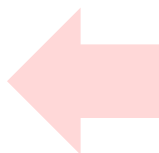
### 二次规划(QP)满足强对偶特性

$$\underbrace{\min_{b,w} \left( \max_{\text{all } \alpha_n \geq 0} \mathcal{L}(b, w, \alpha) \right)}_{\text{equiv.to original SVM Primal Problem}} = \underbrace{\max_{\text{all } \alpha_n \geq 0} \left( \min_{b,w} (\mathcal{L}(b, w, \alpha)) \right)}_{\text{Lagrange Dual Problem}}$$

等式两边对原问题求解  
和对对偶问题求解都能  
得到最优( $b, w, \alpha$ )

- “ $\geq$ ”是一种弱对偶关系(weak duality)
- “=”是一种强对偶关系(strong duality), 如果满足:

二次规划(QP)问题



- 原问题是凸函数
- 原问题存在可行解
- 约束条件为线性的

非线性变换 $\Phi(\mathbf{x}_n)$



## 8.2 对偶支撑向量机的拉格朗日分析

### 对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0} \left( \min_{b, \mathbf{w}} \underbrace{\left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n + b)) \right)}_{\mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})} \right)$$

- “括号”内的问题(*inner problem*)是无约束条件的优化问题

$$\frac{\partial \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})}{\partial b} = - \sum_{n=1}^N \alpha_n y_n = 0$$

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n)) - \sum_{n=1}^N \alpha_n y_n b \right) \right)$$

## 8.2 对偶支撑向量机的拉格朗日分析

### 对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0} \left( \min_{b, \mathbf{w}} \underbrace{\left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n + b)) \right)}_{\mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})} \right)$$

- “括号”内的问题(*inner problem*)是无约束条件的优化问题

$$\frac{\partial \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})}{\partial b} = - \sum_{n=1}^N \alpha_n y_n = 0$$

“括号”内的问题  $b$  取最佳解时:

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n)) \right) \right)$$

## 8.2 对偶支撑向量机的拉格朗日分析

### 对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n)) \right) \right)$$

- “括号”内的问题(*inner problem*)是无约束条件的优化问题

$$\frac{\partial \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})}{\partial w_i} = w_i - \sum_{n=1}^N \alpha_n y_n z_{n,i} = 0 \quad \Rightarrow \quad \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n$$

“括号”内的问题 $\mathbf{w}$ 取最佳解时:

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n - \mathbf{w}^T \mathbf{w} \right) \right)$$

## 8.2 对偶支撑向量机的拉格朗日分析

### 对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n)) \right) \right)$$

- “括号”内的问题(*inner problem*)是无约束条件的优化问题

$$\frac{\partial \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})}{\partial w_i} = w_i - \sum_{n=1}^N \alpha_n y_n z_{n,i} = 0 \quad \Rightarrow \quad \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n$$

“括号”内的问题 $\mathbf{w}$ 取最佳解时:

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n} \left( \min_{b, \mathbf{w}} \left( -\frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n \right) \right)$$

## 8.2 对偶支撑向量机的拉格朗日分析

### 对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0} \left( \min_{b, \mathbf{w}} \left( \frac{1}{2} \mathbf{w}^T \mathbf{w} + \sum_{n=1}^N \alpha_n (1 - y_n (\mathbf{w}^T \mathbf{z}_n)) \right) \right)$$

- “括号”内的问题(*inner problem*)是无约束条件的优化问题

$$\frac{\partial \mathcal{L}(b, \mathbf{w}, \boldsymbol{\alpha})}{\partial w_i} = w_i - \sum_{n=1}^N \alpha_n y_n z_{n,i} = 0 \quad \Rightarrow \quad \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n$$

“括号”内的问题 $\mathbf{w}$ 取最佳解时:

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n} \left( -\frac{1}{2} \left\| \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n \right\|^2 + \sum_{n=1}^N \alpha_n \right)$$

## 8.2 对偶支撑向量机的拉格朗日分析

### KKT条件

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n} \left( -\frac{1}{2} \left\| \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n \right\|^2 + \sum_{n=1}^N \alpha_n \right)$$

如果 $(b, \mathbf{w}, \alpha)$ 是原问题-对偶问题的最佳解(*primal-dual optimal*):

- 原问题可行解(*primal feasible*):  $y_n(\mathbf{w}^T \mathbf{z}_n + b) \geq 1$
- 对偶问题可行解(*dual feasible*):  $\alpha_n \geq 0$
- 对偶“括号”内优化解(*dual-inner optimal*):  $\sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n$
- 原问题“括号”内优化解(*primal-inner optimal*):  $\alpha_n(1 - y_n(\mathbf{w}^T \mathbf{z}_n + b)) = 0$

----称为KKT条件, 为优化的充要条件

利用对偶问题求解最佳 $\alpha$ 后,  
再用KKT条件得到 $(b, \mathbf{w})$

# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)



- 8.1 对偶支撑向量机动机 (*Motivation of Dual SVM*)
- 8.2 对偶支撑向量机的拉格朗日分析 (*Lagrange Dual SVM*)
- 8.3 求解对偶支撑向量机最佳值 (*Solving Dual SVM*)
- 8.4 对偶支撑向量机讨论 (*Messages behind Dual SVM*)
- 8.5 核函数支撑向量机 (*Kernel SVM*)

## 8.3 求解对偶支撑向量机最佳值

### 支撑向量机的对偶问题求解

$$\max_{\text{all } \alpha_n \geq 0, \sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n} \left( -\frac{1}{2} \left\| \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n \right\|^2 + \sum_{n=1}^N \alpha_n \right)$$

标准的硬间隔SVM的对偶问题----求解最佳 $\alpha$ :

$$\min_{\alpha} \quad \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

$$\text{Subject to} \quad \sum_{n=1}^N \alpha_n y_n = 0$$

$$\alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N$$

- $\alpha$ 的目标函数是二次函数、凸函数!  $N$ 个变量
- $\alpha$ 的约束条件是线性函数!  $N+1$ 个约束条件

----二次规划(QP)问题!

QP有成熟方便的办法求优化解!



## 8.3 求解对偶支撑向量机最佳值

SVM对偶问题的一般求解：

最佳的 $\alpha$  = ?

$$\min_{\alpha} \quad \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

二次规划(QP)的求解：

最佳的 $\mathbf{u} \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c}, \mathbf{r}, \mathbf{v})$

$$\min_{\mathbf{u}} \quad \frac{1}{2} \mathbf{u}^T \mathbf{Q} \mathbf{u} + \mathbf{p}^T \mathbf{u}$$

$$\text{Subject to} \quad \mathbf{a}_m^T \mathbf{u} \geq c_m, \mathbf{r}^T \mathbf{u} = \mathbf{v} \\ \text{for } m = 1, 2, \dots, M$$

$$\mathbf{u} = [\alpha_1 \quad \dots \quad \alpha_n \quad \dots \quad \alpha_N]^T \quad \mathbf{Q} = \begin{bmatrix} y_1 y_1 \mathbf{z}_1^T \mathbf{z}_1 & \dots & y_1 y_N \mathbf{z}_1^T \mathbf{z}_N \\ \vdots & y_n y_m \mathbf{z}_n^T \mathbf{z}_m & \vdots \\ y_N y_1 \mathbf{z}_N^T \mathbf{z}_1 & \dots & y_N y_N \mathbf{z}_N^T \mathbf{z}_N \end{bmatrix} \quad \mathbf{p} = -\mathbf{1}_N$$

$$\frac{1}{2} \mathbf{u}^T \mathbf{Q} \mathbf{u} + \mathbf{p}^T \mathbf{u} = \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

## 8.3 求解对偶支撑向量机最佳值

SVM对偶问题的一般求解:

最佳的 $\alpha$  = ?

$$\min_{\alpha} \quad \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

Subject to

$$\alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N$$

$$\mathbf{a}_n^T \mathbf{u} = \alpha_n$$

$$\mathbf{a}_n^T \mathbf{u} \geq c_n \quad \longrightarrow \quad \alpha_n \geq 0$$

二次规划(QP)的求解:

最佳的 $\mathbf{u} \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c}, \mathbf{r}, \mathbf{v})$

$$\min_{\mathbf{u}} \quad \frac{1}{2} \mathbf{u}^T \mathbf{Q} \mathbf{u} + \mathbf{p}^T \mathbf{u}$$

$$\text{Subject to} \quad \mathbf{a}_m^T \mathbf{u} \geq c_m, \quad \text{for } m = 1, 2, \dots, M$$

$$\mathbf{u} = [\alpha_1 \quad \dots \quad \alpha_n \quad \dots \quad \alpha_N]^T \quad q_{nm} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m \quad \mathbf{p} = -\mathbf{1}_N$$

$$\mathbf{a}_1^T = [1 \quad \dots \quad 0 \quad \dots \quad 0],$$

$$\mathbf{a}_n^T = [0 \quad \dots \quad 1 \quad \dots \quad 0], \quad c_n = 0, \quad M = N$$

$$\mathbf{a}_N^T = [0 \quad \dots \quad 0 \quad \dots \quad 1],$$

## 8.3 求解对偶支撑向量机最佳值

SVM对偶问题的一般求解：

最佳的 $\alpha$  = ?

$$\min_{\alpha} \quad \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

Subject to

$$\alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N$$

$$\mathbf{u} = [\alpha_1 \quad \dots \quad \alpha_n \quad \dots \quad \alpha_N]^T \quad q_{nm} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m \quad \mathbf{p} = -\mathbf{1}_N$$

$$\mathbf{r}^T = [y_1 \quad \dots \quad y_n \quad \dots \quad y_N], \quad v = 0$$

$$\mathbf{r}^T \mathbf{u} = \sum_{n=1}^N \alpha_n y_n = v = 0$$

二次规划(QP)的求解：

最佳的 $\mathbf{u} \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c}, \mathbf{r}, v)$

$$\min_{\mathbf{u}} \quad \frac{1}{2} \mathbf{u}^T \mathbf{Q} \mathbf{u} + \mathbf{p}^T \mathbf{u}$$

$$\text{Subject to} \quad \mathbf{a}_m^T \mathbf{u} \geq c_m, \quad \text{for } m = 1, 2, \dots, M$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \dots \\ \mathbf{a}_n^T \\ \dots \\ \mathbf{a}_N^T \end{bmatrix}$$

$$\mathbf{c} = \mathbf{0}_N$$

$$M = N$$

## 8.3 求解对偶支撑向量机最佳值

### 利用二次规划(QP)实现支撑向量机对偶问题求解

$$\begin{aligned}
 \textcircled{1} \quad \mathbf{Q} &= \begin{bmatrix} y_1 y_1 \mathbf{z}_1^T \mathbf{z}_1 & \cdots & y_1 y_N \mathbf{z}_1^T \mathbf{z}_N \\ \vdots & y_n y_m \mathbf{z}_n^T \mathbf{z}_m & \vdots \\ y_N y_1 \mathbf{z}_N^T \mathbf{z}_1 & \cdots & y_N y_N \mathbf{z}_N^T \mathbf{z}_N \end{bmatrix}, \quad \mathbf{p} = -\mathbf{1}_N, \\
 \mathbf{a}_n^T &= [0 \quad \cdots \quad 1 \quad \cdots \quad 0], \quad \mathbf{c}_n = 0, \quad \longrightarrow \quad \mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \\ \vdots \\ \mathbf{a}_N^T \end{bmatrix}, \quad \mathbf{c} = \mathbf{0}_N \\
 \mathbf{r}^T &= [y_1 \cdots y_n \cdots y_N], \quad \mathbf{v} = 0
 \end{aligned}$$

$$\textcircled{2} \quad [\alpha_1 \quad \cdots \quad \alpha_n \quad \cdots \quad \alpha_N]^T \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c}, \mathbf{r}, \mathbf{v})$$

③ 返回最终的 $\alpha$

## 8.3 求解对偶支撑向量机最佳值

### 求解最佳的 $(b, w)$

如果 $(b, w, \alpha)$ 是原问题-对偶问题的最佳解(*primal-dual optimal*):

- 原问题可行解(*primal feasible*):  $y_n(\mathbf{w}^T \mathbf{z}_n + b) \geq 1$
- 对偶问题可行解(*dual feasible*):  $\alpha_n \geq 0$
- 对偶“括号”内优化解(*dual-inner optimal*):  $\sum_{n=1}^N \alpha_n y_n = 0, \mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n$
- 原问题“括号”内优化解(*primal-inner optimal*):  $\alpha_n (1 - y_n(\mathbf{w}^T \mathbf{z}_n + b)) = 0$

---称为KKT条件，为优化的充要条件

挑选任意一个  $\alpha_n > 0$  的样本,  $b = y_n - \mathbf{w}^T \mathbf{z}_n$   $\alpha_n > 0$  的样本是支撑向量

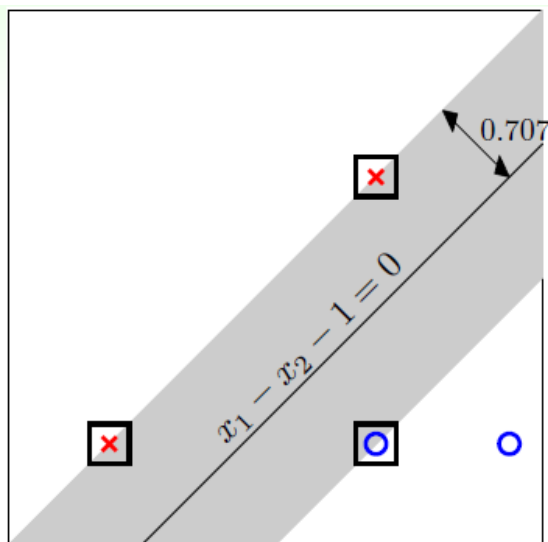
# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)



- 8.1 对偶支撑向量机动机 (*Motivation of Dual SVM*)
- 8.2 对偶支撑向量机的拉格朗日分析 (*Lagrange Dual SVM*)
- 8.3 求解对偶支撑向量机最佳值 (*Solving Dual SVM*)
- 8.4 对偶支撑向量机讨论 (*Messages behind Dual SVM*)
- 8.5 核函数支撑向量机 (*Kernel SVM*)

## 7.3 支撑向量机

为什么叫支撑向量机？



$$\mathbf{X} = \begin{bmatrix} 0 & 0 \\ 2 & 2 \\ 2 & 0 \\ 3 & 0 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} -1 \\ -1 \\ +1 \\ +1 \end{bmatrix}$$

$$\min_{\mathbf{w}} \quad \frac{1}{2} \mathbf{w}^T \mathbf{w}$$

$$\text{Subject to} \quad y_n(\mathbf{w}^T \mathbf{x}_n + b) \geq 1, \text{ for all } n$$

优化得到的解为：

$$w_1 = 1, w_2 = -1, b = -1$$

$$g(\mathbf{x}) = \text{sign}(x_1 - x_2 - 1)$$

$$\text{margin}(\mathbf{w}) = \frac{1}{\|\mathbf{w}\|} = \frac{1}{\sqrt{2}}$$

- 分类面由边界上的样本确定，其他样本不起作用
- 边界上的样本被称为支撑向量(候选)

支撑向量机(SVM)—Support Vector Machine  
----借助支撑向量学到间隔最大分类面

## 8.4 对偶支撑向量机讨论

### SVM原问题

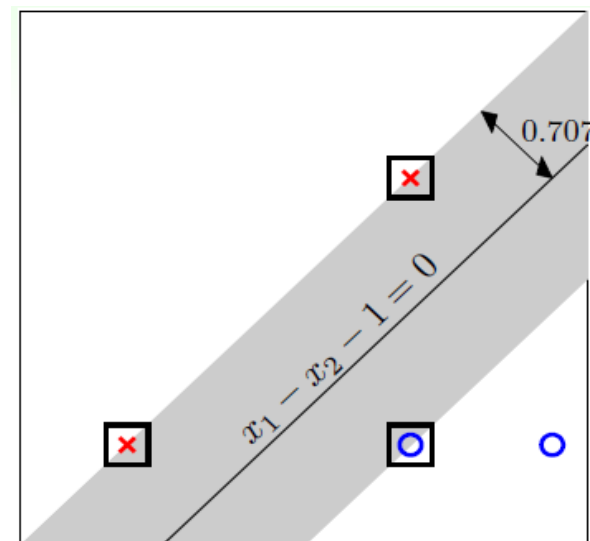
- 分类面由边界上的样本确定，其他样本不起作用
- 边界上的样本被称为支撑向量(候选)

### SVM对偶问题

- $\alpha_n > 0$  的样本落在分类面的边界上
  - $\alpha_n > 0$  的样本  $(\mathbf{z}_n, y_n)$  被称为支撑向量(候选)
- $$SV(\alpha_n > 0 \text{ 的样本}) \subseteq SV(\text{边界上的样本})$$

求解  $\mathbf{w}$  时，只需要支撑向量:  $\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{SV} \alpha_n y_n \mathbf{z}_n$

求解  $b$  时，只需要任一个支撑向量  $(\mathbf{z}_n, y_n)$ :  $b = y_n - \mathbf{w}^T \mathbf{z}_n$



### SVM

----利用对偶问题的最佳解确定支撑向量，从而找到间隔最大的分类面



## 8.4 对偶支撑向量机讨论

SVM:

$$\mathbf{w}_{SVM} = \sum_{n=1}^N \alpha_n (y_n \mathbf{z}_n)$$

$\alpha_n$  由对偶问题的解确定

$\mathbf{w}$  是  $y_n \mathbf{z}_n$  的线性组合

$\mathbf{w}$  被数据  $y_n \mathbf{z}_n$  所表达

$\mathbf{w}$  体现出“模式”

**SVM:** 仅通过SV表达 $\mathbf{w}$

PLA:

$$\mathbf{w}_{PLA} = \sum_{n=1}^N \beta_n (y_n \mathbf{z}_n)$$

$\beta_n$  由分错的样本确定



## 8.4 对偶支撑向量机讨论

SVM的原问题求解:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} \\ \text{Subject to} \quad & y_n (\mathbf{w}^T \mathbf{z}_n + b) \geq 1 \\ & \text{for } n = 1, 2, \dots, N \end{aligned}$$

- $(\tilde{d} + 1)$  个变量和  $N$  个约束条件
- 求解最佳  $(b, \mathbf{w})$

SVM的对偶问题求解:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n \\ \text{Subject to} \quad & \sum_{n=1}^N \alpha_n y_n = 0 \\ & \alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N \end{aligned}$$

- $N$  个变量和  $N+1$  个约束条件
- 求解最佳  $\alpha$ , 确定支撑向量

两种方法都能得到最佳解  $(b, \mathbf{w})$  获得最大间隔分类面  $g_{SVM} = \text{sign}(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n) + b)$

# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)



- 8.1 对偶支撑向量机动机 (*Motivation of Dual SVM*)
- 8.2 对偶支撑向量机的拉格朗日分析 (*Lagrange Dual SVM*)
- 8.3 求解对偶支撑向量机最佳值 (*Solving Dual SVM*)
- 8.4 对偶支撑向量机讨论 (*Messages behind Dual SVM*)
- 8.5 核函数支撑向量机 (*Kernel SVM*)

## 8.5 核函数支撑向量机

研究对偶SVM的动机是不想依赖 $\tilde{d}$

$$\min_{\alpha} \quad \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^N \alpha_n$$

$$\text{Subject to} \quad \sum_{n=1}^N \alpha_n y_n = 0$$

$$\alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N$$

## 8.5 核函数支撑向量机

研究对偶SVM的动机是不想依赖 $\tilde{d}$

$$\min_{\alpha} \quad \frac{1}{2} \alpha^T Q \alpha - \mathbf{1}^T \alpha$$

$$\text{Subject to} \quad \mathbf{y}^T \alpha = 0$$

$$\alpha_n \geq 0, \text{ for } n = 1, 2, \dots, N$$

$$Q = \begin{bmatrix} y_1 y_1 \mathbf{z}_1^T \mathbf{z}_1 & \dots & y_1 y_N \mathbf{z}_1^T \mathbf{z}_N \\ \vdots & y_n y_m \mathbf{z}_n^T \mathbf{z}_m & \vdots \\ y_N y_1 \mathbf{z}_N^T \mathbf{z}_1 & \dots & y_N y_N \mathbf{z}_N^T \mathbf{z}_N \end{bmatrix}$$

对偶SVM真的不依赖于 $\tilde{d}$ ?

- $N$  个变量和  $N+1$  个约束条件
- 对偶问题的 $Q$  是个稠密矩阵
- 每个元素 $q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m$  偶都要做内积运算, 计算代价 $O(\tilde{d})$

能提高 $\mathbf{z}_n^T \mathbf{z}_m = \Phi(\mathbf{x}_n)^T \Phi(\mathbf{x}_m)$   
计算效率吗?

## 8.5 核函数支撑向量机

二次多项式 $\Phi_2(\mathbf{x})$ 的快速内积计算

$$\Phi_2(\mathbf{x}) = (1, x_1, x_2, \dots, x_d, x_1^2, x_1x_2, \dots, x_1x_d, x_2x_1, x_2^2, \dots, x_2x_d, x_dx_1, x_dx_2, \dots, x_d^2)$$

$$\begin{aligned}\Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') &= 1 + \sum_{i=1}^d x_i x'_i + \sum_{i=1}^d \sum_{j=1}^d x_i x_j x'_i x'_j \\ &= 1 + \sum_{i=1}^d x_i x'_i + \sum_{i=1}^d x_i x'_i \sum_{j=1}^d x_j x'_j \\ &= 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')\end{aligned}$$

计算复杂度从 $O(\tilde{d})$ 下降到 $O(d)$

## 8.5 核函数支撑向量机

核函数:  $\Phi$  变换+内积计算  $\longrightarrow K_{\Phi}(\mathbf{x}, \mathbf{x}') = \Phi^T(\mathbf{x})\Phi(\mathbf{x}')$

$$K_{\Phi_2}(\mathbf{x}, \mathbf{x}') = \Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') = 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')$$

---

$$q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m = y_n y_m \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m) = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$$

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{SV} \alpha_n y_n \mathbf{z}_n$$

支撑SV( $\mathbf{z}_m, y_m$ ):  $b = y_m - \mathbf{w}^T \mathbf{z}_m = y_m - \left( \sum_{SV} \alpha_n y_n \mathbf{z}_n \right)^T \mathbf{z}_m$

输入测试样本  $\mathbf{x}$ :  $g_{SVM} = \text{sign}(\mathbf{w}^T \Phi(\mathbf{x}) + b)$

## 8.5 核函数支撑向量机

核函数:  $\Phi$ 变换+内积计算  $\longrightarrow K_{\Phi}(\mathbf{x}, \mathbf{x}') = \Phi^T(\mathbf{x})\Phi(\mathbf{x}')$

$$K_{\Phi_2}(\mathbf{x}, \mathbf{x}') = \Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') = 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')$$

---

$$q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m = y_n y_m \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m) = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$$

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{\text{SV}} \alpha_n y_n \mathbf{z}_n$$

支撑SV( $\mathbf{z}_m, y_m$ ):  $b = y_m - \mathbf{w}^T \mathbf{z}_m = y_m - \sum_{\text{SV}} \alpha_n y_n \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m)$

输入测试样本  $\mathbf{x}$ :  $g_{\text{SVM}} = \text{sign}(\mathbf{w}^T \Phi(\mathbf{x}) + b)$



## 8.5 核函数支撑向量机

核函数:  $\Phi$ 变换+内积计算  $\longrightarrow K_{\Phi}(\mathbf{x}, \mathbf{x}') = \Phi^T(\mathbf{x})\Phi(\mathbf{x}')$

$$K_{\Phi_2}(\mathbf{x}, \mathbf{x}') = \Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') = 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')$$

---

$$q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m = y_n y_m \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m) = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$$

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{SV} \alpha_n y_n \mathbf{z}_n$$

支撑SV( $\mathbf{z}_m, y_m$ ):  $b = y_m - \mathbf{w}^T \mathbf{z}_m = y_m - \sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$

输入测试样本  $\mathbf{x}$ :  $g_{SVM} = \text{sign}(\mathbf{w}^T \Phi(\mathbf{x}) + b) = \text{sign}\left(\left(\sum_{SV} \alpha_n y_n \mathbf{z}_n\right)^T \Phi(\mathbf{x}) + b\right)$

## 8.5 核函数支撑向量机

核函数:  $\Phi$  变换+内积计算  $\longrightarrow K_{\Phi}(\mathbf{x}, \mathbf{x}') = \Phi^T(\mathbf{x})\Phi(\mathbf{x}')$

$$K_{\Phi_2}(\mathbf{x}, \mathbf{x}') = \Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') = 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')$$


---

$$q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m = y_n y_m \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m) = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$$

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{SV} \alpha_n y_n \mathbf{z}_n$$

支撑SV( $\mathbf{z}_m, y_m$ ):  $b = y_m - \mathbf{w}^T \mathbf{z}_m = y_m - \sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$

输入测试样本  $\mathbf{x}$ :  $g_{SVM} = \text{sign}(\mathbf{w}^T \Phi(\mathbf{x}) + b) = \text{sign}\left(\sum_{SV} \alpha_n y_n \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}) + b\right)$

## 8.5 核函数支撑向量机

核函数:  $\Phi$  变换+内积计算  $\longrightarrow K_{\Phi}(\mathbf{x}, \mathbf{x}') = \Phi^T(\mathbf{x})\Phi(\mathbf{x}')$

$$K_{\Phi_2}(\mathbf{x}, \mathbf{x}') = \Phi_2^T(\mathbf{x})\Phi_2(\mathbf{x}') = 1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')(\mathbf{x}^T \mathbf{x}')$$

$$q_{n,m} = y_n y_m \mathbf{z}_n^T \mathbf{z}_m = y_n y_m \Phi^T(\mathbf{x}_n) \Phi(\mathbf{x}_m) = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$$

$$\mathbf{w} = \sum_{n=1}^N \alpha_n y_n \mathbf{z}_n = \sum_{SV} \alpha_n y_n \mathbf{z}_n$$

支撑SV  $(\mathbf{z}_m, y_m)$ :  $b = y_m - \mathbf{w}^T \mathbf{z}_m = y_m - \sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$

输入测试样本  $\mathbf{x}$ :  $g_{SVM} = \text{sign}(\mathbf{w}^T \Phi(\mathbf{x}) + b) = \text{sign}(\sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}) + b)$

利用核函数, 将  
计算复杂度从  
 $O(\tilde{d})$  下降到  $O(d)$   
避免依赖  $\tilde{d}$

## 8.5 核函数支撑向量机

### 利用二次规划(QP)实现核函数支撑向量机求解

- ①  $q_{n,m} = y_n y_m K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m)$ ,  $\mathbf{p} = -\mathbf{1}_N$ , 由约束条件得到 $(\mathbf{A}, \mathbf{c}, \mathbf{r}, v)$
- ②  $\alpha \leftarrow \text{QP}(\mathbf{Q}, \mathbf{p}, \mathbf{A}, \mathbf{c}, \mathbf{r}, v)$
- ③ 任选一支撑SV $(\mathbf{x}_m, y_m)$ :  $b \leftarrow (y_m - \sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}_m))$
- ④ 对于测试样本  $\mathbf{x}$ :  $g_{SVM} = \text{sign}(\sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}) + b)$

- 步骤①的时间复杂度:  $O(N^2) \cdot (\text{kernel evaluation})$
- 步骤②的开销:  $N$ 个变量,  $N + 1$ 个约束
- 步骤③和④的复杂度:  $O(\#SV) \cdot (\text{kernel evaluation})$

## 8.5 核函数支撑向量机

### 二次多项式核函数的一般表达式

$\Phi_2(\mathbf{x})$	$K_2(\mathbf{x}, \mathbf{x}')$
$(1, x_1, \dots, x_d, x_1^2, x_1x_2, \dots, x_d^2)$	$1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')^2$
$(1, \sqrt{2}x_1, \dots, \sqrt{2}x_d, x_1^2, x_1x_2, \dots, x_d^2)$	$1 + 2\mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')^2$
$(1, \sqrt{2\gamma}x_1, \dots, \sqrt{2\gamma}x_d, x_1^2, x_1x_2, \dots, x_d^2)$	$1 + 2\gamma\mathbf{x}^T \mathbf{x}' + \gamma^2(\mathbf{x}^T \mathbf{x}')^2$

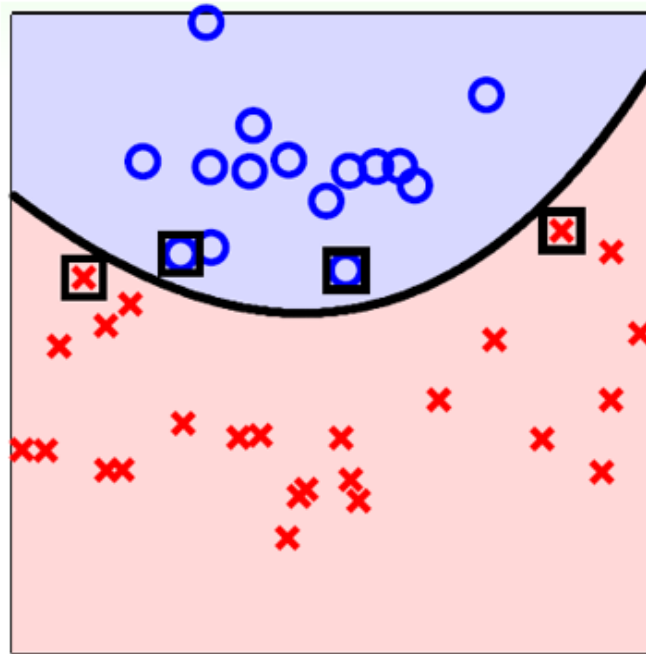
$$K_2(\mathbf{x}, \mathbf{x}') = (1 + \gamma\mathbf{x}^T \mathbf{x}')^2 \quad \gamma > 0$$

不同的二次多项式变换：

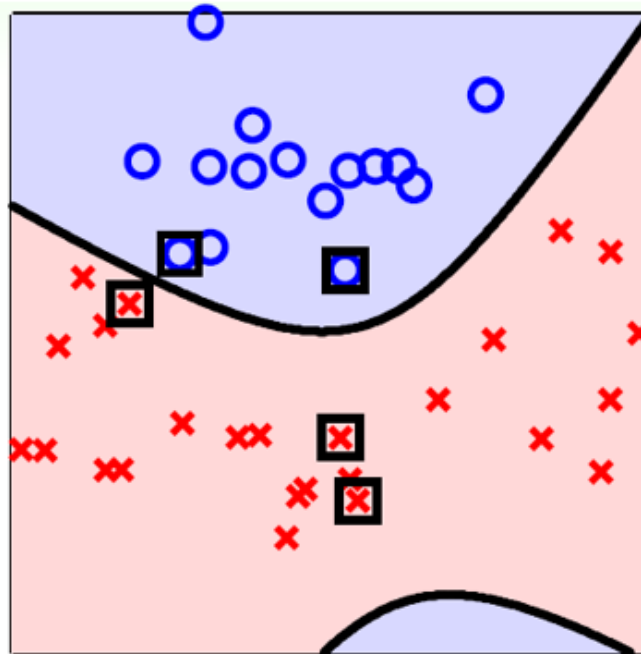
- 升维次数相同
- 内积结果不同  $\longrightarrow$  分类面不同

## 8.5 核函数支撑向量机

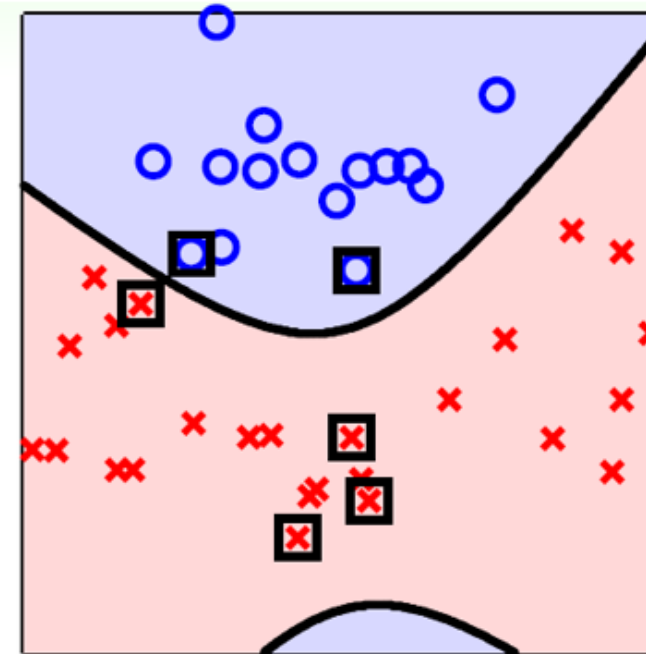
### 二次多项式核函数的一般表达式



$$(1 + 0.001 \mathbf{x}^T \mathbf{x}')^2$$



$$1 + \mathbf{x}^T \mathbf{x}' + (\mathbf{x}^T \mathbf{x}')^2$$



$$(1 + 1000 \mathbf{x}^T \mathbf{x}')^2$$

- 核函数不同  $\longrightarrow$  支撑向量(SVs)不同, 分类面( $g_{SVM}$ )不同
- 核函数变化  $\longrightarrow$  Margin也会变化

## 8.5 核函数支撑向量机

### 多项式核函数的一般表达式

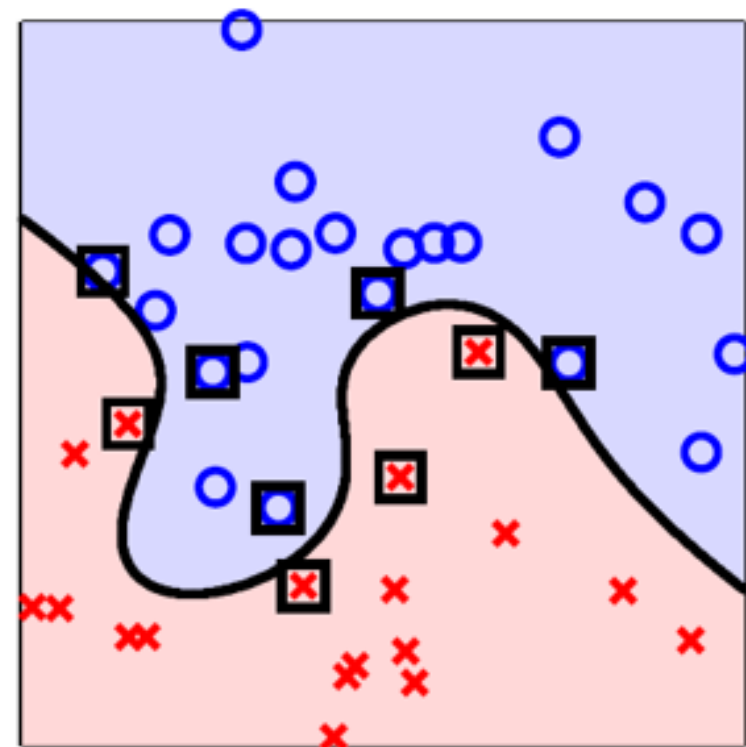
$$K_2(\mathbf{x}, \mathbf{x}') = (\zeta + \gamma \mathbf{x}^T \mathbf{x}')^2 \quad \gamma > 0, \zeta \geq 0$$

$$K_3(\mathbf{x}, \mathbf{x}') = (\zeta + \gamma \mathbf{x}^T \mathbf{x}')^3 \quad \gamma > 0, \zeta \geq 0$$

⋮

$$K_Q(\mathbf{x}, \mathbf{x}') = (\zeta + \gamma \mathbf{x}^T \mathbf{x}')^Q \quad \gamma > 0, \zeta \geq 0$$

- $Q$ 、 $\gamma$ 、 $\zeta$  确定了多项式核函数的形式
- 利用核函数可不依赖于  $\tilde{d}$  获得大间隔分类面



Margin为1的10次多项式

$SVM + \text{Polynomial Kernel} = \text{Polynomial SVM}$

## 8.5 核函数支撑向量机

### 核函数能实现无穷维变换?

当 $\Phi(\mathbf{x})$ 为无穷维变换时, 借助核函数 $K(\mathbf{x}, \mathbf{x}')$ 技巧仍可进行有效计算

为简单起见, 考虑一维变量, 即:  $\mathbf{x} = (x)$

$$\begin{aligned} K(x, x') &= \exp(-(x - x')^2) = \exp(-(x)^2) \exp(-(x')^2) \exp(2xx') \\ &= \exp(-(x)^2) \exp(-(x')^2) \left( \sum_{i=0}^{\infty} \frac{(2xx')^i}{i!} \right) \\ &= \sum_{i=0}^{\infty} \left( \exp(-(x)^2) \exp(-(x')^2) \sqrt{\frac{2^i}{i!}} \sqrt{\frac{2^i}{i!}} (x)^i (x')^i \right) \\ \Phi(x) &= \exp(-(x)^2) \cdot \left( 1, \sqrt{\frac{2^1}{1!}} x, \sqrt{\frac{2^2}{2!}} x^2, \sqrt{\frac{2^3}{3!}} x^3, \dots, \dots \right) \end{aligned}$$



## 8.5 核函数支撑向量机

### 核函数能实现无穷维变换?

当 $\Phi(\mathbf{x})$ 为无穷维变换时, 借助核函数 $K(\mathbf{x}, \mathbf{x}')$ 技巧仍可进行有效计算

为简单起见, 考虑一维变量, 即:  $\mathbf{x} = (x)$

$$K(x, x') = \Phi(x)^T \Phi(x') = \exp(-(x - x')^2) = \exp(-(x)^2) \exp(-(x')^2) \exp(2xx')$$

$$= \exp(-(x)^2) \exp(-(x')^2) \left( \sum_{i=0}^{\infty} \frac{(2xx')^i}{i!} \right)$$

$$= \sum_{i=0}^{\infty} (\exp(-(x)^2) \exp(-(x')^2) \sqrt{\frac{2^i}{i!}} \sqrt{\frac{2^i}{i!}} (x)^i (x')^i)$$

$$\Phi(x) = \exp(-(x)^2) \cdot (1, \sqrt{\frac{2^1}{1!}} x, \sqrt{\frac{2^2}{2!}} x^2, \sqrt{\frac{2^3}{3!}} x^3, \dots, \dots)$$

高斯核函数 $K(\mathbf{x}, \mathbf{x}')$ 的一般式

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2)$$

$$\gamma > 0$$

## 8.5 核函数支撑向量机

### 高斯核函数SVM

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2) \quad \gamma > 0$$

输入测试样本  $\mathbf{x}$ :

$$g_{SVM} = \text{sign}\left(\sum_{SV} \alpha_n y_n K_{\Phi}(\mathbf{x}_n, \mathbf{x}) + b\right)$$
$$= \text{sign}\left(\sum_{SV} \alpha_n y_n \exp(-\gamma \|\mathbf{x} - \mathbf{x}_n\|^2) + b\right)$$

- 以所有支撑向量  $\mathbf{x}_n$  为中心的高斯函数的线性组合
- 也被称之为径向基核函数(Radial Basis Function, RBF)

高斯核函数SVM: 通过求取  $\alpha_n$  确定所有支撑向量  $\mathbf{x}_n$  , 构造以支撑向量为中心的高斯函数的线性组合, 实现在无穷维空间获得最大间隔分类面

# 第八讲 对偶支撑向量机与核支撑向量机(*Dual SVM & Kernel SVM*)

## 8.1 对偶支撑向量机动机(*Motivation of Dual SVM*)

希望不依赖于非线性变换后升维的  $\tilde{d}$

## 8.2 对偶支撑向量机的拉格朗日分析(*Lagrange Dual SVM*)

通过 **KKT** 条件 将原问题和对偶问题相关联获得最佳分类面

## 8.3 求解对偶支撑向量机最佳值(*Solving Dual SVM*)

仍然是二次规划问题，可以方便求解

## 8.4 对偶支撑向量机讨论(*Messages behind Dual SVM*)

由支撑向量确定最大间隔分类面

## 8.5 核函数支撑向量机(*Kernel SVM*)

利用核函数避免了对升维后  $\tilde{d}$  的依赖，且高效求解