

### 目录

精通数据科学

从线性回归初深度管

ONE 从非线性到线性

精通数据科学

从绝对国的多洲深度管

空间变换 从编馆回的粉珠度管的

TWO 拉格朗日对偶 

THREE 支持向量。

SVM名字来源

精通数据科学

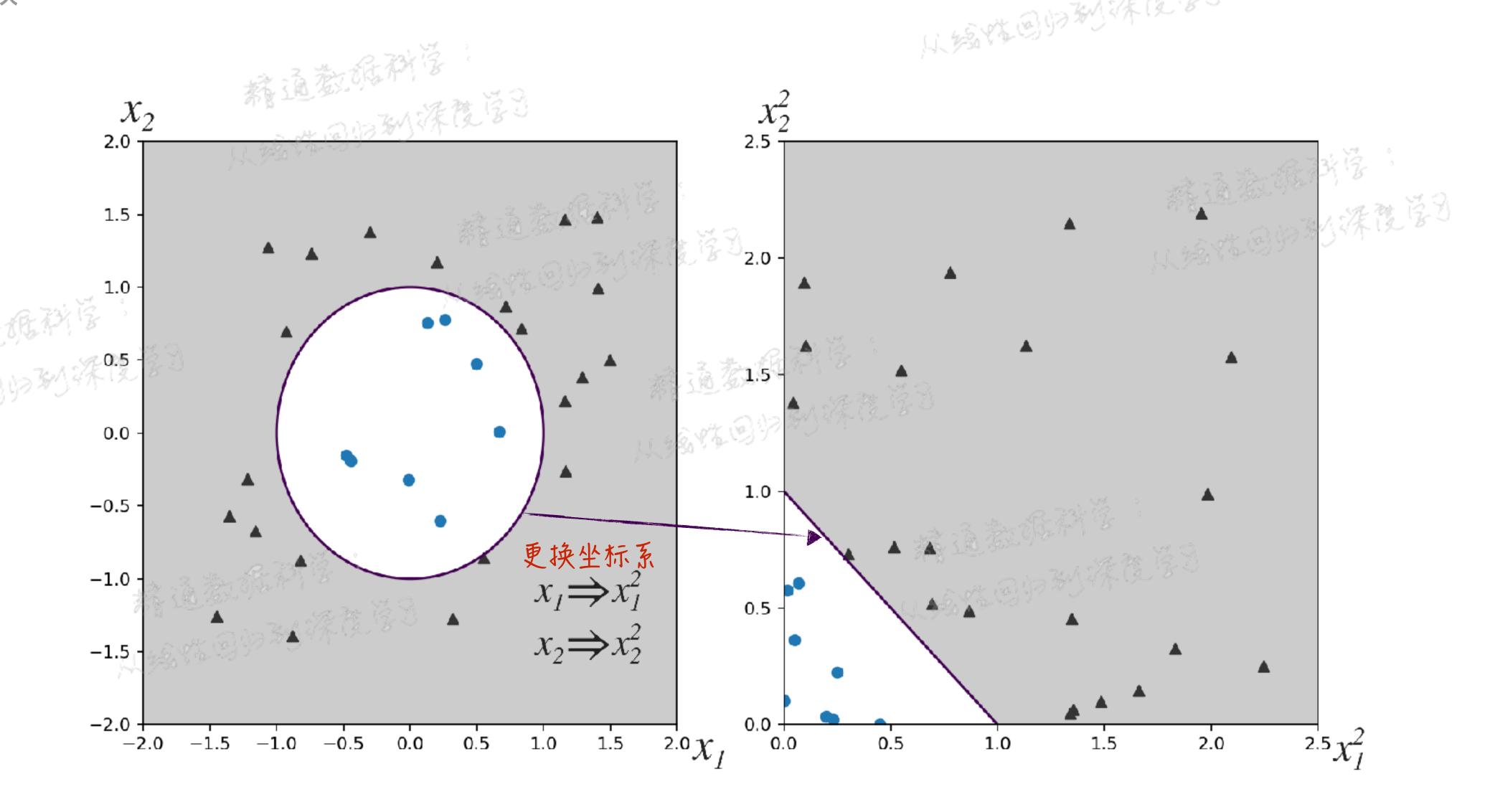
以给此回的那么深度管

精通数据科学

从绝对回的物体度管

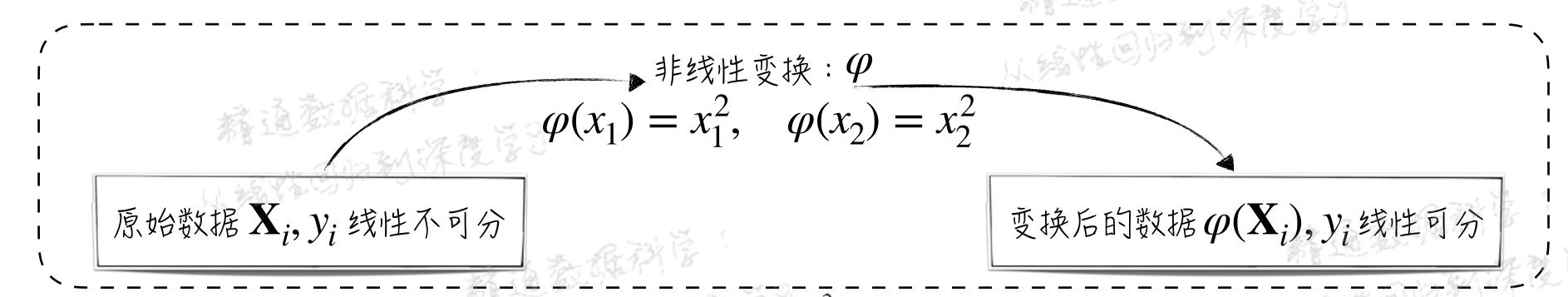
# 从非线性到线性

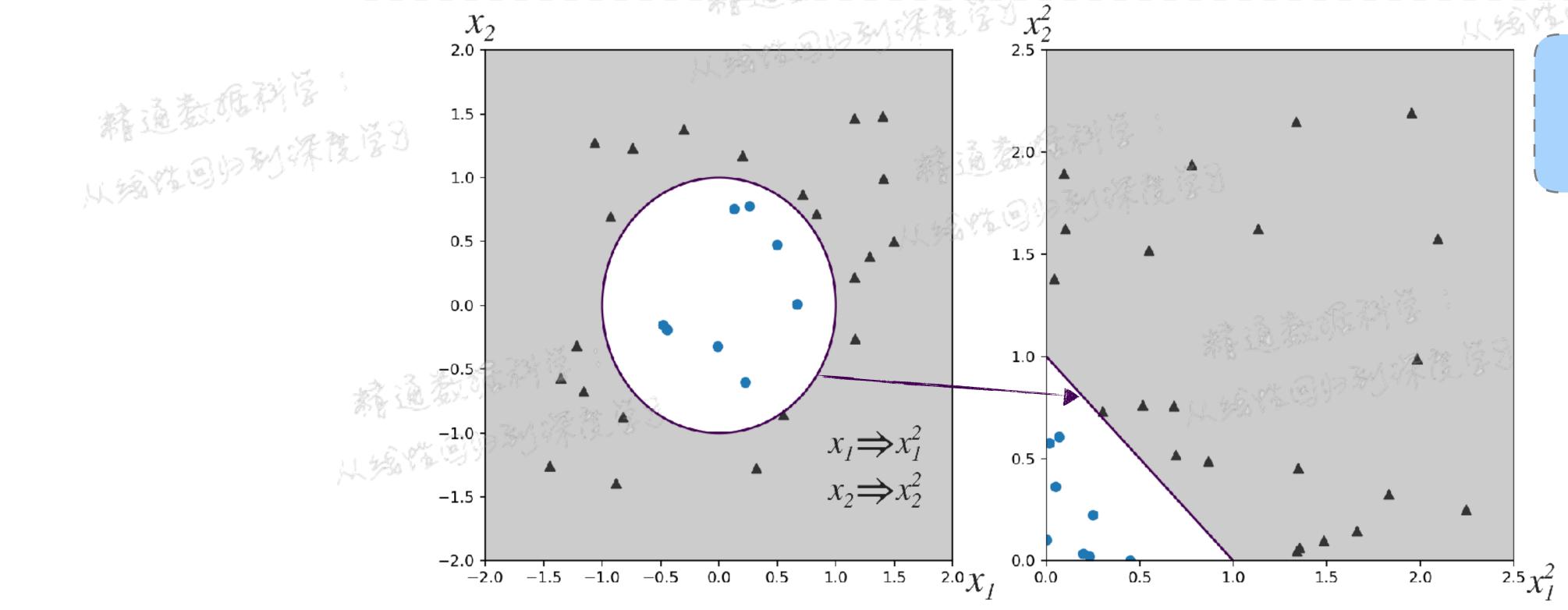
空间变换



### 从非线性到线性

空间变换





- · 变换函数很难定义
- ·运算复杂程度高

核函数解决的问题

### 目录

精通数据科学 了NE 从非线性到线性

空间整换 从编馆回的别深度管

精通数据科学。 从绝路的多处深度管

> 精通数据科学 从绝对回的粉珠度管

精通数据科学 从线性回归到深度管别

TWO 拉格朗日对偶 

精通数据科学 从细胞回的多种深度管 THREE 支持向量。海外深度等

SVM名字来源

什么是拉格朗日对偶

伟大的法国数学家



Joseph Louis de Lagrange

精通数据和资





从绝路的多别深度管

解决约束条件下的最优化问题

定义原始问题

SVM模型对应的最优化问题

最优化问题

$$\min \frac{1}{2} \parallel \mathbf{w} \parallel^2 + C \sum_{i} \xi_i$$

限制条件 
$$y_i(\mathbf{w} \cdot \mathbf{X}_i + c) \ge 1 - \xi_i$$
,  $\xi_i \ge 0$ 



### 模型的预测损失(hinge loss)

$$\min \left| \frac{1}{2} \parallel \mathbf{w} \parallel^2 \right| + C \sum_{i} \max(0, 1 - y_i(\mathbf{w} \cdot \mathbf{X}_i + c))$$

无限制条件的最优化问题

 $\min f(\theta)$ 

$$g_i(\theta) \le 0, \quad i = 1,...,k$$

$$h_i(\theta) = 0, \quad i = 1, ..., l$$

原始问题

primal optimization problem



$$\min_{\theta} P(\theta) = \min_{\theta} \max_{\alpha \ge 0, \beta} L(\theta, \alpha, \beta)$$

### 定义拉格朗日函数

$$L(\theta, \alpha, \beta) = f(\theta) + \sum_{i} \alpha_{i} g_{i}(\theta) + \sum_{i} \beta_{i} h_{i}(\theta)$$

定义对偶问题

### 两个问题在一定条件下等价

两个问题的最优解一定存在,且满足KKT条件

$$\min f(\theta)$$

$$g_i(\theta) \le 0, \quad i = 1,...,k$$

$$h_i(\theta) = 0, \quad i = 1,...,l$$

$$L(\theta, \alpha, \beta) = f(\theta) + \sum_{i} \alpha_{i} g_{i}(\theta) + \sum_{i} \beta_{i} h_{i}(\theta)$$

原始问题

primal optimization problem



$$\min_{\theta} P(\theta) = \min_{\theta} \max_{\alpha \ge 0, \beta} L(\theta, \alpha, \beta)$$

对偶问题 dual optimization problem

$$D(\alpha, \beta) = \min_{\theta} L(\theta, \alpha, \beta)$$

$$\max_{\alpha \ge 0, \beta} D(\alpha, \beta) = \max_{\alpha \ge 0, \beta} \min_{\theta} L(\theta, \alpha, \beta)$$

$$\alpha \ge 0, \beta$$

$$L(\hat{\theta}, \hat{\alpha}, \hat{\beta}) = \max_{\alpha \ge 0, \beta} \min_{\theta} L(\theta, \alpha, \beta) = \min_{\theta} \max_{\alpha \ge 0, \beta} L(\theta, \alpha, \beta)$$

$$\frac{\frac{\partial L}{\partial \theta}(\widehat{\theta}, \widehat{\alpha}, \widehat{\beta})}{\frac{\partial L}{\partial \beta}(\widehat{\theta}, \widehat{\alpha}, \widehat{\beta})} = 0$$

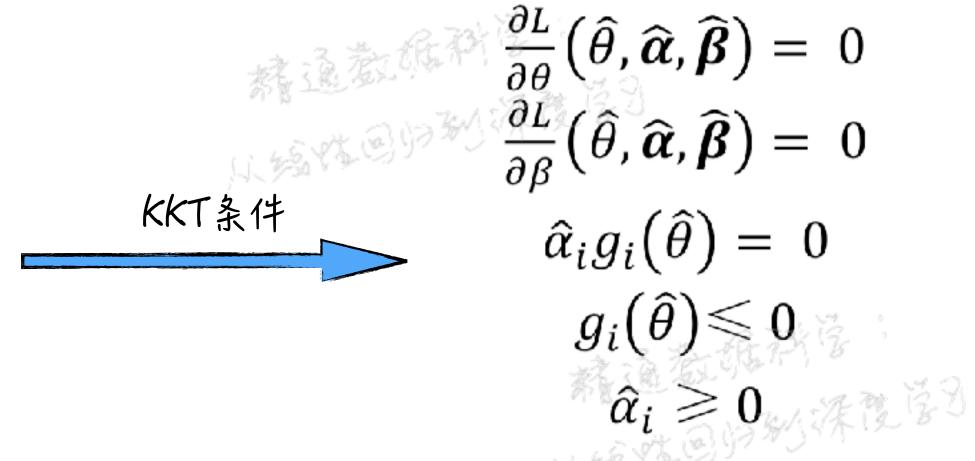
$$\widehat{\alpha}_{i}g_{i}(\widehat{\theta}) = 0$$

$$g_{i}(\widehat{\theta}) \leq 0$$

$$\widehat{\alpha}_{i} \geq 0$$

SVM的对偶问题

$$L(\hat{\theta}, \hat{\alpha}, \hat{\beta}) = \max_{\alpha \ge 0, \beta} \min_{\theta} L(\theta, \alpha, \beta) = \min_{\theta} \max_{\alpha \ge 0, \beta} L(\theta, \alpha, \beta)$$
$$L(\theta, \alpha, \beta) = f(\theta) + \sum_{i} \alpha_{i} g_{i}(\theta) + \sum_{i} \beta_{i} h_{i}(\theta)$$



### SVM的拉格朗日函数

$$L(w, c, \xi, \alpha, \gamma) = \frac{1}{2} \| \mathbf{w} \|^2 + C \sum_{i} \xi_i - \sum_{i} \alpha_i [y_i(w \cdot X_i + c) - 1 + \xi_i] - \sum_{i} \gamma_i \xi_i$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow \hat{w} = \sum_{i} \hat{\alpha}_{i} y_{i} X_{i}$$

$$\frac{\partial L}{\partial \xi} = 0 \Rightarrow C - \hat{\alpha}_{i} - \hat{\gamma}_{i} = 0$$

$$\frac{\partial L}{\partial c} = 0 \Rightarrow \sum_{i} y_{i} \hat{\alpha}_{i} = 0$$

$$D(\alpha, \gamma) = \min_{\theta} L(\theta, \alpha, \beta)$$

$$D(\alpha, \gamma) = \sum_{i} \alpha_{i} - \frac{1}{2} \sum_{i,j} y_{i} y_{j} \alpha_{i} \alpha_{j} (X_{i} \cdot X_{j})$$

$$0 \le \alpha_{i} \le C; \qquad \sum_{i} \alpha_{i} y_{i} = 0$$

SVM的对偶问题

Original SVM

参数估计公式

$$\min_{\mathbf{w},c} \frac{1}{2} ||\mathbf{w}||^2 + C \sum_{i} \xi_i$$
s.t.  $y_i(\mathbf{w} \cdot \mathbf{X}_i + c) \ge 1 - \xi_i$ 

$$\xi_i \ge 0, \forall i$$

预测公式

$$\hat{y}_j = \operatorname{sign}(\hat{\boldsymbol{w}} \cdot \boldsymbol{X_j} + \hat{c})$$

### Dual problem

参数估计公式

所以估计公式 
$$\max_{\alpha} \sum_{i} \alpha_{i} - \frac{1}{2} \sum_{i,j} y_{i} y_{j} \alpha_{i} \alpha_{j} (\boldsymbol{X_{i}} \cdot \boldsymbol{X_{j}})$$
  $s.t.$   $0 \leq \alpha_{i} \leq C$  
$$\sum_{i} \alpha_{i} y_{i} = 0, \forall i$$

$$D(\alpha, \gamma) = \sum_{i} \alpha_{i} - \frac{1}{2} \sum_{i,j} y_{i} y_{j} \alpha_{i} \alpha_{j} (X_{i} \cdot X_{j})$$

预测公式 
$$\hat{y}_j = \text{sign}(\sum_i \hat{\alpha}_i y_i (X_i \cdot X_j) + \hat{c})$$

$$0 \le \alpha_i \le C; \qquad \sum_i \alpha_i y_i = 0$$

### 目录

精通数据科学 了NE 从非线性到线性

空间整规查证 从编馆回的到深度管

精通数据科学 从绝对回的粉件度管的

精通数纸料学。

从绝路的的秘珠度管的

精通数据科学 从线性回归初深度管引

TVVO 拉格朗图对偶

KKT

精通激源和学: 从细胞的多种性 THREE 支持向量。如果使用的

SVM名字来源

# 支持向量

SVM名字来源

# Dual problem

$$\max_{\alpha} \sum_{i} \alpha_{i} - \frac{1}{2} \sum_{i,j} y_{i} y_{j} \alpha_{i} \alpha_{j} (X_{i} \cdot X_{j})$$

$$s.t. \quad 0 \leq \alpha_{i} \leq C$$

$$\sum_{i} \alpha_{i} y_{i} = 0, \forall i$$

$$\hat{y}_j = \operatorname{sign}(\sum_i \hat{\alpha}_i y_i (\boldsymbol{X}_i \cdot \boldsymbol{X}_j) + \hat{c})$$



- · 用内积表示"点"与"点"之间的相似度
- ·用内积作为权重去平均被预测量y,得到预测结果

# 支持向量

SVM名字来源

$$\min_{\substack{\mathbf{w},c \ \mathbf{w}}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i} \xi_i \\
\mathbf{s.t.} \quad y_i(\mathbf{w} \cdot \mathbf{X}_i + c) \ge 1 - \xi_i \\
\xi_i \ge 0, \forall i$$

$$\frac{\partial L}{\partial \theta} (\widehat{\theta}, \widehat{\alpha}, \widehat{\beta}) = 0 \\
\frac{\partial L}{\partial \beta} (\widehat{\theta}, \widehat{\alpha}, \widehat{\beta}) = 0 \\
\widehat{\alpha}_i g_i(\widehat{\theta}) = 0 \\
g_i(\widehat{\theta}) \le 0$$

$$\frac{\partial L}{\partial \theta}(\widehat{\theta}, \widehat{\alpha}, \widehat{\beta}) = 0$$

$$\frac{\partial L}{\partial L}(\widehat{\theta}, \widehat{\alpha}, \widehat{\beta}) = 0$$

$$\widehat{\alpha}_{i}g_{i}(\widehat{\theta}) = 0$$

$$g_{i}(\widehat{\theta}) \leq 0$$

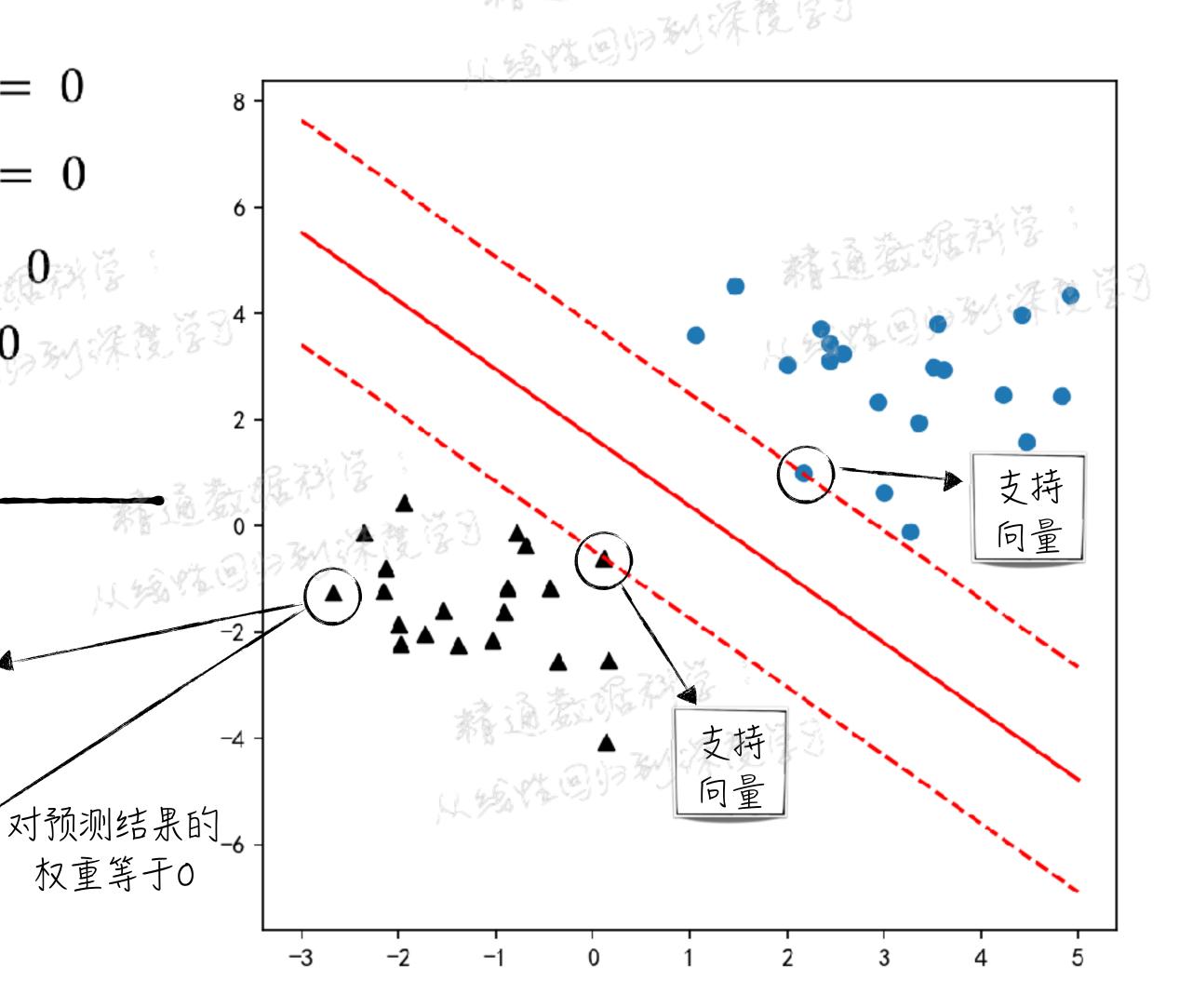
$$\widehat{\alpha}_{i} \geq 0$$
2

 $g_i < 0$ 

 $\hat{\alpha}_i = 0$ 

$$g_i = 1 - y_i(w \cdot X_i + c) - \xi_i$$
$$\hat{\alpha}_i g_i = 0$$

$$\hat{y}_j = \text{sign}(\sum_i \hat{\alpha}_i y_i (X_i \cdot X_j) + \hat{c})$$



精通数据科学。 从验验证到的秘证不改资

# THANKSOUS

務通数据科学 从给您回归和深度管

村通教师和强。

精通数据科学。 从绝路的多处深度管

精通数据科学

精通数派科学 从给你回的秘况