什么是SVM?

万能分类器

小胖

目录

前情回顾与后续预告

数据模型与算法模型

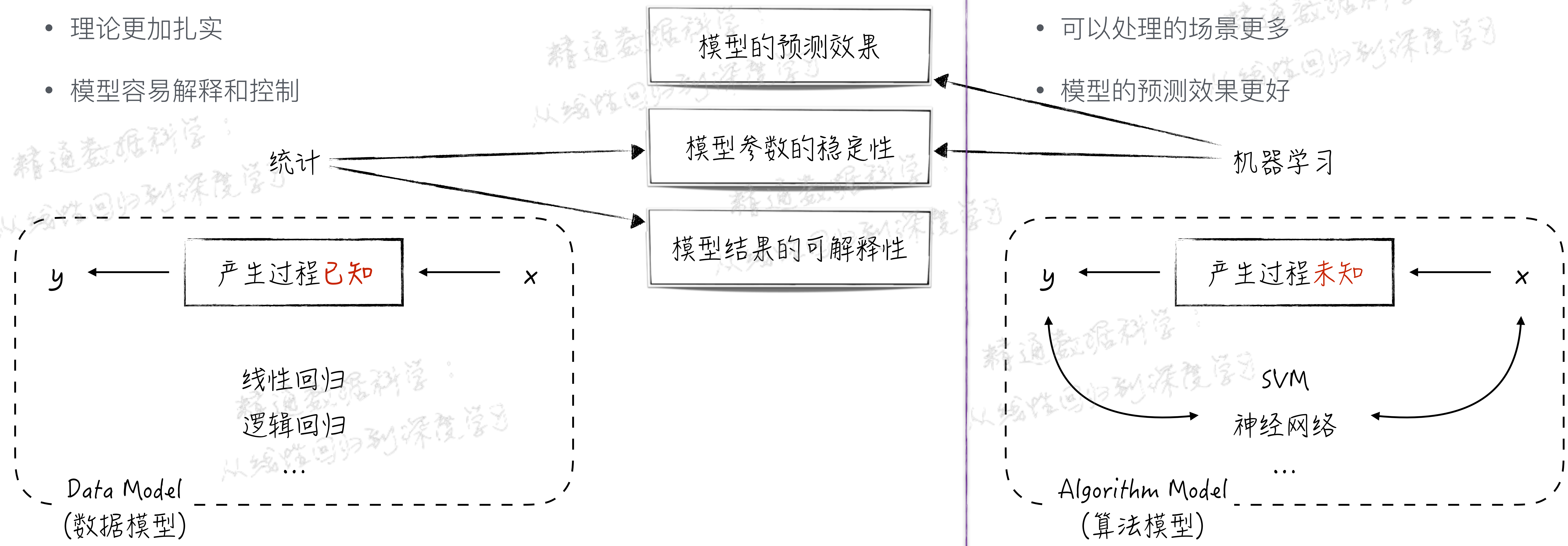
后续重点

注重用**数学的方法**来搭建模型

- 理论更加扎实
- 模型容易解释和控制

注重用**工程的方法**来搭建模型

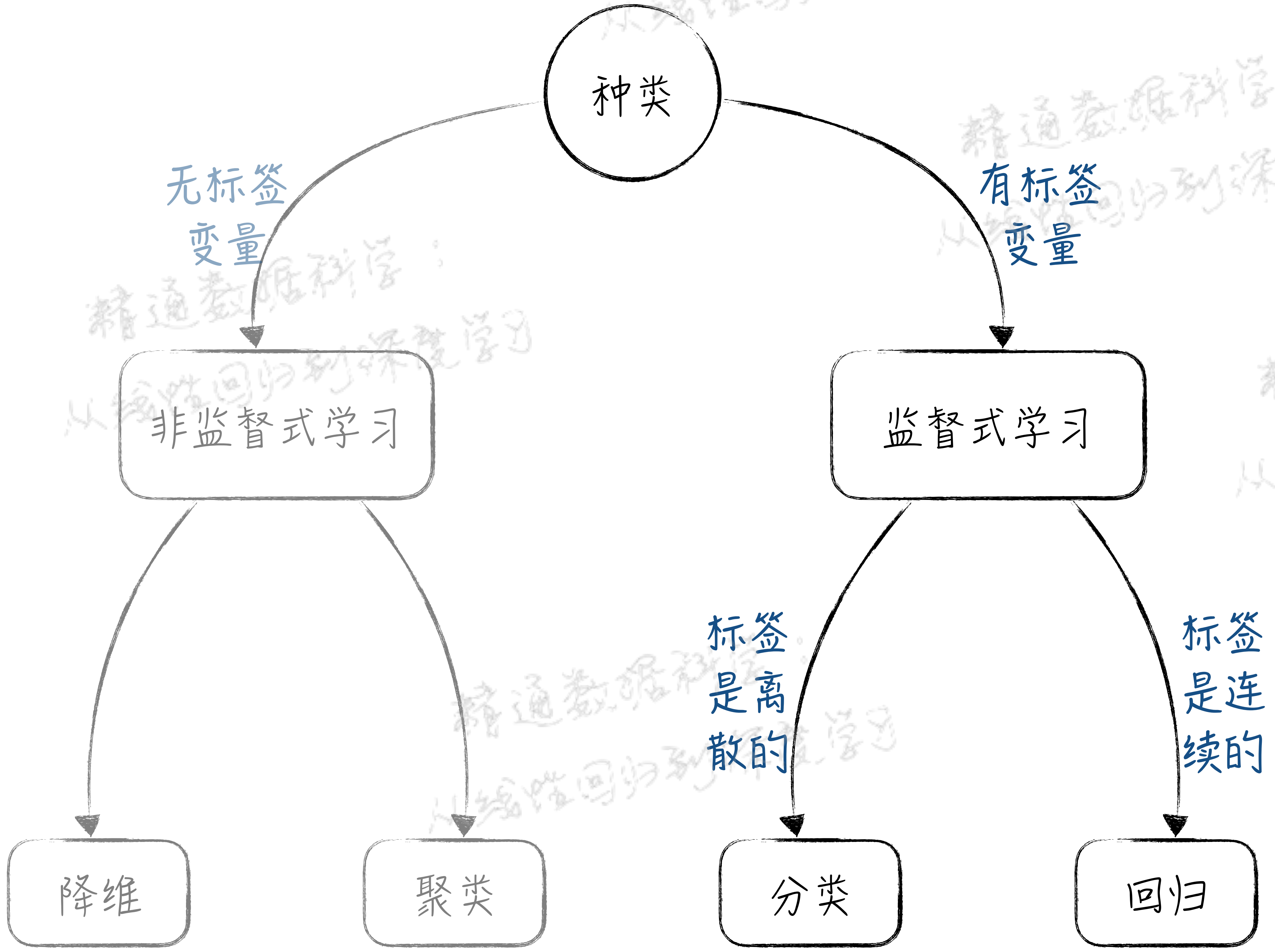
- 可以处理的场景更多
- 模型的预测效果更好



前情回顾与后续预告

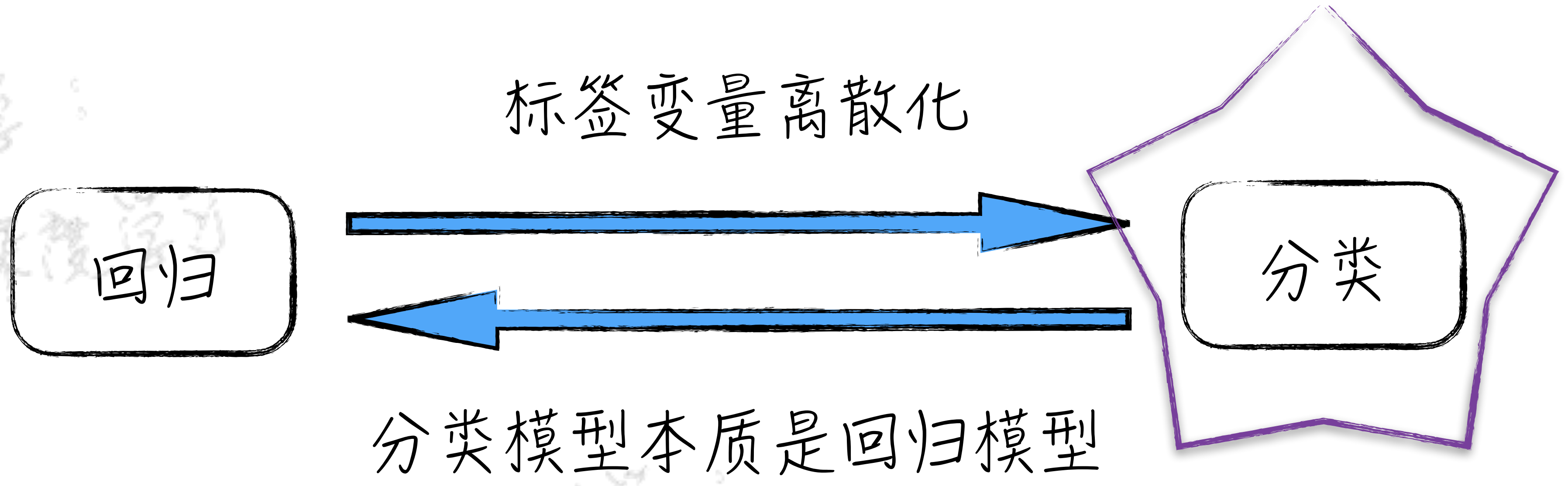
回归与分类

模型的种类



分类与回归的相互转换

后续重点



目录

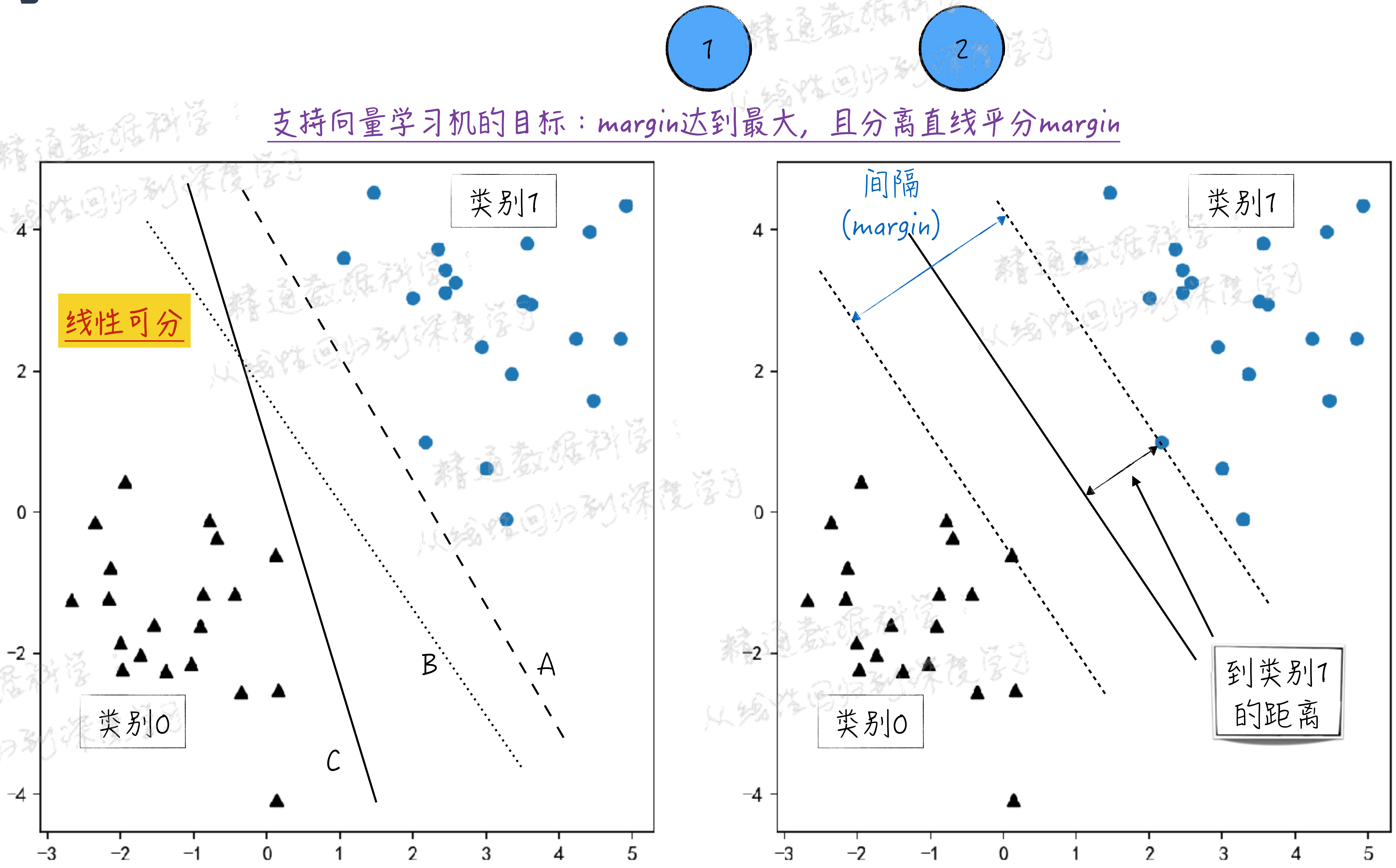
一个直观的例子

什么是SVM?

SVM的全称是支持向量学习机（Support Vector Machine）

SVM是一个解决分类问题的模型，被不少人誉为“万能分类器”

支持向量学习机的目标：margin达到最大，且分离直线平分margin



一个直观的例子

向量内积回顾

终点坐标 - 起点坐标

$$\mathbf{a} = (a_1, a_2)$$

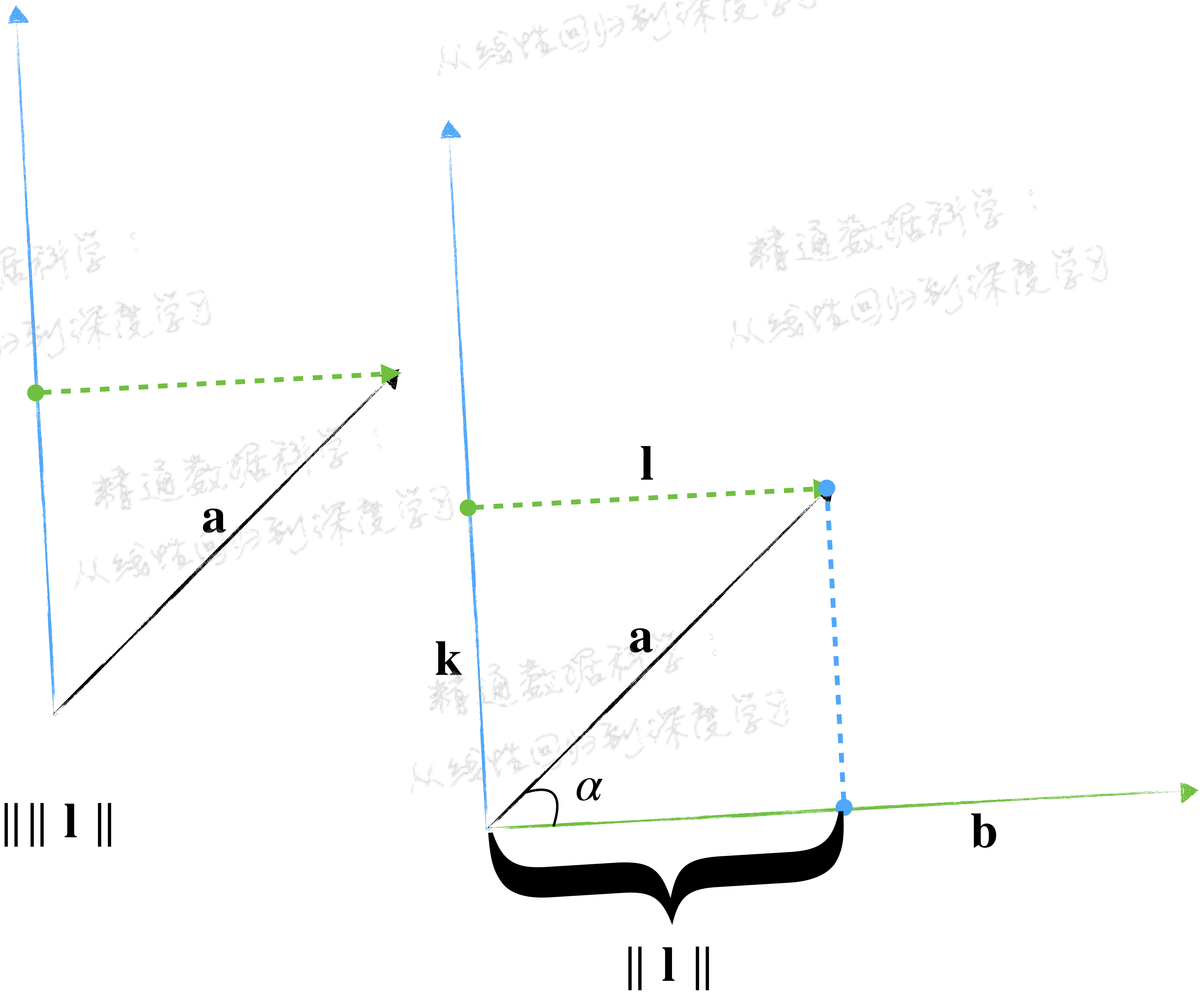
$$\mathbf{b} = (b_1, b_2)$$

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2$$

$$\mathbf{a} \cdot \mathbf{b} = \| \mathbf{b} \| \| \mathbf{a} \| \, cos\alpha$$

$$\mathbf{a} \cdot \mathbf{b} = \| \mathbf{b} \| \| \mathbf{l} \|$$

$$\mathbf{a} \cdot \mathbf{b} = (\mathbf{k} + \mathbf{l}) \cdot \mathbf{b} = \mathbf{k} \cdot \mathbf{b} + \mathbf{l} \cdot \mathbf{b} = \| \mathbf{b} \| \| \mathbf{l} \|$$



一个直观的例子

用数学理解直观

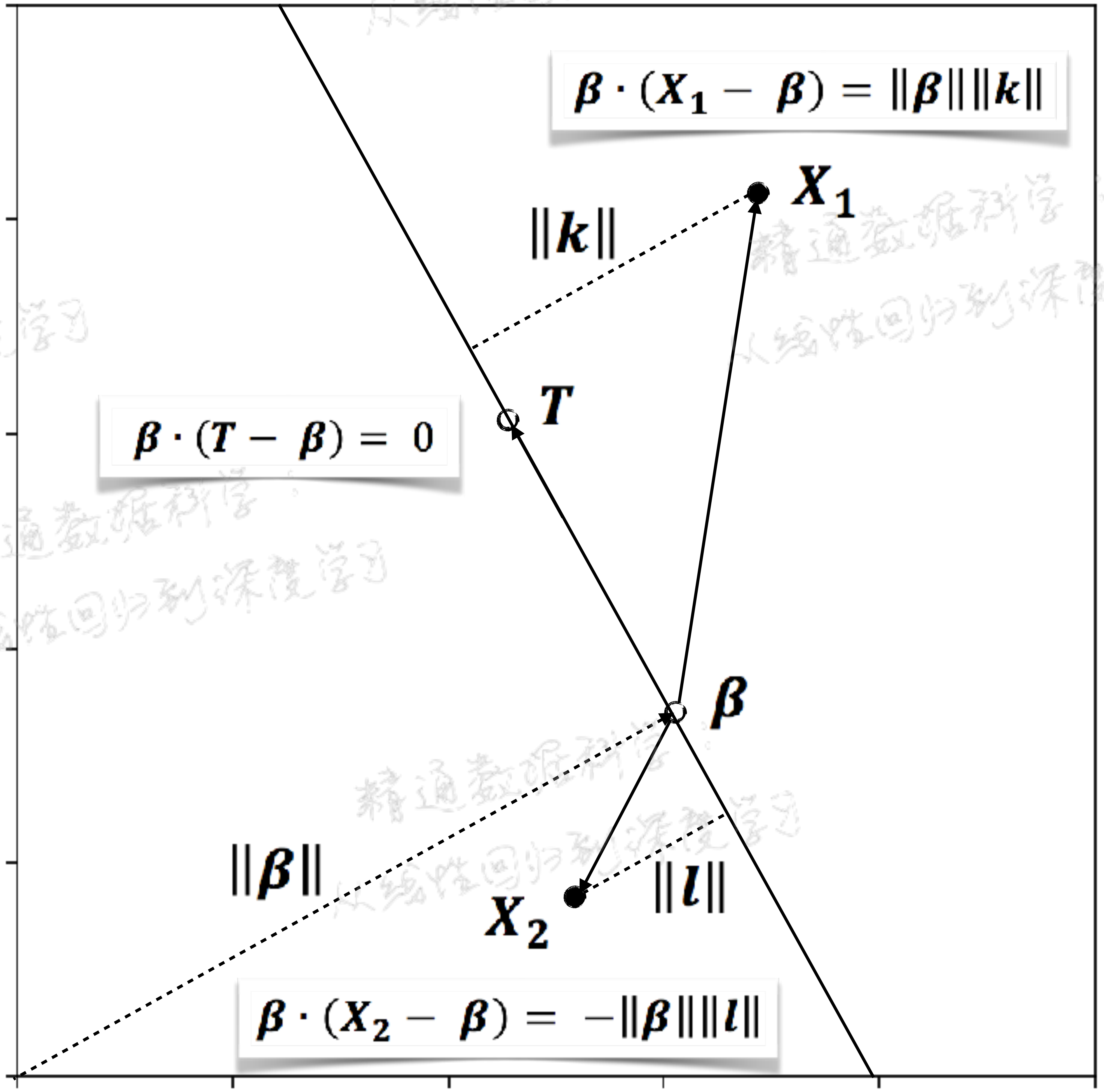
对于向量beta，定义如下的线性函数

$$f(\mathbf{x}) = \beta \cdot (\mathbf{x} - \beta)$$

- f(x) = 0表示垂直于beta的直线
- f(x)的绝对值与点x到上述直线的距离成正比

这个结论可以推广到任意线性函数以及高维空间

将几何直观转换为代数

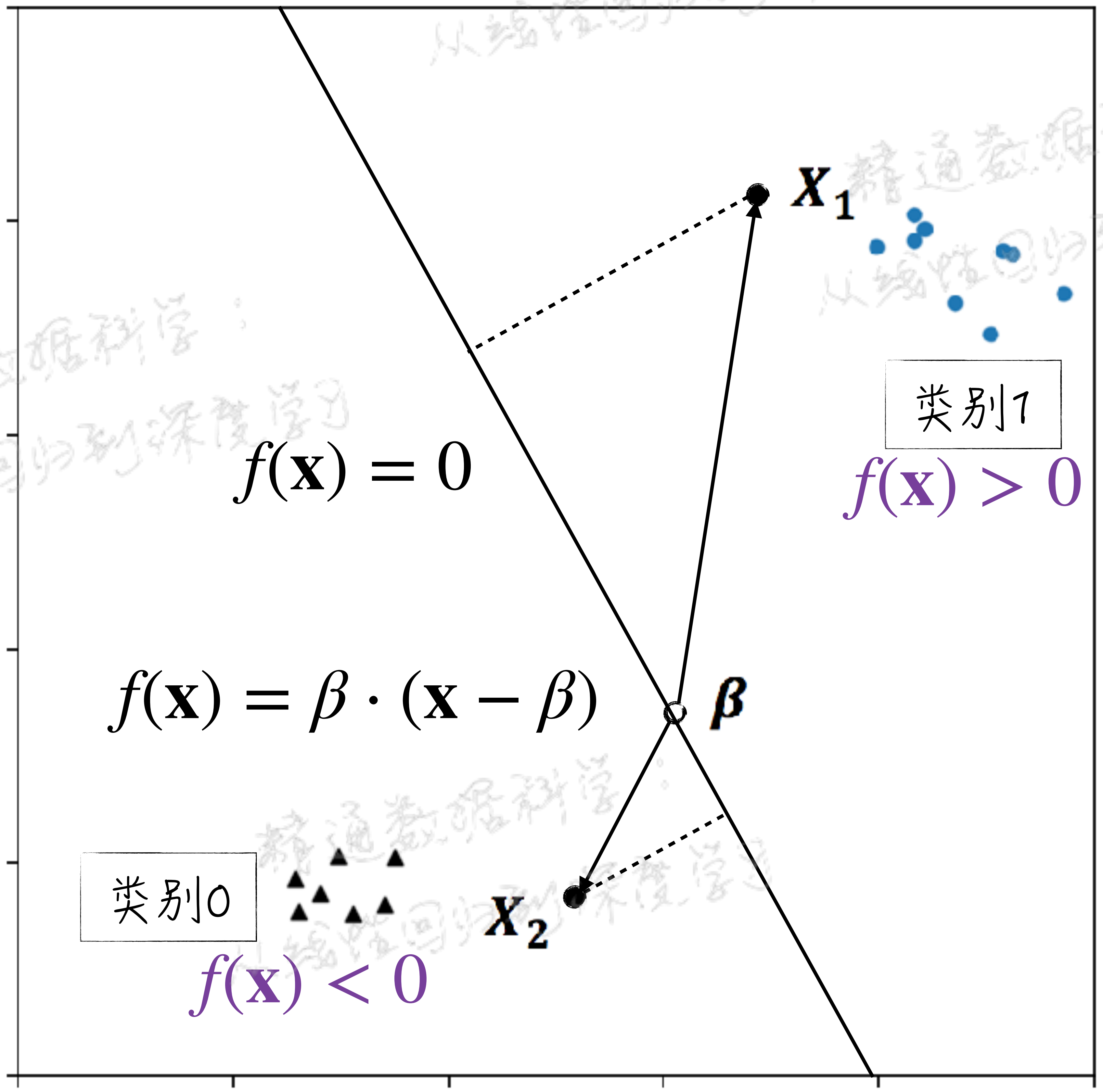


一个直观的例子

SVM与逻辑回归

殊途同归

SVM的预测逻辑



逻辑回归的预测逻辑

通常Alpha等于0.5

$$\hat{y}_i = \begin{cases} 1, & \frac{1}{1+e^{-X_i\beta}} > \alpha \\ 0, & else \end{cases}$$

g(X)越大，y=1的概率越大

$g(\mathbf{X}) > 0$ 预测结果为y =1

$$g(\mathbf{X}) = \mathbf{X}\beta$$

$g(\mathbf{X}) < 0$ 预测结果为y =0

g(X)越小，y=0的概率越大

目录

从几何直观到数学表达

SVM的数学表达

将SVM的原则翻译成数学

$$\max 2\|\boldsymbol{l}\|$$

任一$\boldsymbol{X}$属于类别 1，$\boldsymbol{\beta} \cdot \boldsymbol{X} + b \geqslant \|\boldsymbol{l}\|\|\boldsymbol{\beta}\|$

任一$\boldsymbol{X}$属于类别 0，$\boldsymbol{\beta} \cdot \boldsymbol{X} + b \leqslant -\|\boldsymbol{l}\|\|\boldsymbol{\beta}\|$

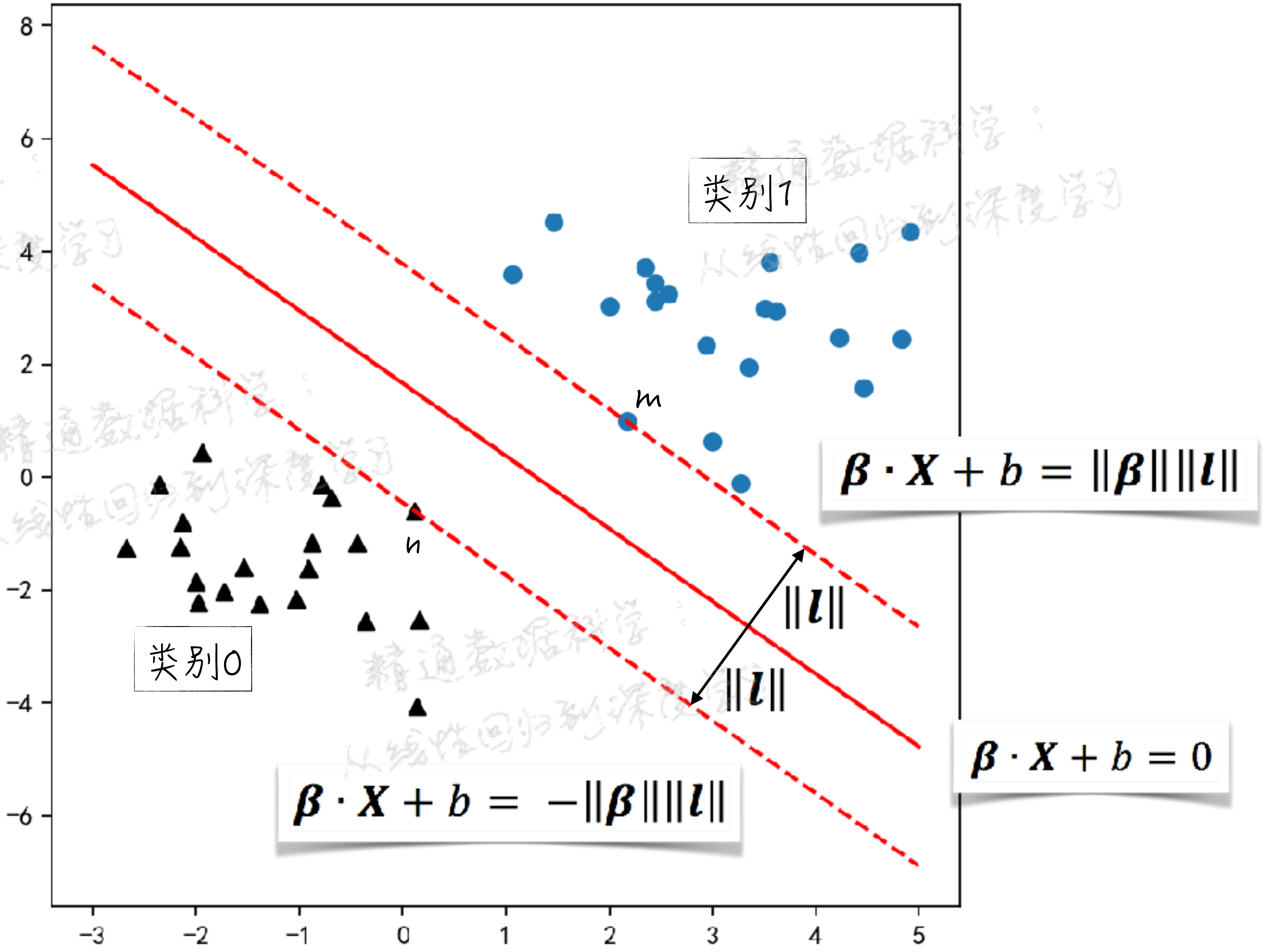
令 $w = \dfrac{\beta}{\|\mathbf{l}\|\|\beta\|}$

$$\max \frac{2}{\|\boldsymbol{w}\|}$$

任一$\boldsymbol{X}$属于类别 1，$\boldsymbol{w} \cdot \boldsymbol{X} + c \geqslant 1$

任一$\boldsymbol{X}$属于类别 0，$\boldsymbol{w} \cdot \boldsymbol{X} + c \leqslant -1$

线性可分



从几何直观到数学表达

带限制条件的最优化问题

$$\max \frac{2}{\|\boldsymbol{w}\|}$$

任一$\boldsymbol{X}$属于类别 1，$\boldsymbol{w} \cdot \boldsymbol{X} + c \geqslant 1$

任一$\boldsymbol{X}$属于类别 0，$\boldsymbol{w} \cdot \boldsymbol{X} + c \leqslant -1$

对于类别1，令$y = 1$；

对于类别0，令$y = -1$

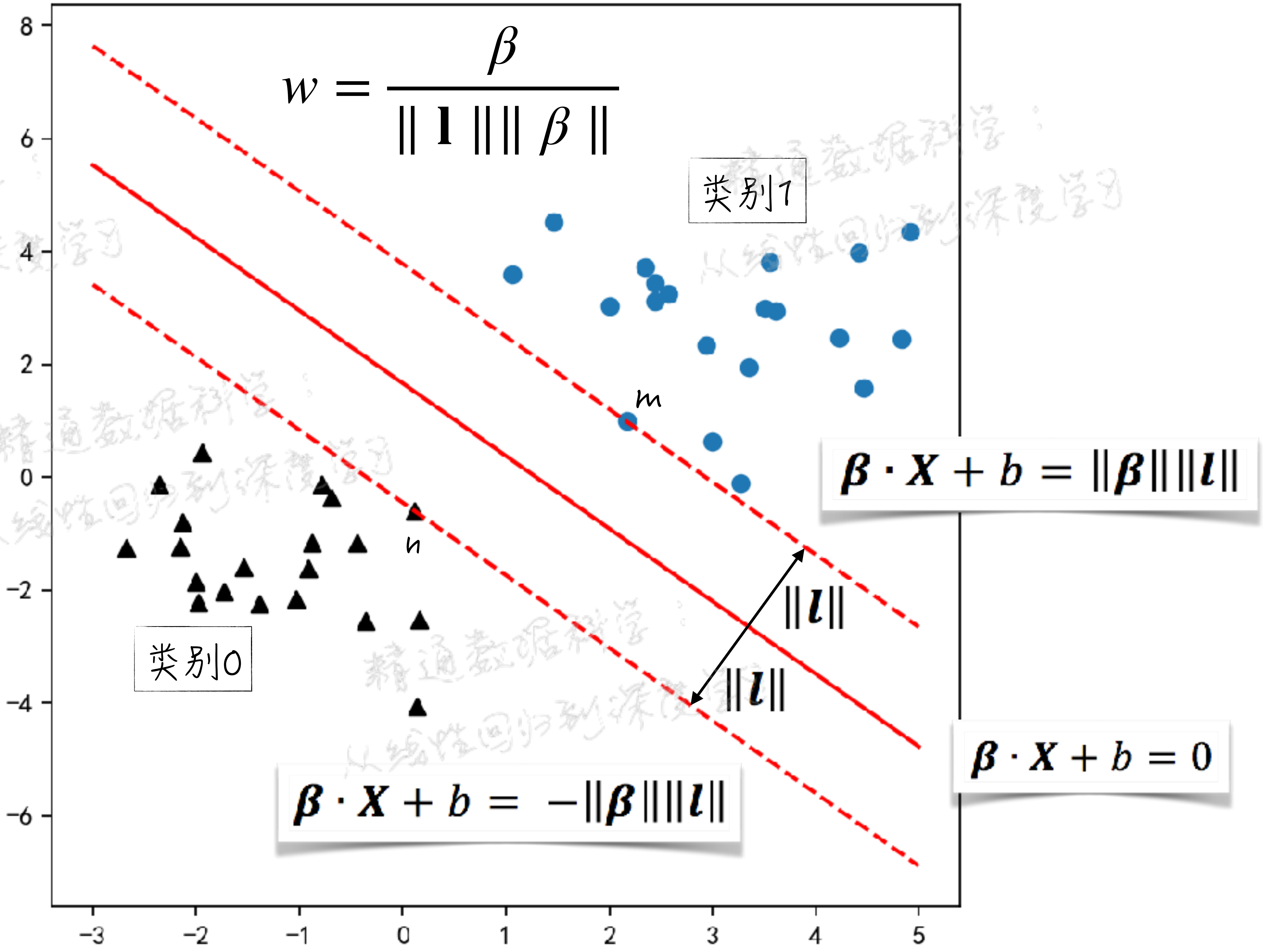
最优化问题 $\min \frac{1}{2}\|\boldsymbol{w}\|^2$

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

SVM预测公式：

$$\hat{y}_i = sign(\hat{w} \cdot \mathbf{X}_i + \hat{c})$$

线性可分



THANK YOU