如何处理线性不可分的问题

SVM的数学进化

小胖

目录

加入预测损失项

误分类的损失项

对于线性可分，我们有如下的数学表达：

对于类别1，令$y = 1$；
对于类别0，令$y = -1$

$$\min \frac{1}{2}\|w\|^2$$

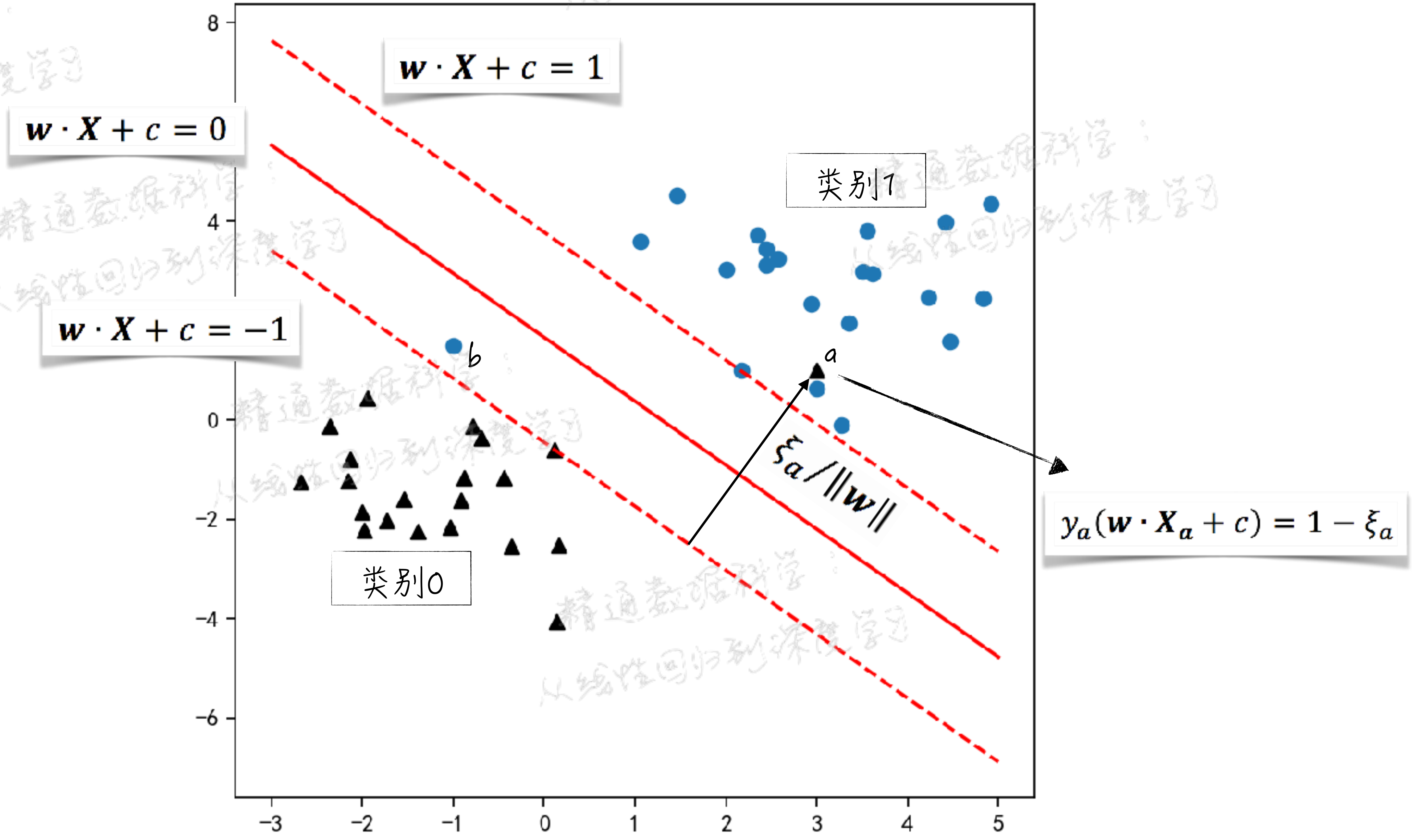
限制条件 $y_i(w \cdot X_i + c) \geqslant 1$

对于线性不可分，适当放宽限制条件

$$y_i(\mathbf{w} \cdot \mathbf{X}_i + c) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

衡量模型违反自身分类原则的程度

数据线性不可分时，加入误分类的损失



加入预测损失项

误分类的损失项

$$\min \frac{1}{2} \parallel \mathbf{w} \parallel^2$$

衡量模型违反自身分类原则的程度

$$y_i(\mathbf{w} \cdot \mathbf{X}_i + c) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

越小越好

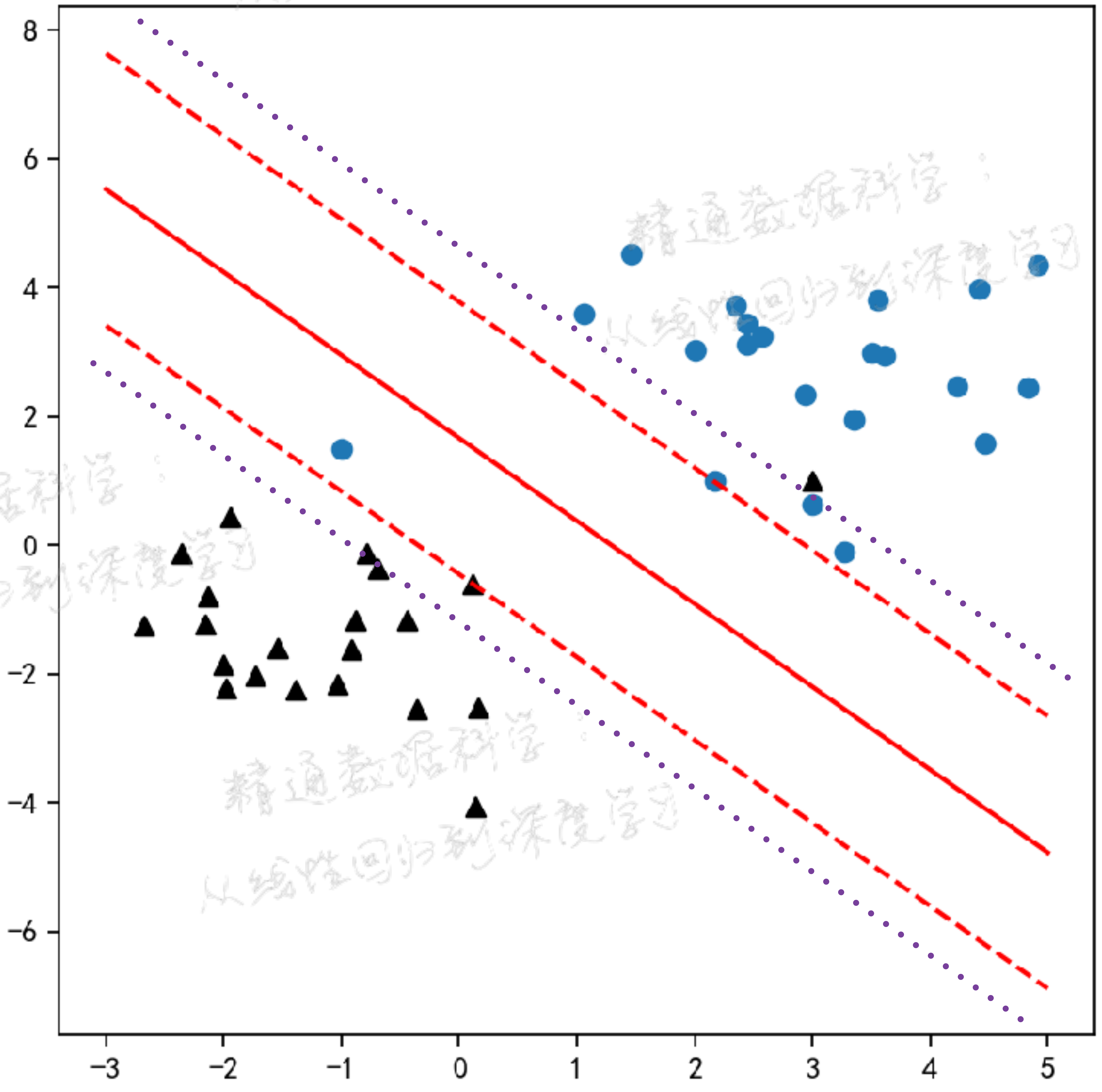
越小越好

$$\sum_i \xi_i \longleftrightarrow \frac{1}{2} \parallel \mathbf{w} \parallel^2$$

此消彼长

$$\min \tfrac{1}{2}\|\boldsymbol{w}\|^2 + C \textstyle\sum_i \xi_i$$

$$y_i(\boldsymbol{w} \cdot \boldsymbol{X}_i + c) \geqslant 1 - \xi_i;\ \xi_i \geqslant 0$$



目录

损失函数与惩罚项

定义SVM的损失函数

$$\min \tfrac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_i \xi_i$$

$$y_i(\boldsymbol{w}\cdot\boldsymbol{X}_i + c) \geqslant 1 - \xi_i;\ \xi_i \geqslant 0$$

$$\xi_i \geq 1 - y_i(\mathbf{w}\cdot\mathbf{X}_i + c), \quad \xi_i \geq 0$$

$$\xi_i \geq \max(0, \quad 1 - y_i(\mathbf{w}\cdot\mathbf{X}_i + c))$$

$$\rightarrow \quad \xi_i = \max(0, \quad 1 - y_i(\mathbf{w}\cdot\mathbf{X}_i + c))$$

惩罚项

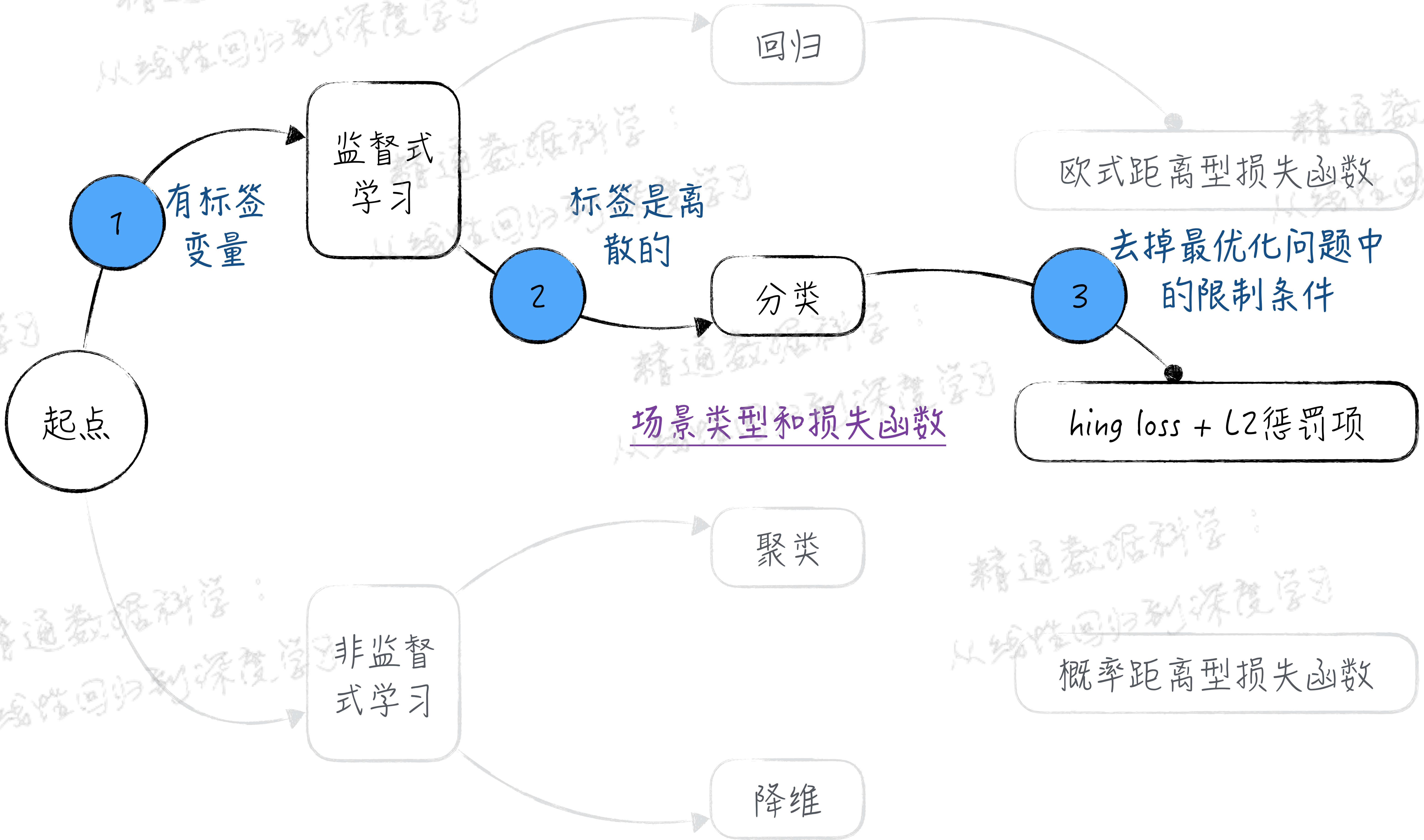
模型的预测损失(hinge loss)

$$\min \boxed{\frac{1}{2}\|\mathbf{w}\|^2} + \boxed{C\sum_i \max(0, 1 - y_i(\mathbf{w}\cdot\mathbf{X}_i + c))}$$

无限制条件的最优化问题

损失函数与惩罚项

定义SVM的损失函数

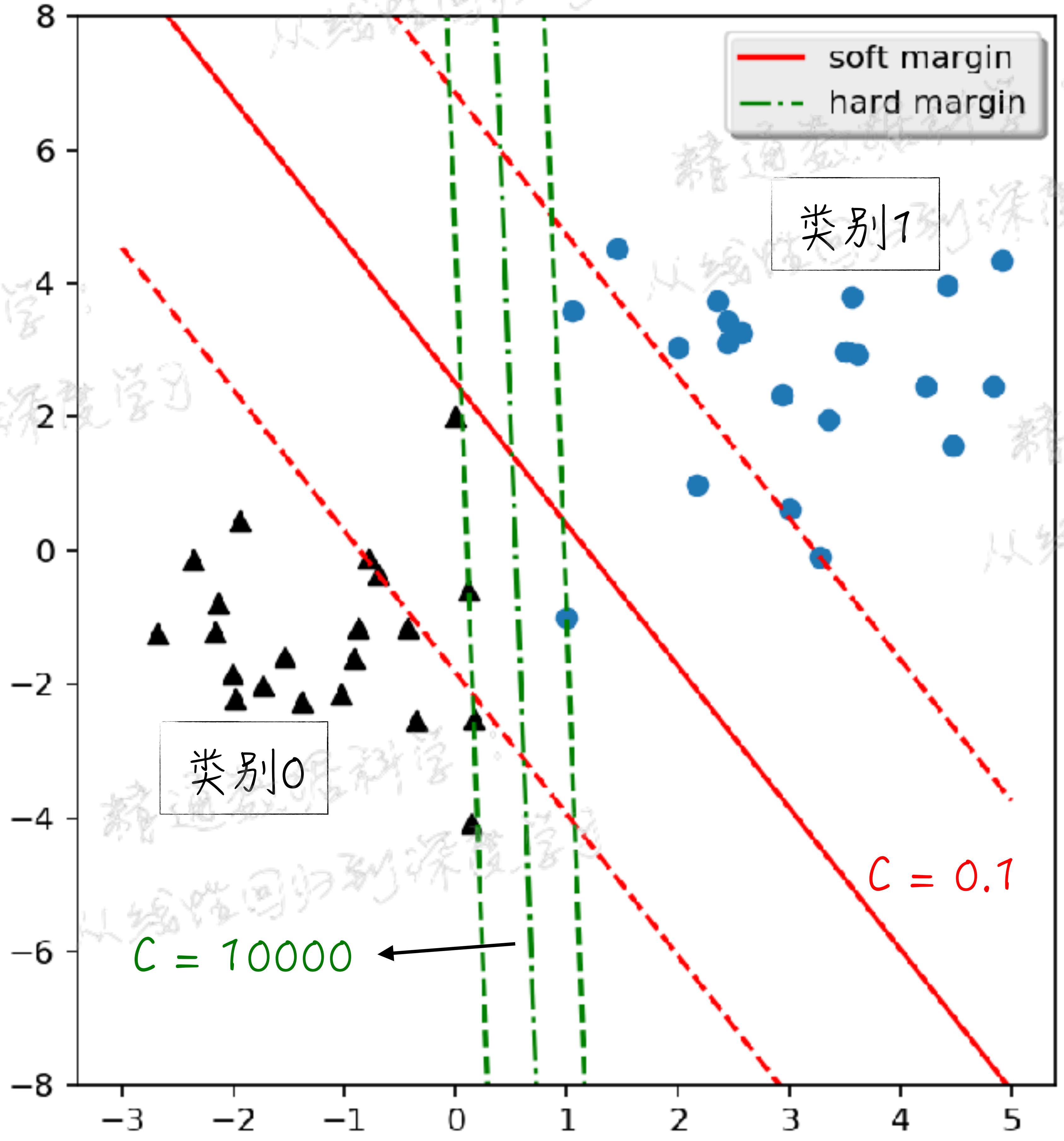


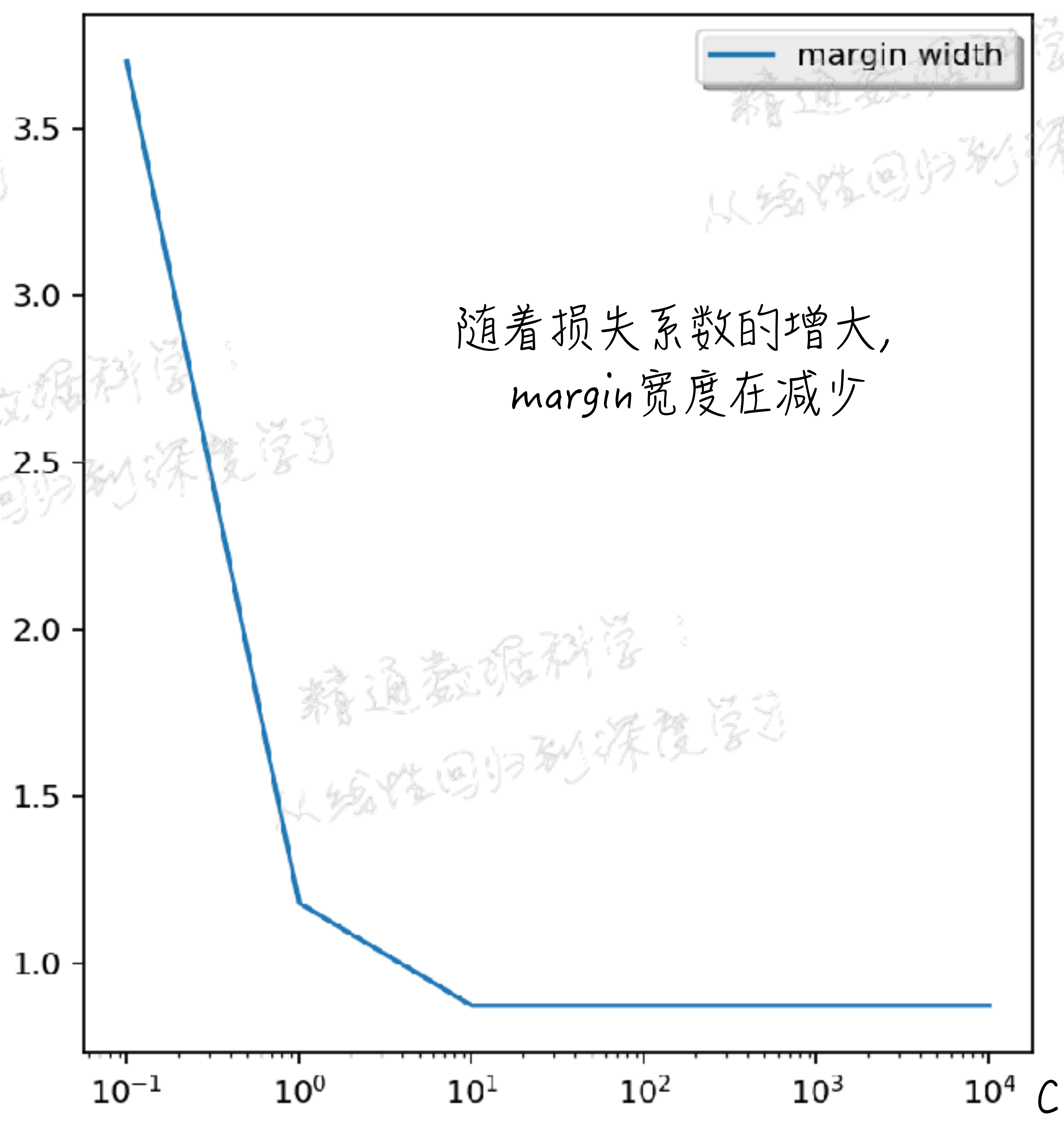
目录

Hard margin v.s. Soft margin

损失系数

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

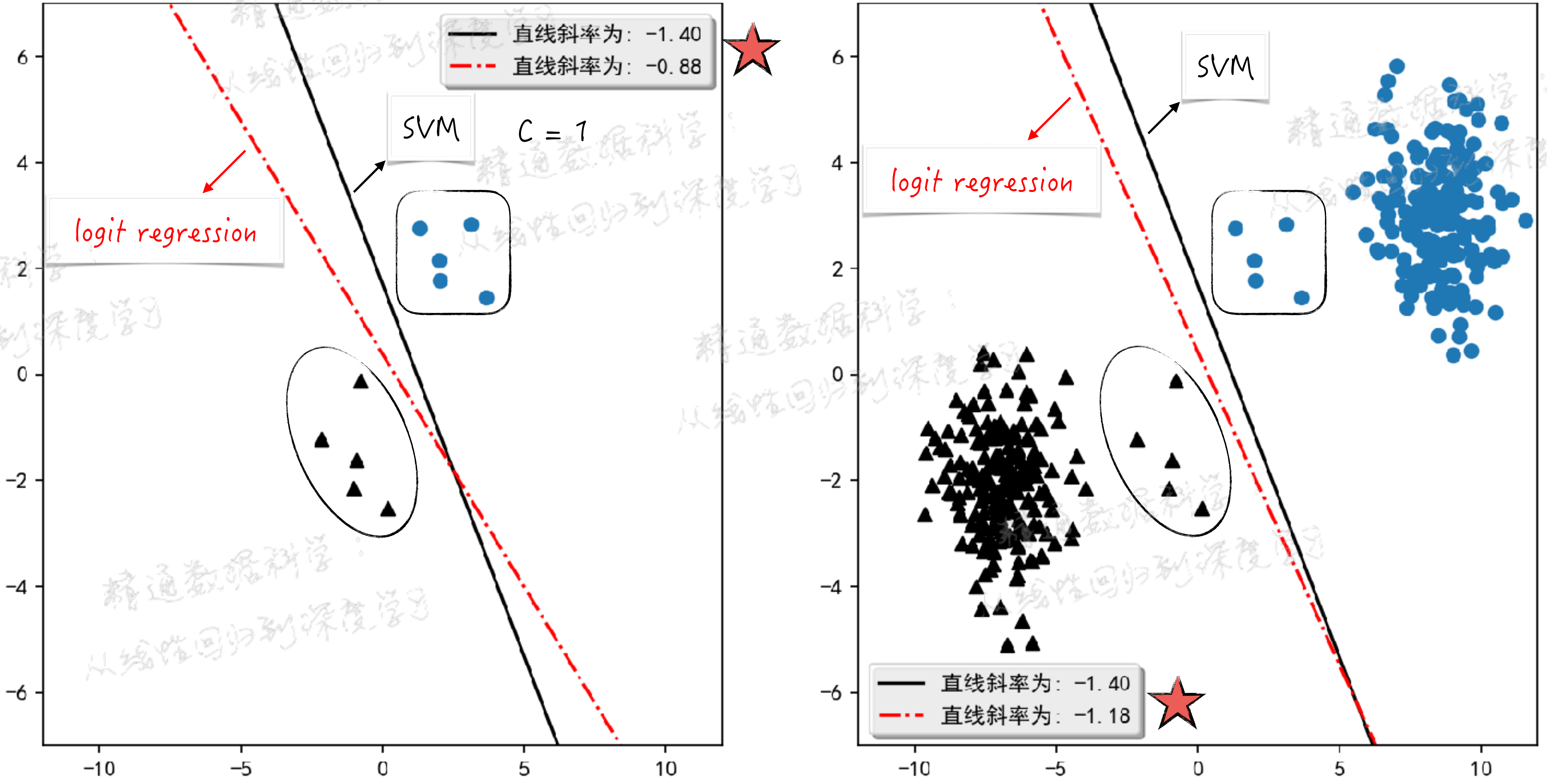




Hard margin v.s. Soft margin

SVM隐藏的模型假设

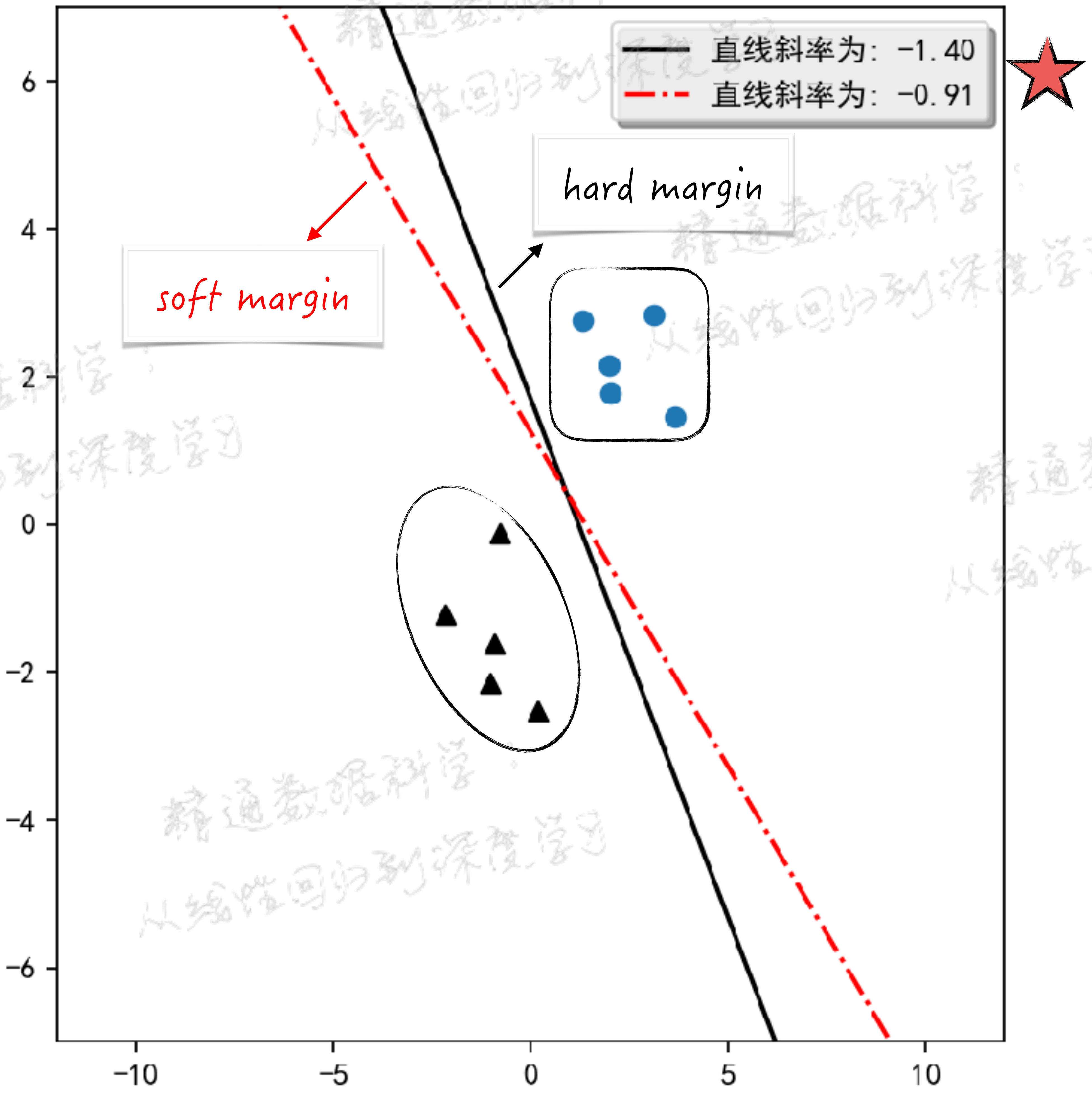
SVM“只用”了框中的数据；
logit regression使用了所有数据

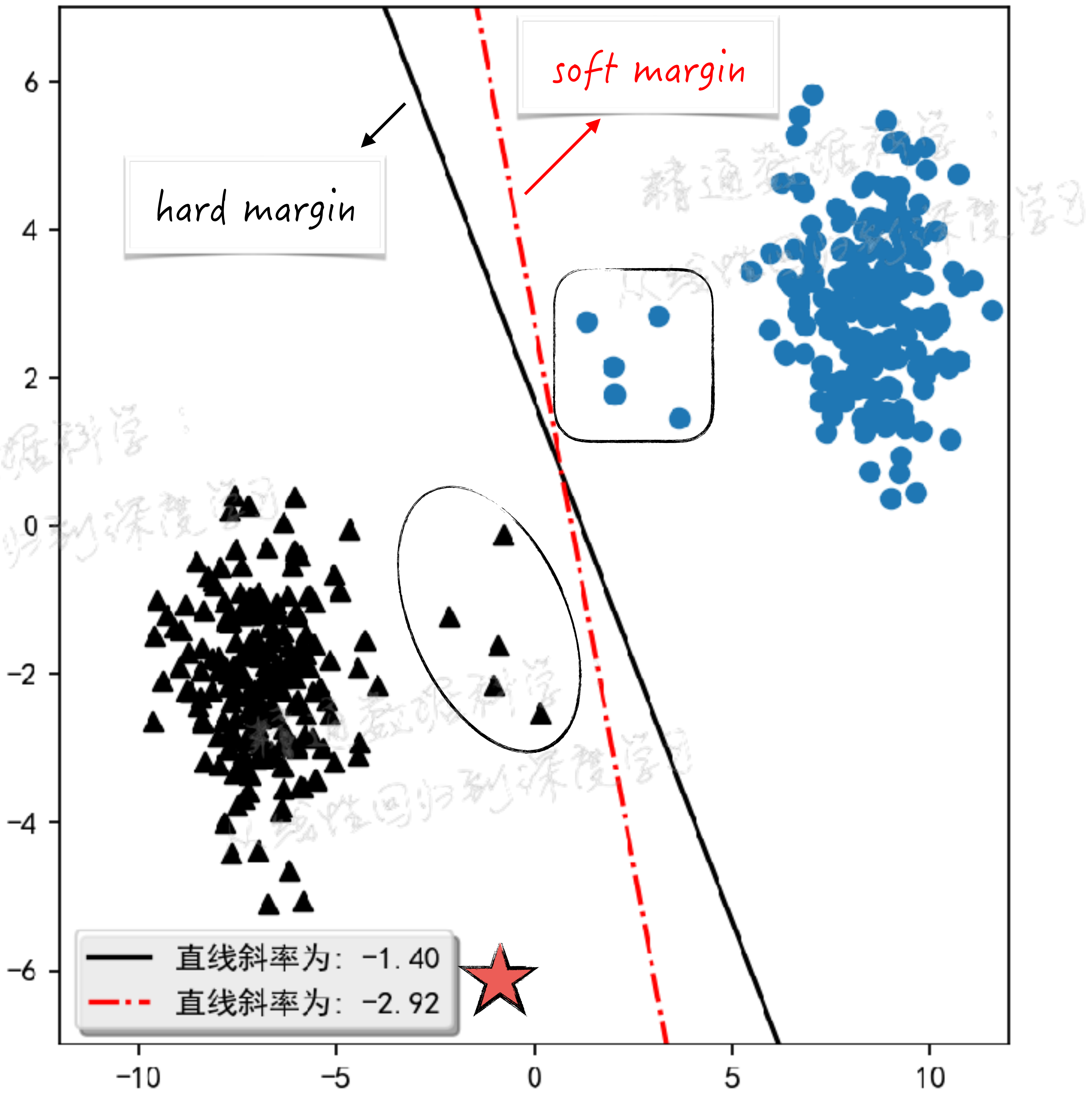


Hard margin v.s. Soft margin

SVM隐藏的模型假设

soft margin使用了所有数据；
hard margin"只用"了框中的数据





THANK YOU