Python 3 玩儿转机器学习

liuyubobobo

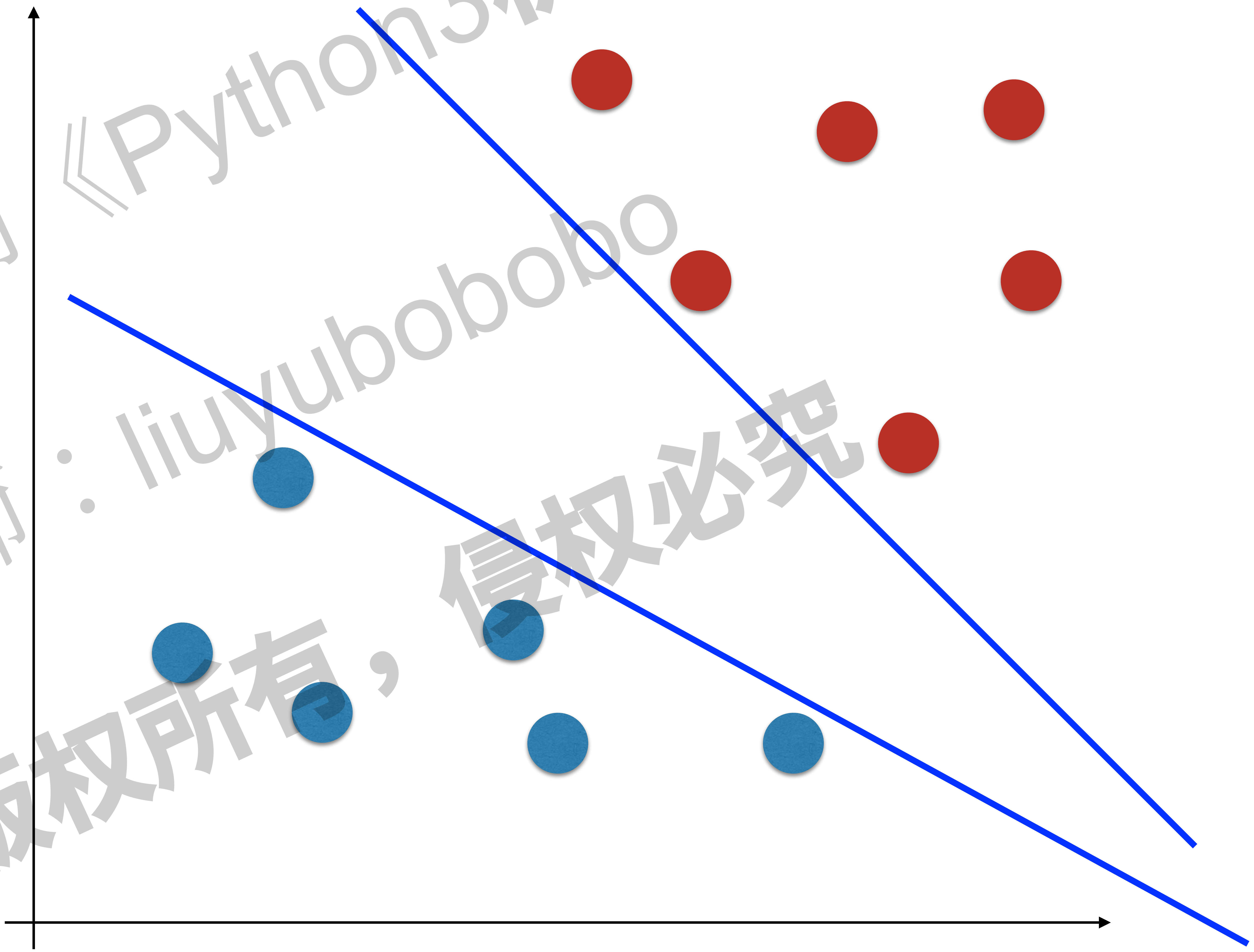
支撑向量机 SVM

什么是支撑向量机?

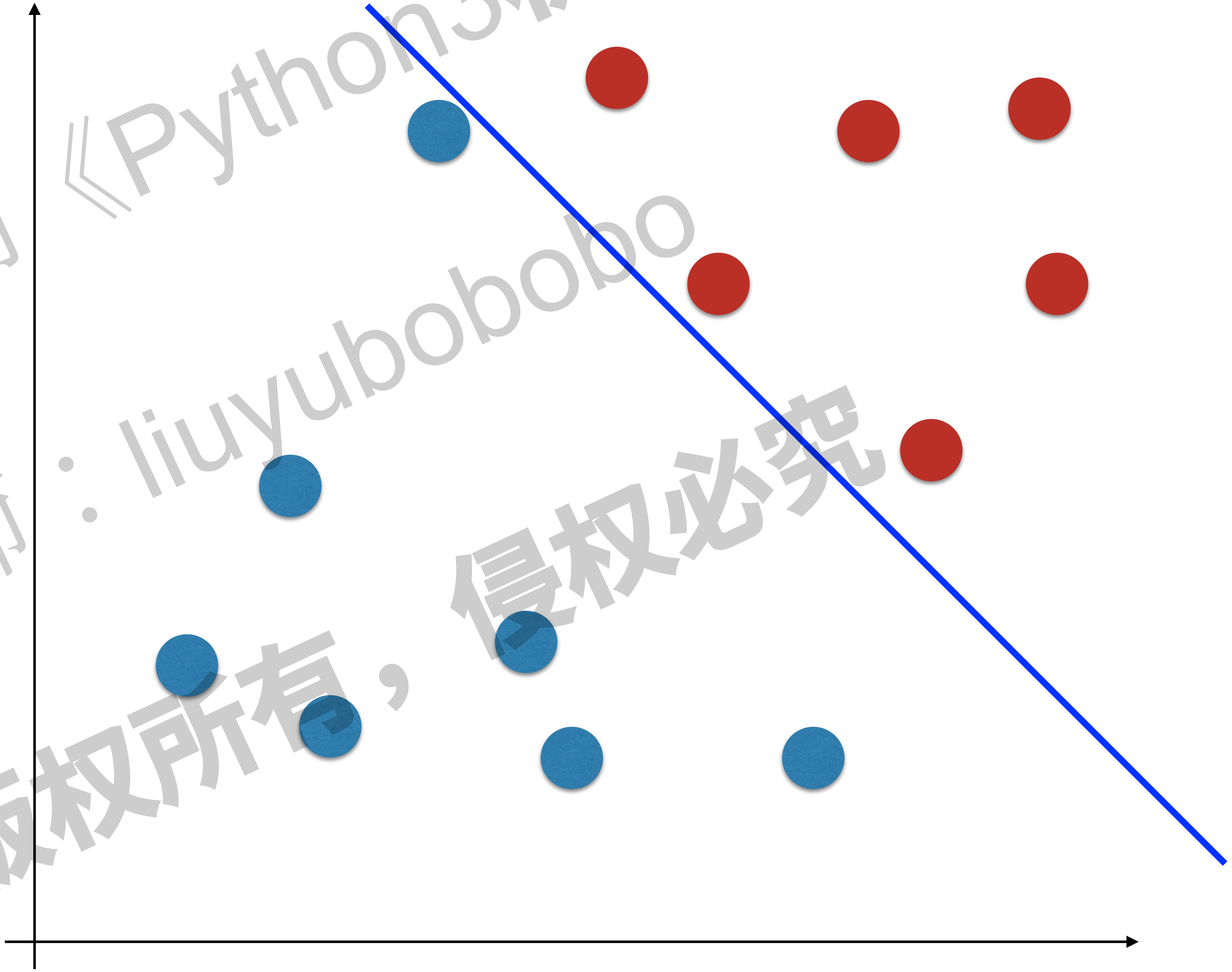
Support Vector Machine

支撑向量机 SVM

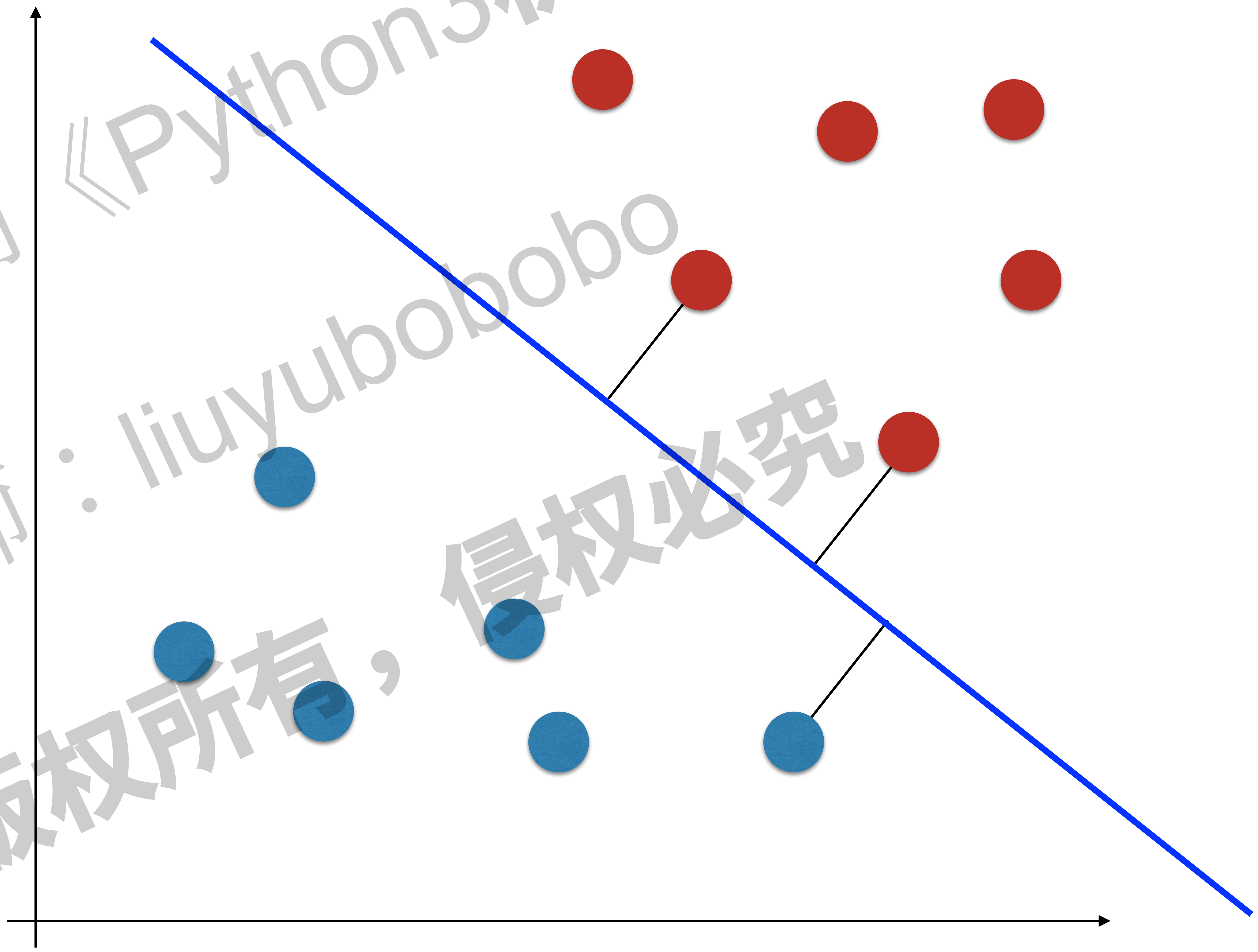
不适定问题



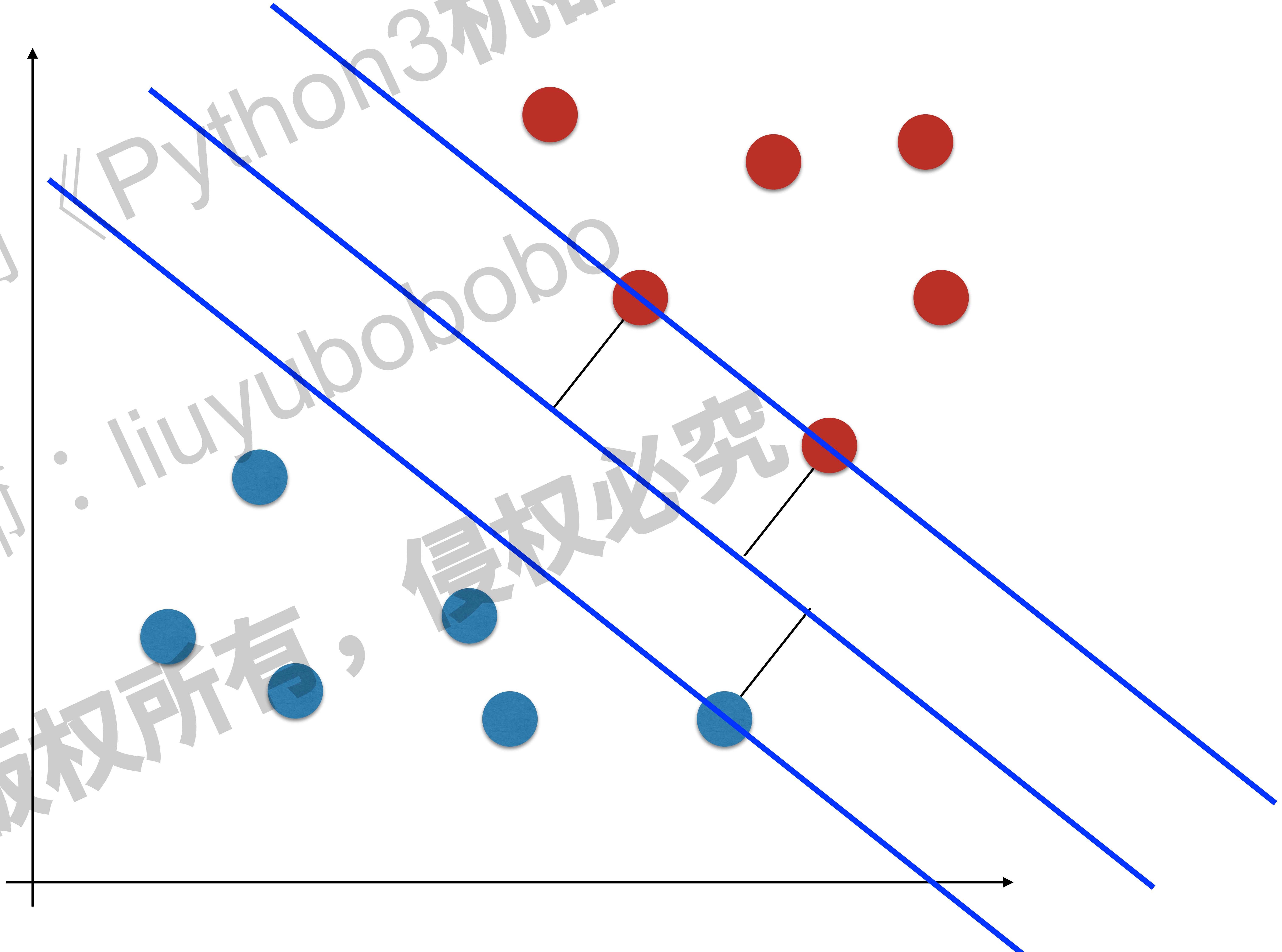
支撑向量机 SVM



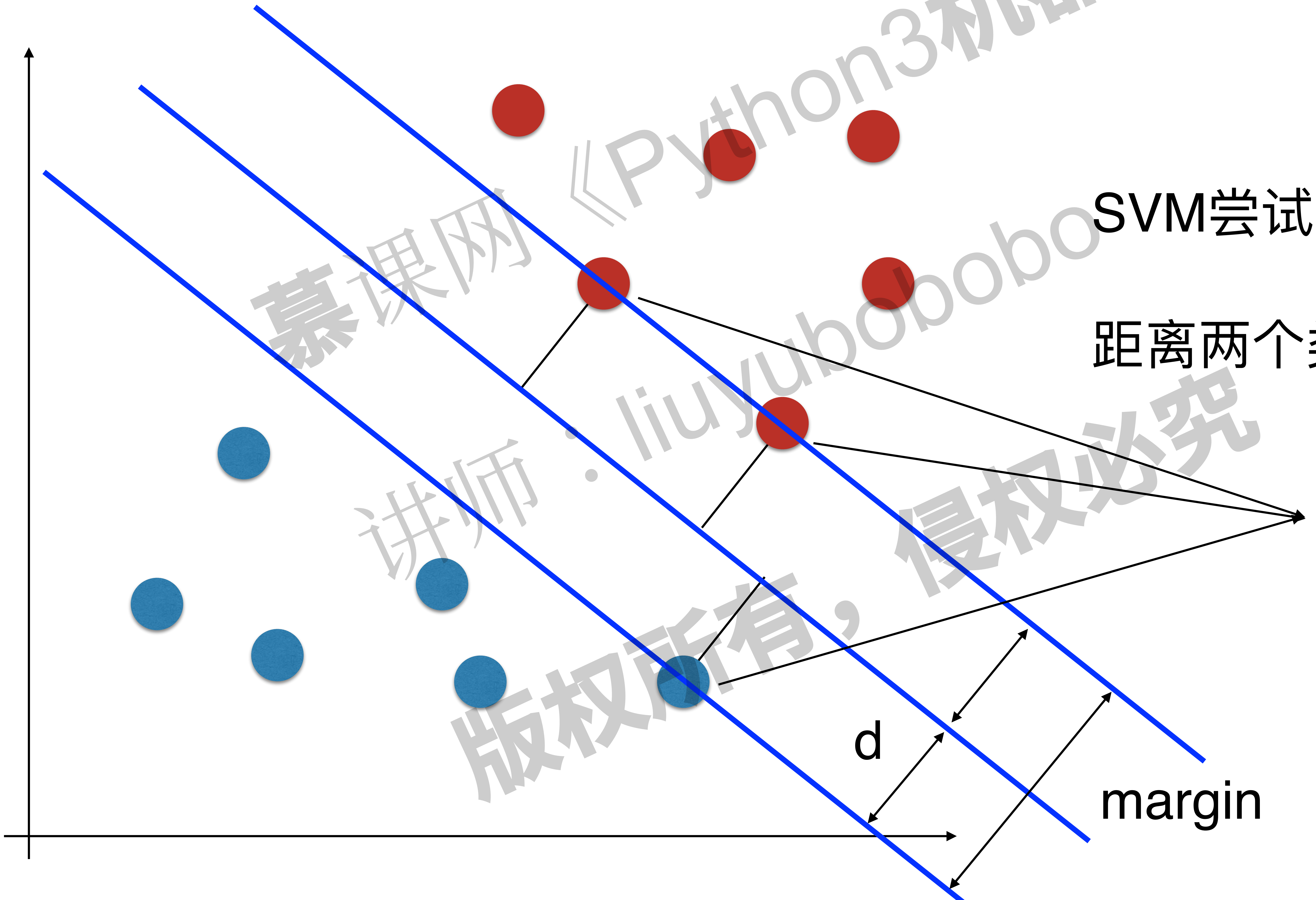
支撑向量机 SVM



支撑向量机 SVM



支撑向量机SVM

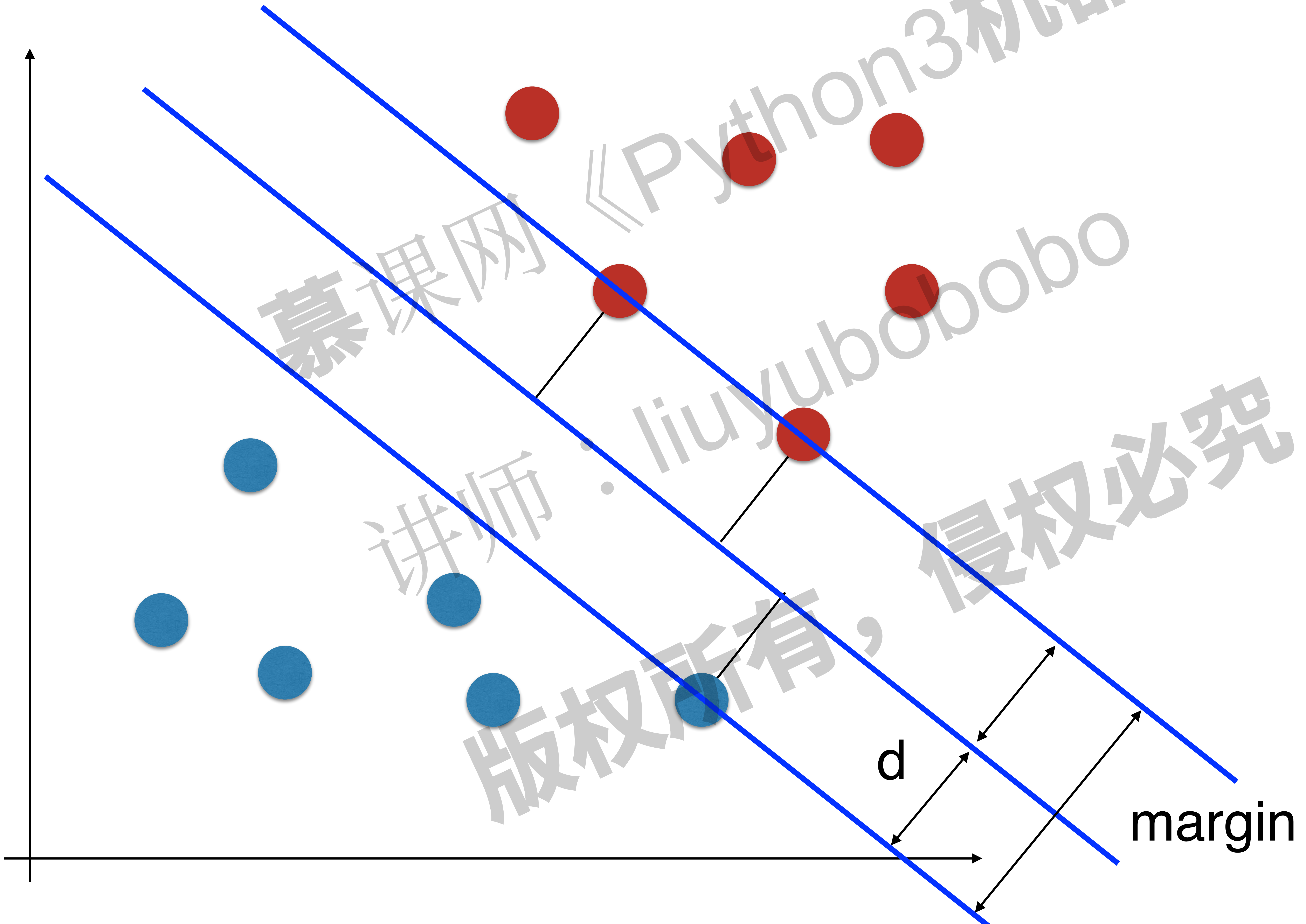


SVM尝试寻找一个最优的决策边界

距离两个类别的最近的样本最远

SVM要最大化margin

支撑向量机SVM



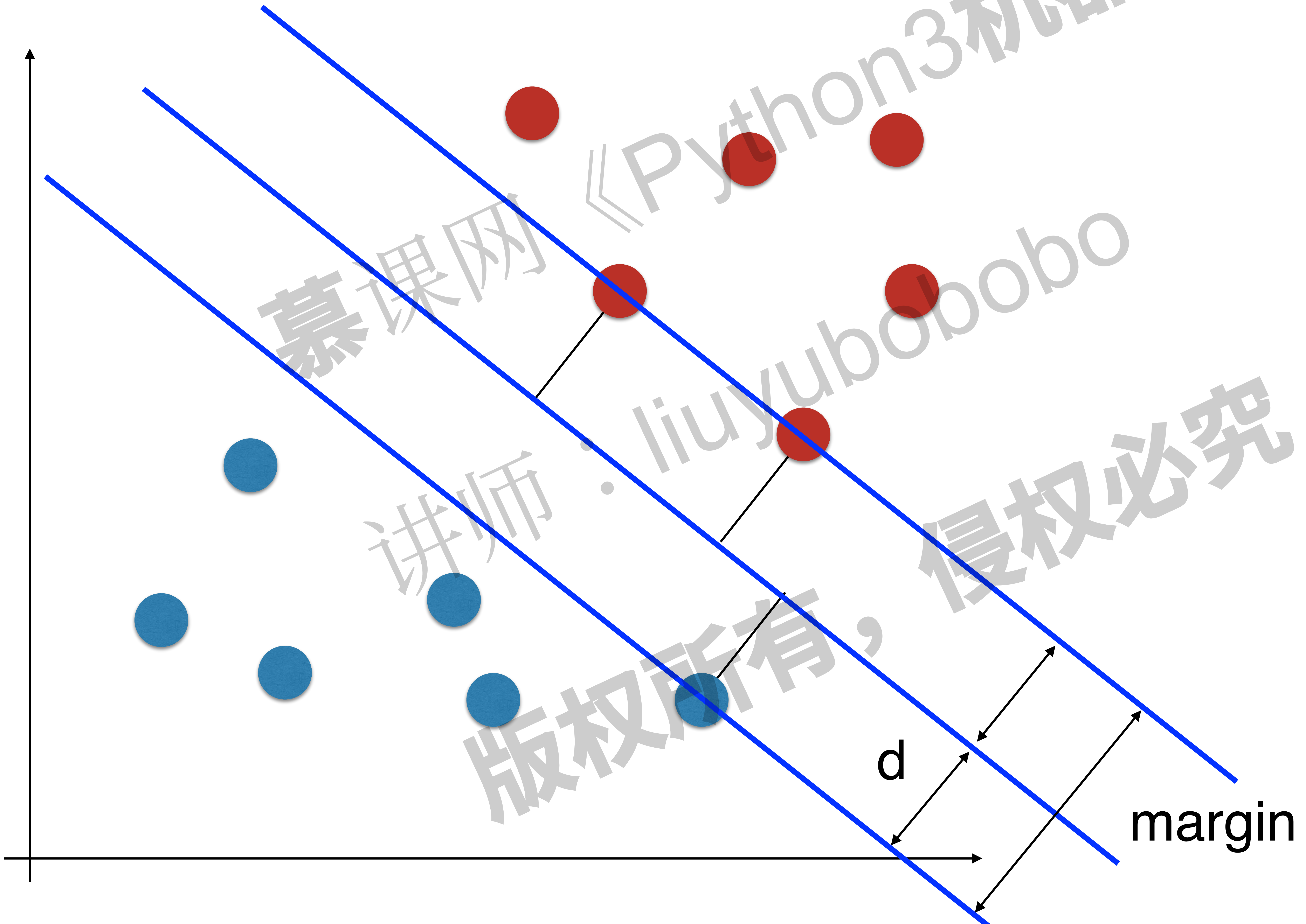
SVM要最大化margin

解决的是线性可分问题

Hard Margin SVM

Soft Margin SVM

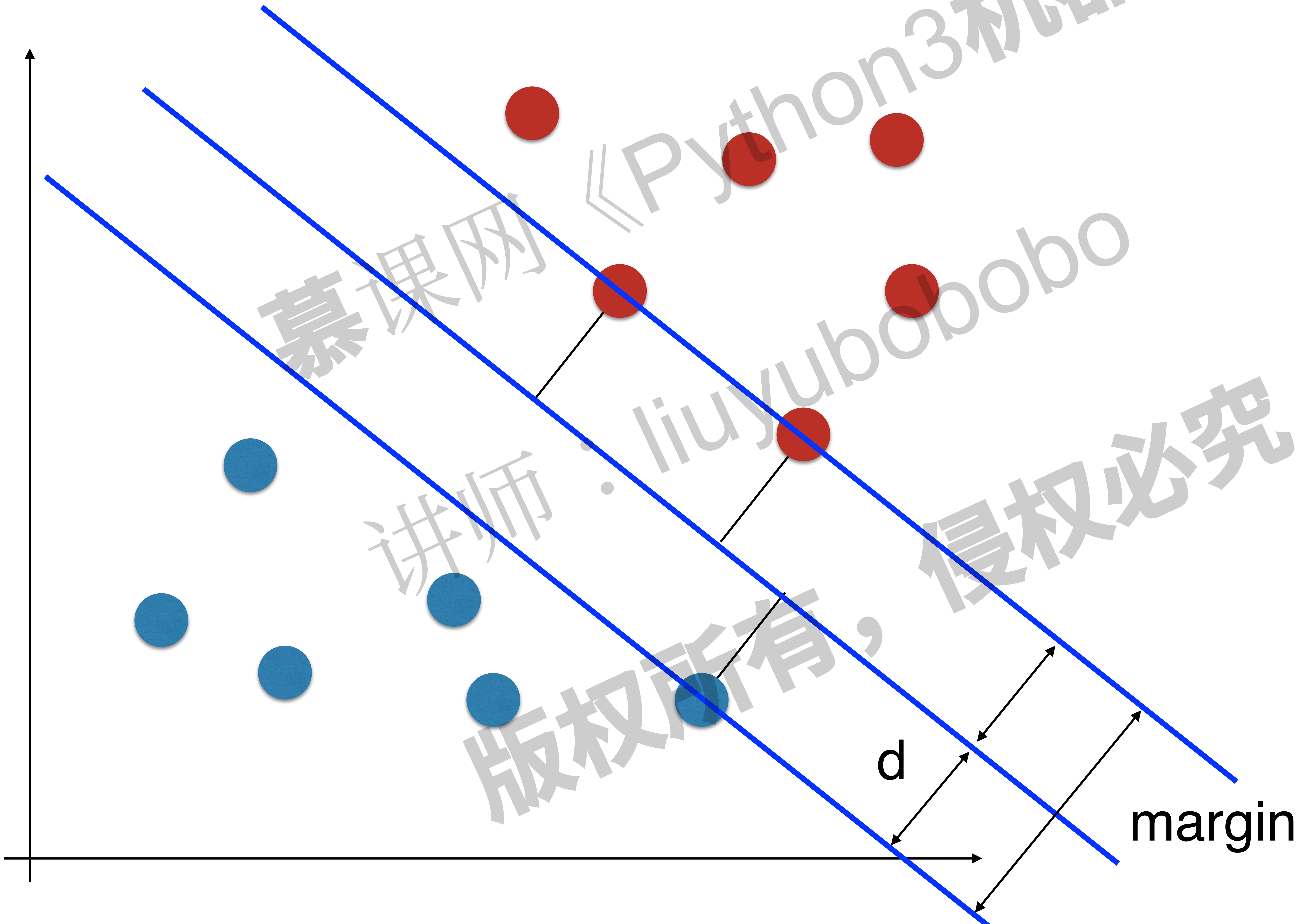
支撑向量机SVM



SVM要最大化margin

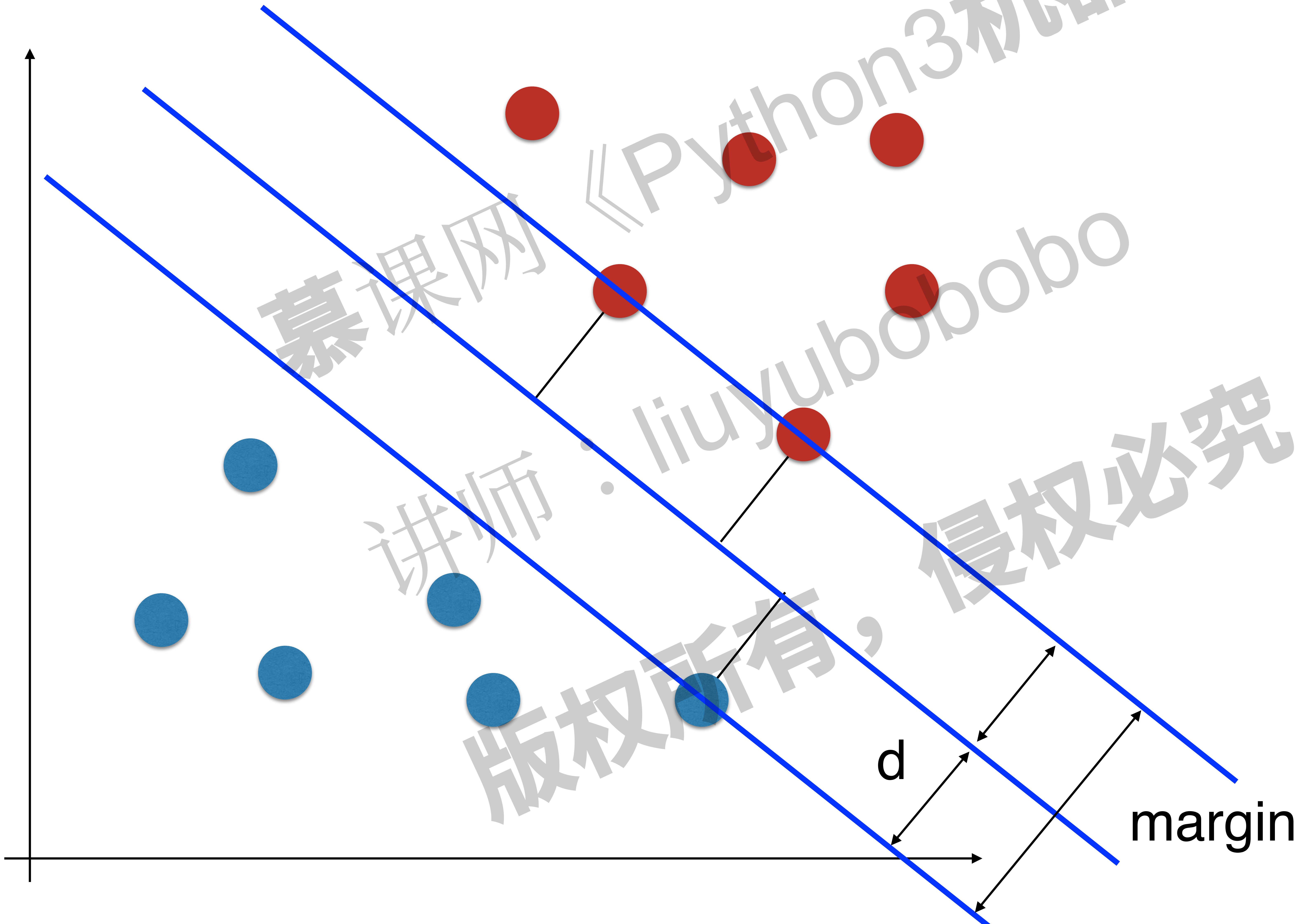
SVM的最优化目标

支撑向量机 SVM



SVM要最大化margin

支撑向量机 SVM



SVM要最大化margin

margin = 2d

SVM要最大化d

支撑向量机SVM

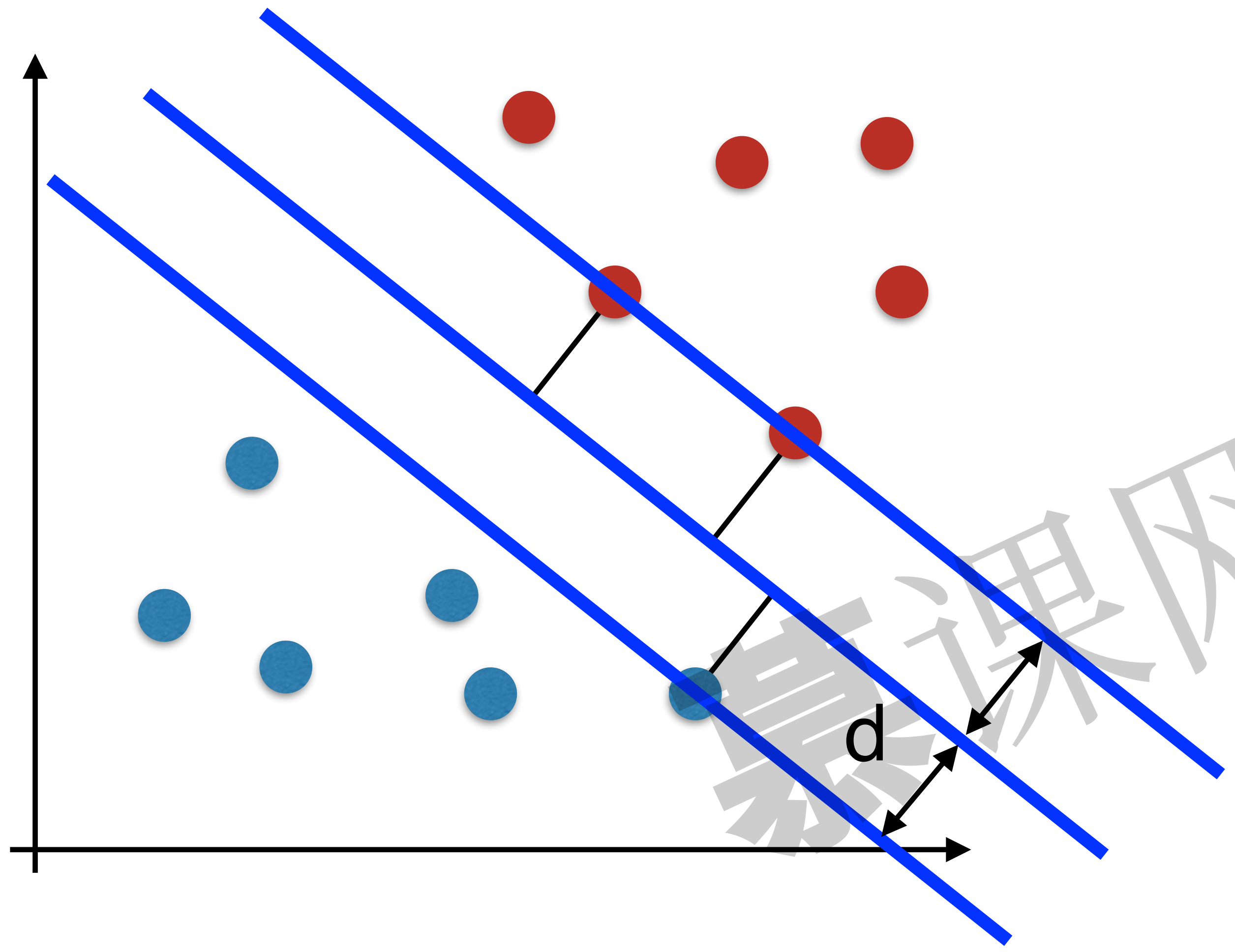
回忆解析几何，点到直线的距离

(x, y) 到 Ax + By + C = 0 的距离 $\frac{|Ax+By+C|}{\sqrt{A^2+B^2}}$

拓展到n维空间 $\theta^T x_b = 0$ ➡ $w^T x + b = 0$

$$\frac{|w^T x + b|}{\|w\|} \qquad \|w\| = \sqrt{w_1^2 + w_2^2 + \ldots + w_n^2}$$

支撑向量机SVM



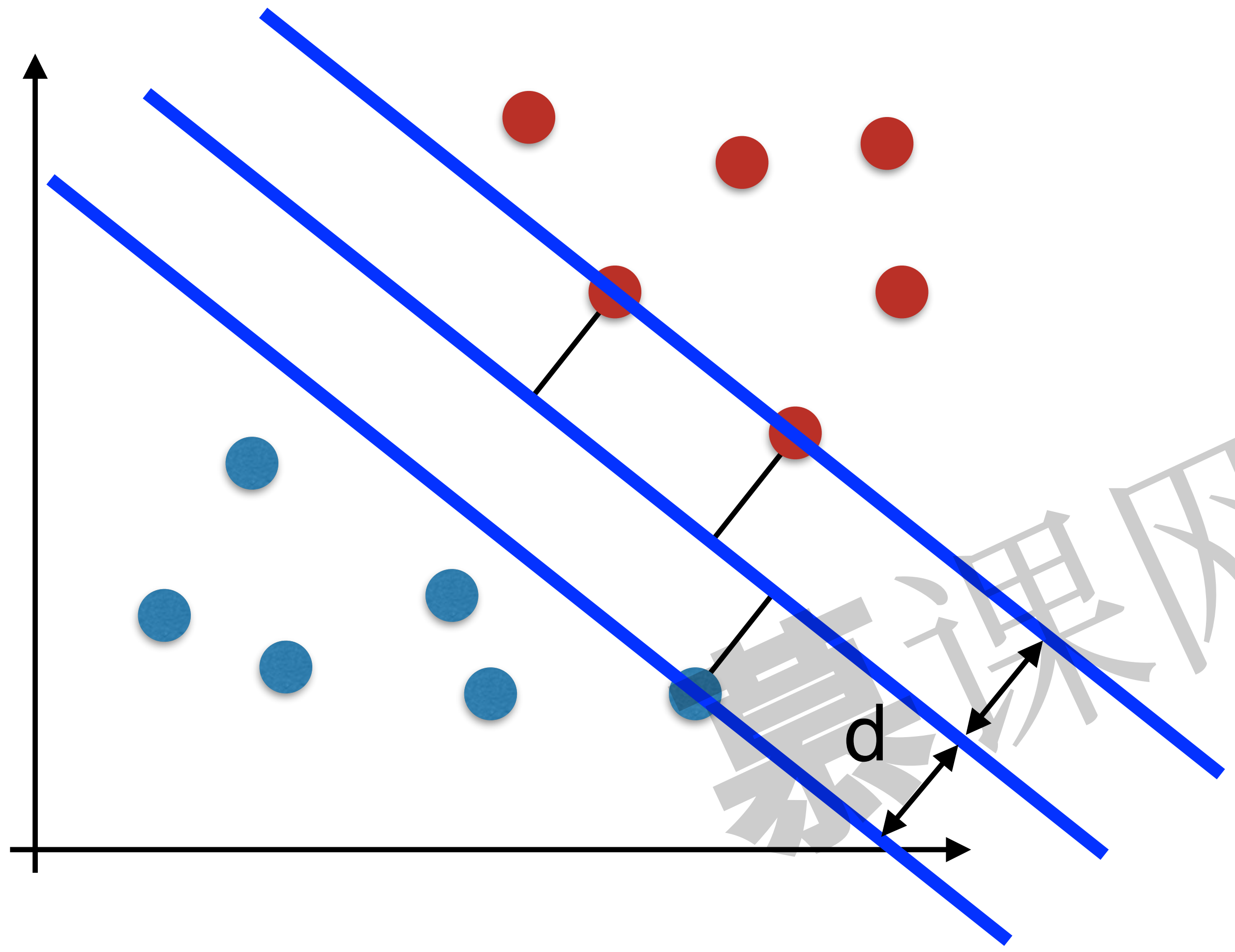
$$w^T x + b = 0$$

$$\frac{|w^T x + b|}{\|w\|}$$

$$\begin{cases} \dfrac{w^T x^{(i)} + b}{\|w\|} \geq d & \forall y^{(i)} = 1 \\ \dfrac{w^T x^{(i)} + b}{\|w\|} \leq -d & \forall y^{(i)} = -1 \end{cases}$$

$$\begin{cases} \dfrac{w^T x^{(i)} + b}{\|w\|d} \geq 1 & \forall y^{(i)} = 1 \\ \dfrac{w^T x^{(i)} + b}{\|w\|d} \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

支撑向量机SVM



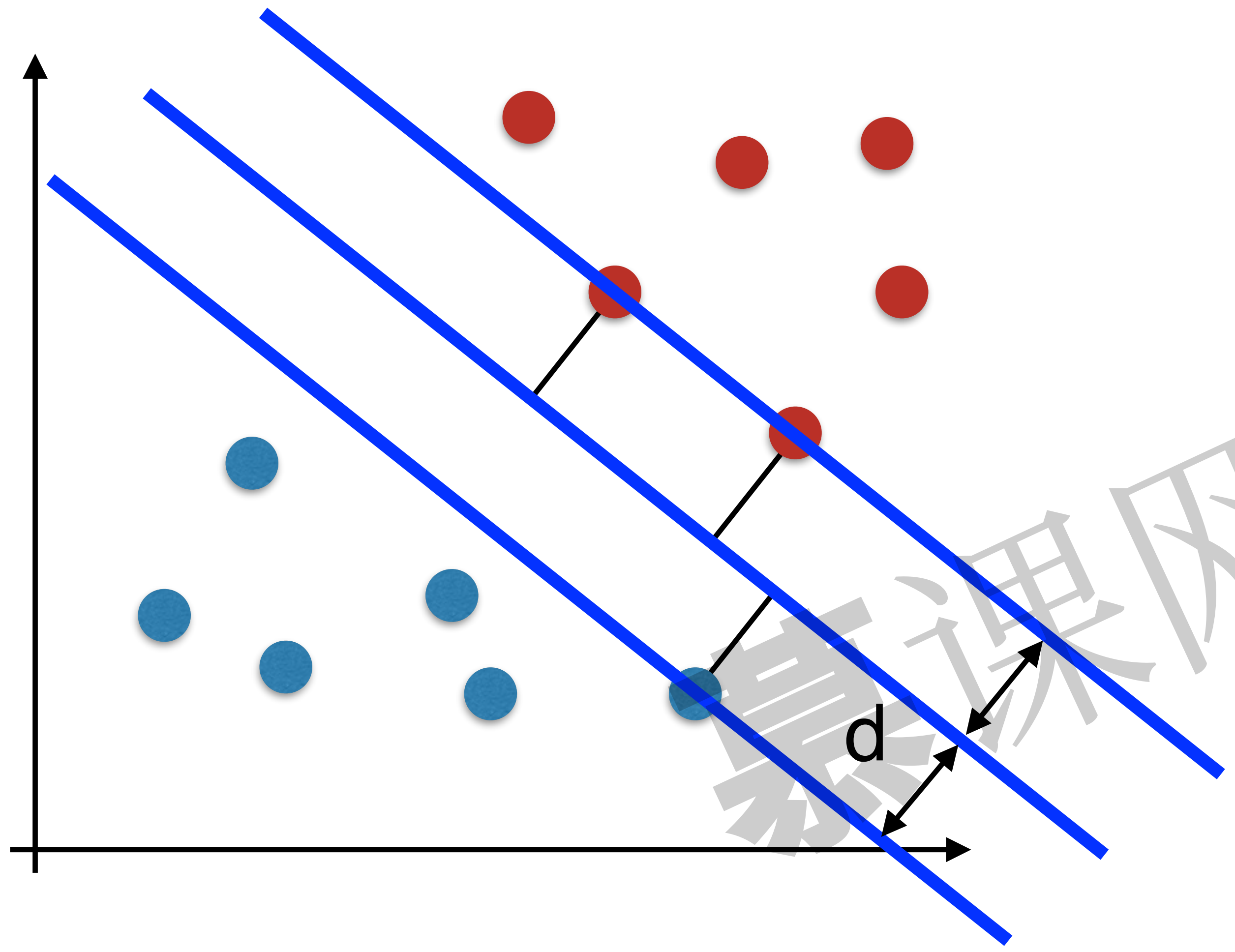
$$w^T x + b = 0$$

$$\frac{|w^T x + b|}{\|w\|}$$

$$\begin{cases} \dfrac{w^T x^{(i)} + b}{\|w\| d} \geq 1 & \forall y^{(i)} = 1 \\ \dfrac{w^T x^{(i)} + b}{\|w\| d} \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

$$\begin{cases} w_d^T x^{(i)} + b_d \geq 1 & \forall y^{(i)} = 1 \\ w_d^T x^{(i)} + b_d \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

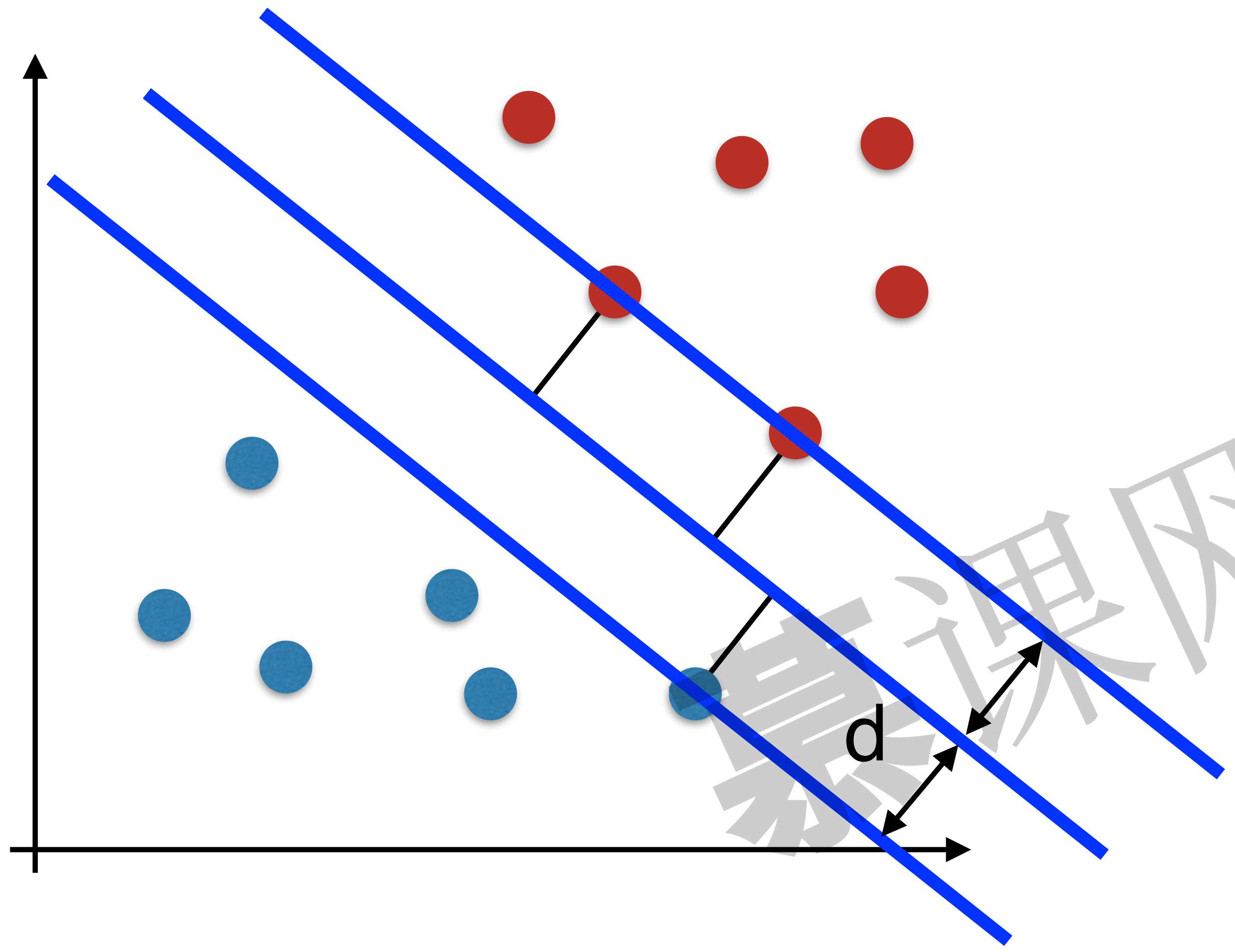
支撑向量机SVM



$$\begin{cases} \dfrac{w^T x^{(i)} + b}{\|w\| d} \geq 1 & \forall y^{(i)} = 1 \\ \dfrac{w^T x^{(i)} + b}{\|w\| d} \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

$$\begin{cases} w_d^T x^{(i)} + b_d \geq 1 & \forall y^{(i)} = 1 \\ w_d^T x^{(i)} + b_d \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

支撑向量机SVM



$$w_d^T x + b_d = 1$$

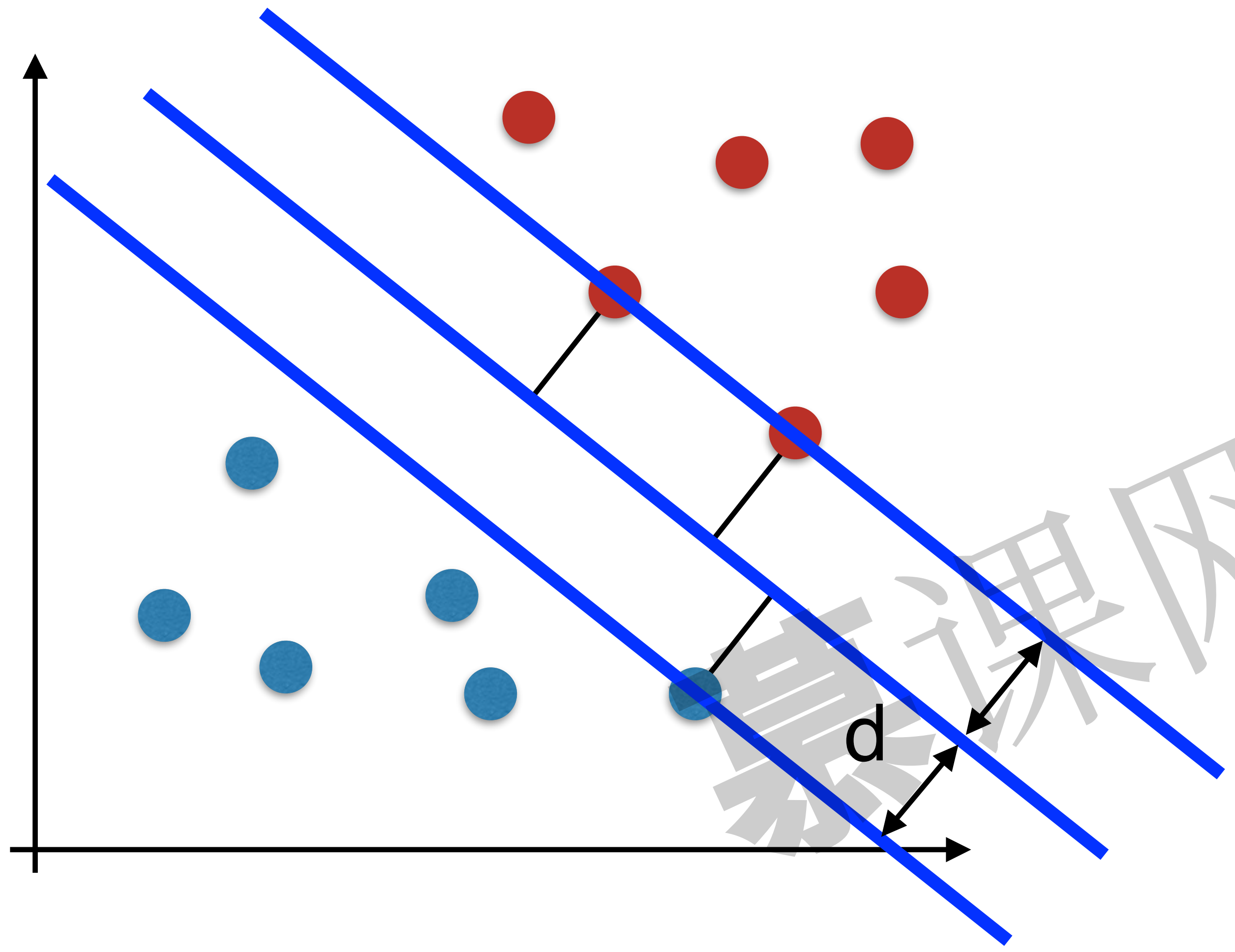
$$w_d^T x + b_d = 0$$

$$w_d^T x + b_d = -1$$

$$\begin{cases} \dfrac{w^T x^{(i)} + b}{\|w\| d} \geq 1 & \forall y^{(i)} = 1 \\ \dfrac{w^T x^{(i)} + b}{\|w\| d} \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

$$\begin{cases} w_d^T x^{(i)} + b_d \geq 1 & \forall y^{(i)} = 1 \\ w_d^T x^{(i)} + b_d \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

支撑向量机 SVM



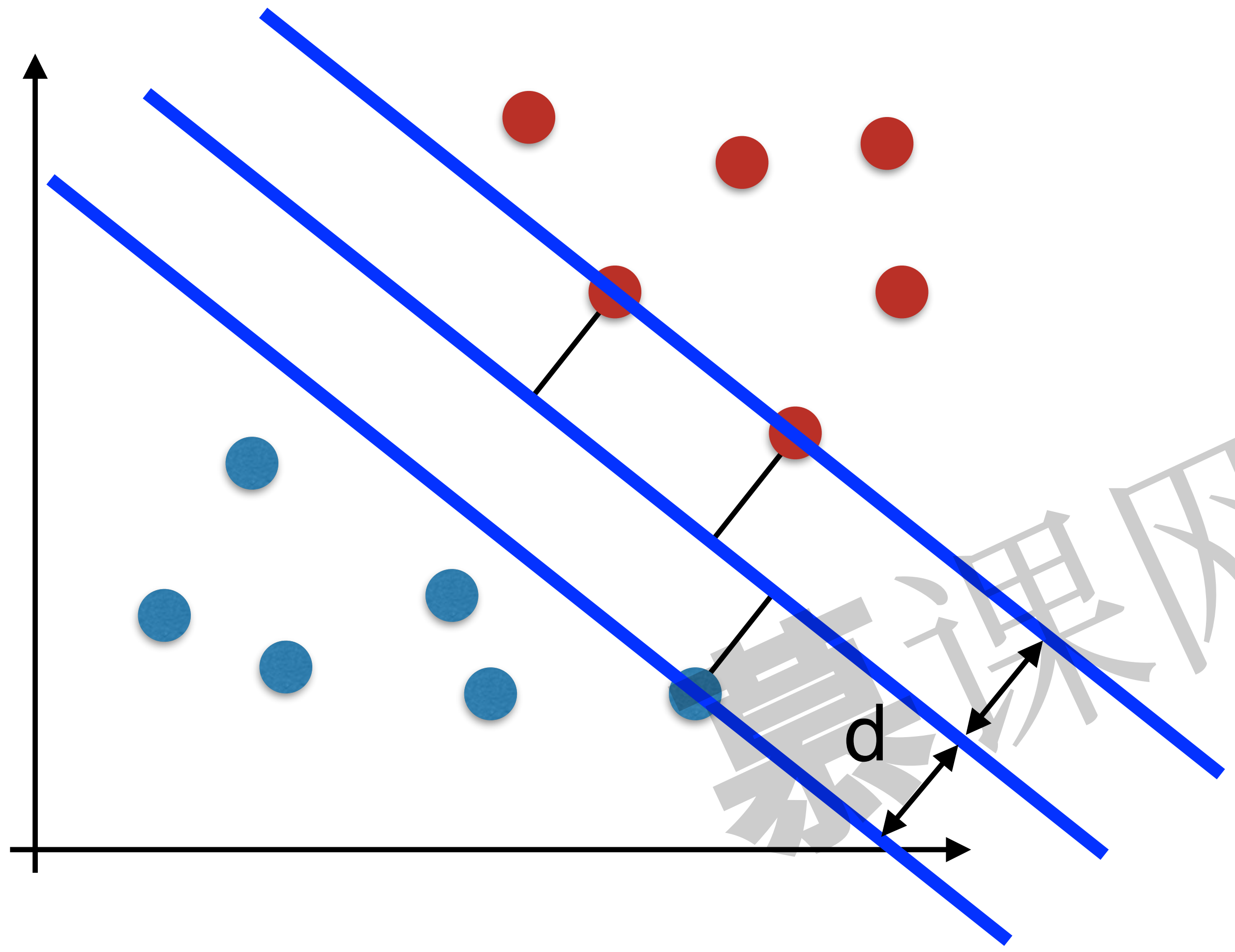
$$w_d^T x + b_d = 1$$

$$w_d^T x + b_d = 0$$

$$w_d^T x + b_d = -1$$

$$\begin{cases} w_d^T x^{(i)} + b_d \geq 1 & \forall y^{(i)} = 1 \\ w_d^T x^{(i)} + b_d \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

支撑向量机SVM



$$w^T x + b = 1$$

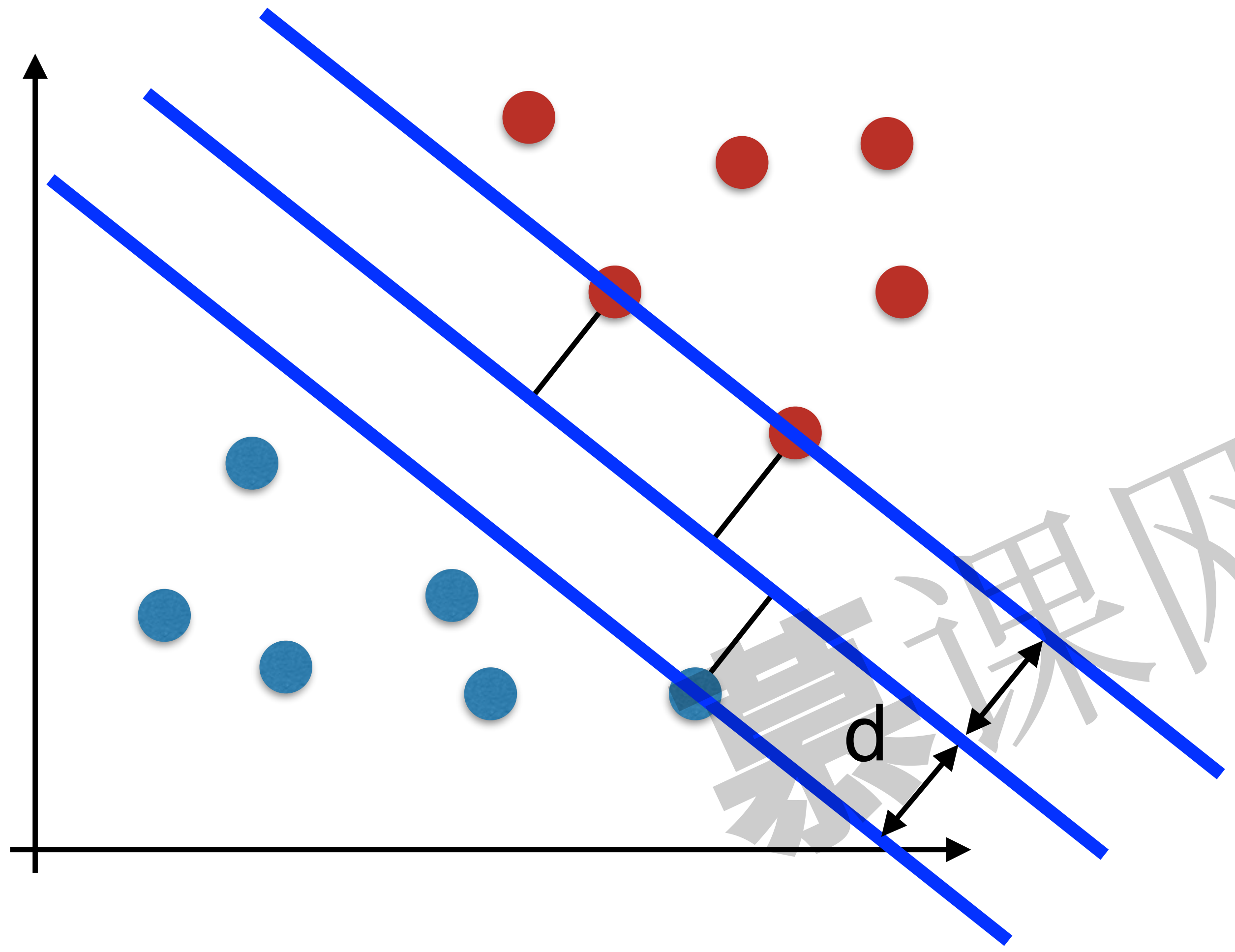
$$w^T x + b = 0$$

$$w^T x + b = -1$$

$$\begin{cases} w^T x^{(i)} + b \geq 1 & \forall y^{(i)} = 1 \\ w^T x^{(i)} + b \leq -1 & \forall y^{(i)} = -1 \end{cases}$$

$$y^{(i)}(w^T x^{(i)} + b) \geq 1$$

支撑向量机 SVM



$$w^T x + b = 1$$

$$w^T x + b = 0$$

$$w^T x + b = -1$$

$$y^{(i)}(w^T x^{(i)} + b) \geq 1$$

对于任意支撑向量x

$$\max \frac{|w^T x + b|}{||w||} \qquad \max \frac{1}{||w||}$$

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

支撑向量机 SVM

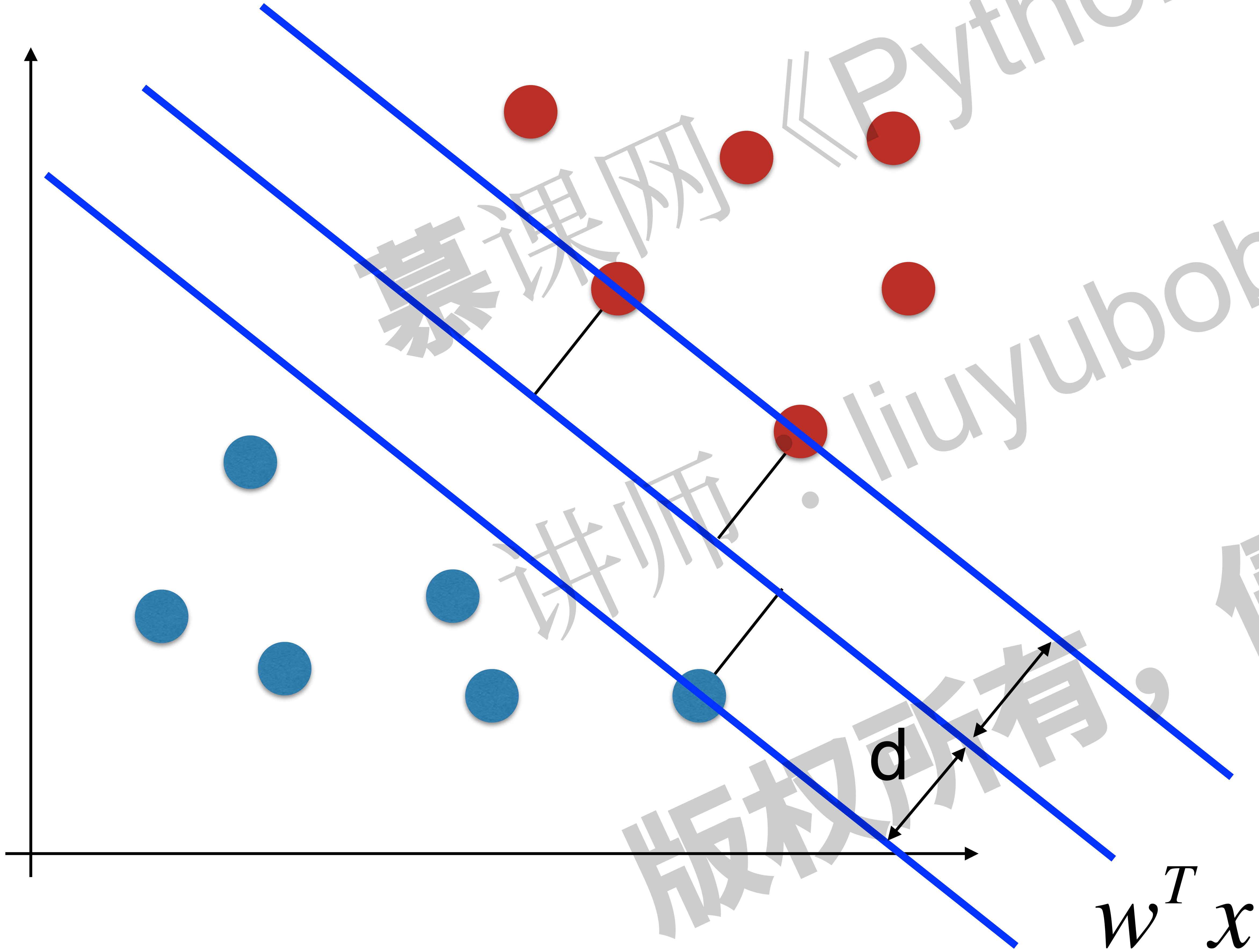
$$\min \frac{1}{2}||w||^2$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \geq 1$$

有条件的最优化问题

Soft Margin和SVM的正则化

支撑向量机SVM

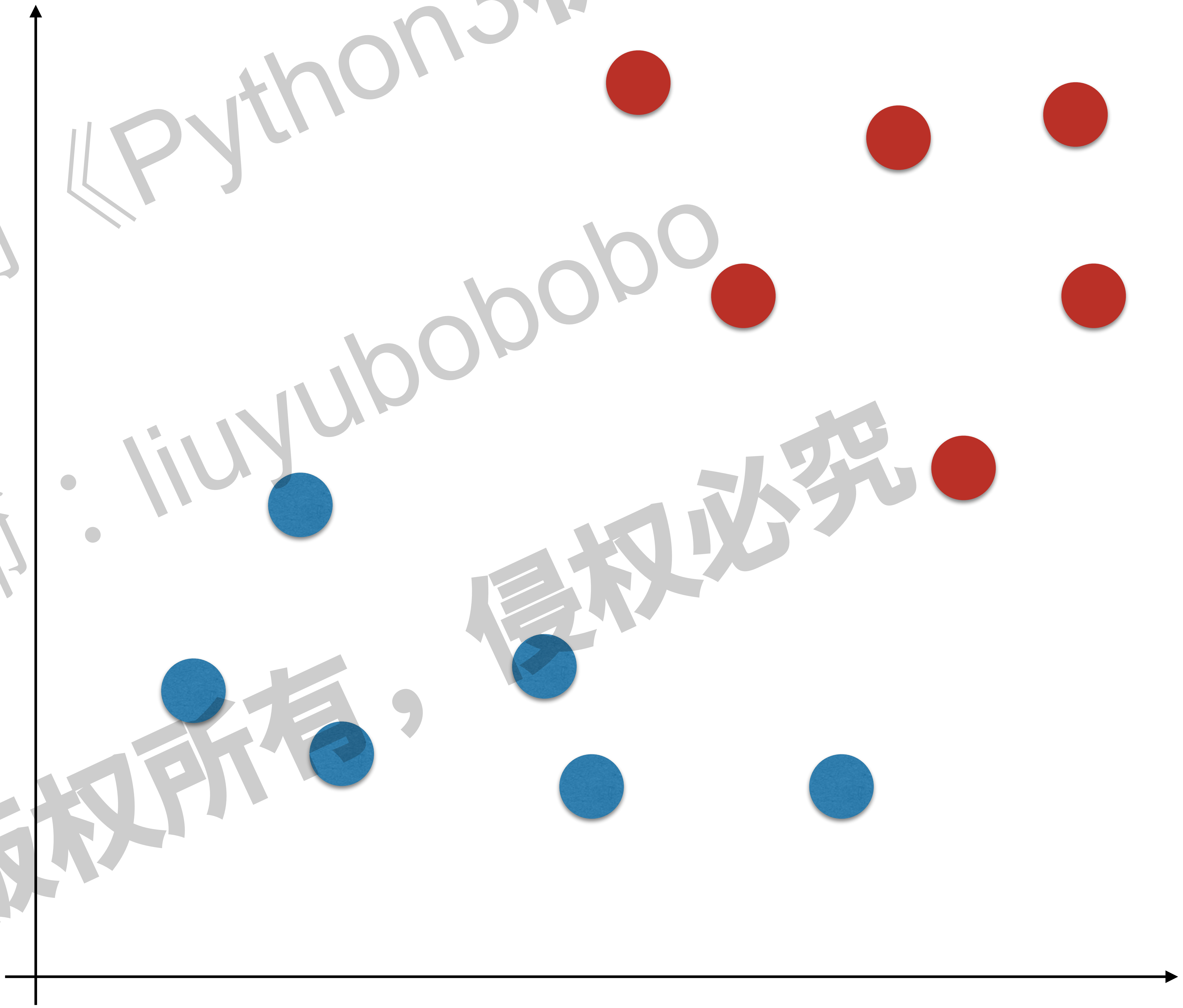


$$w^T x + b = 0$$

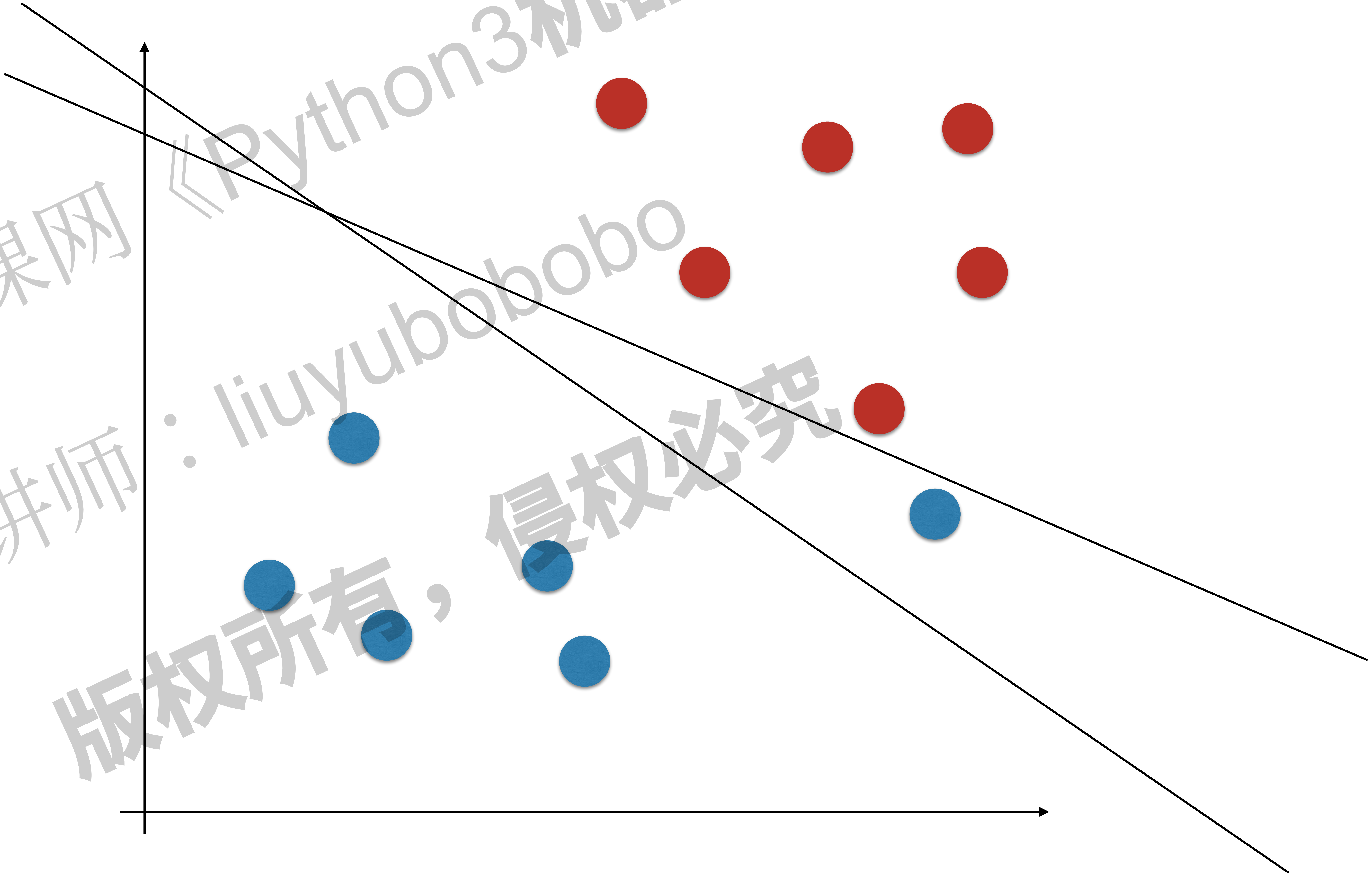
$$\min \frac{1}{2}||w||$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \geq 1$$

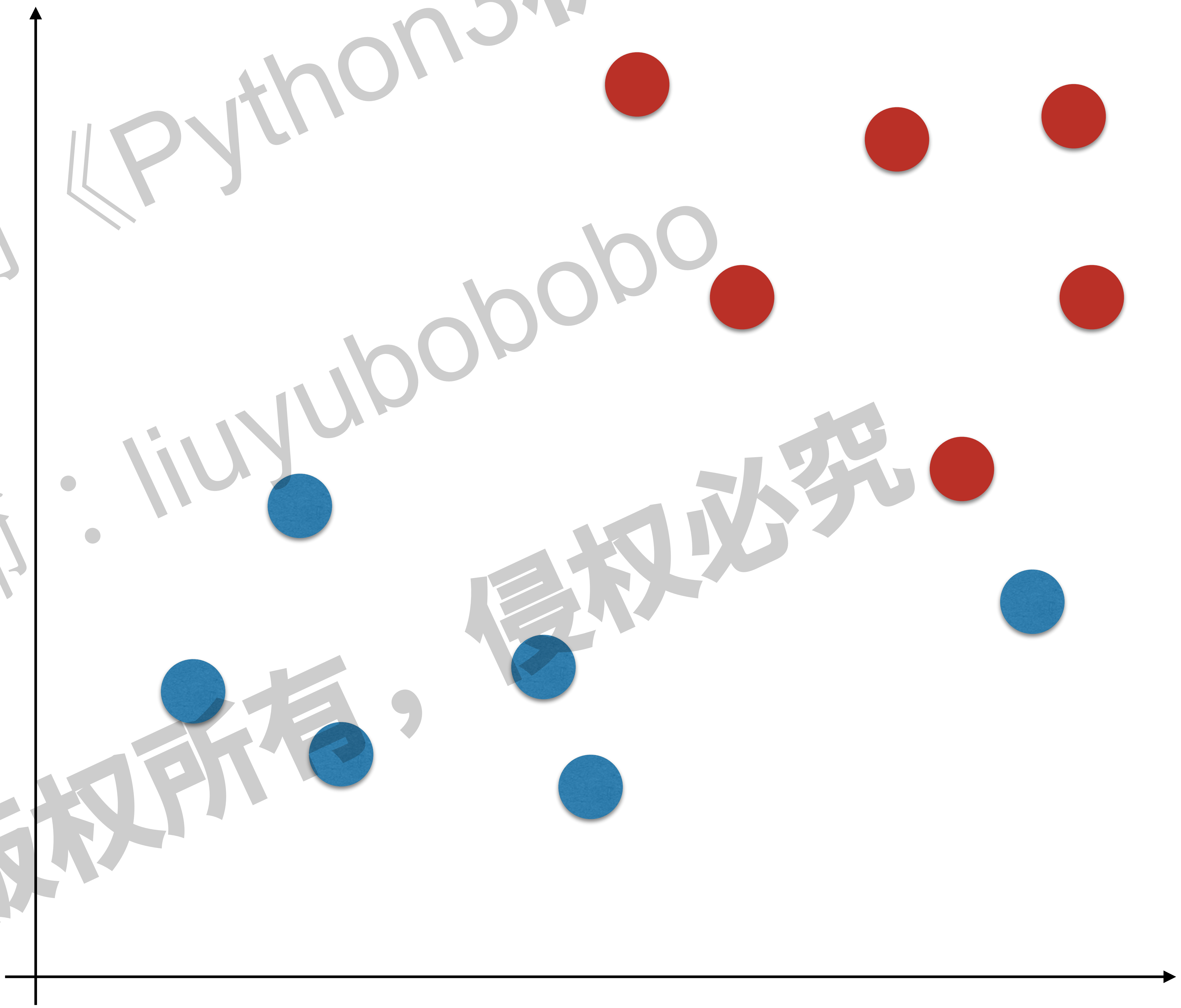
支撑向量机 SVM



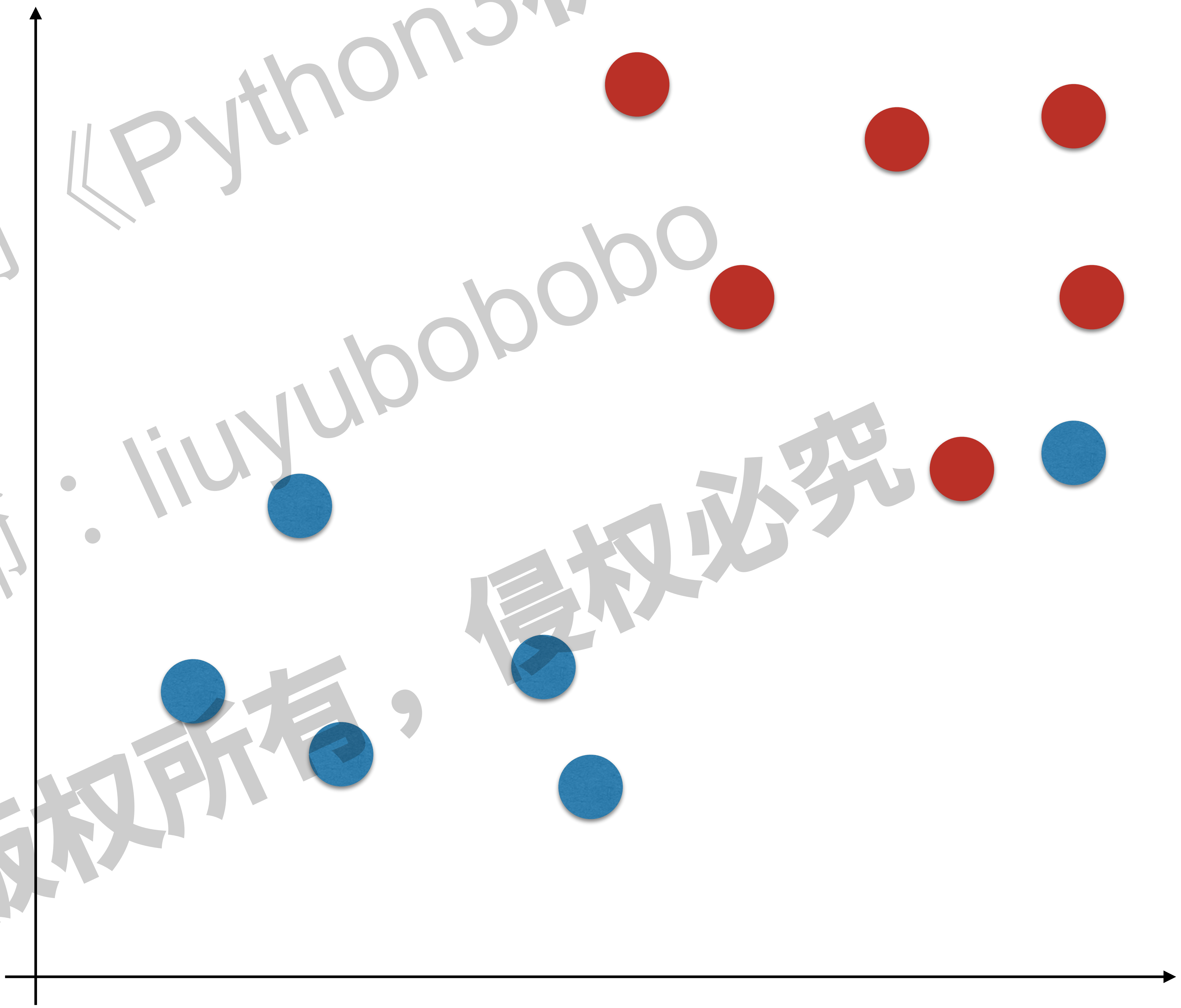
支撑向量机 SVM



支撑向量机 SVM



支撑向量机 SVM



Soft Margin SVM

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

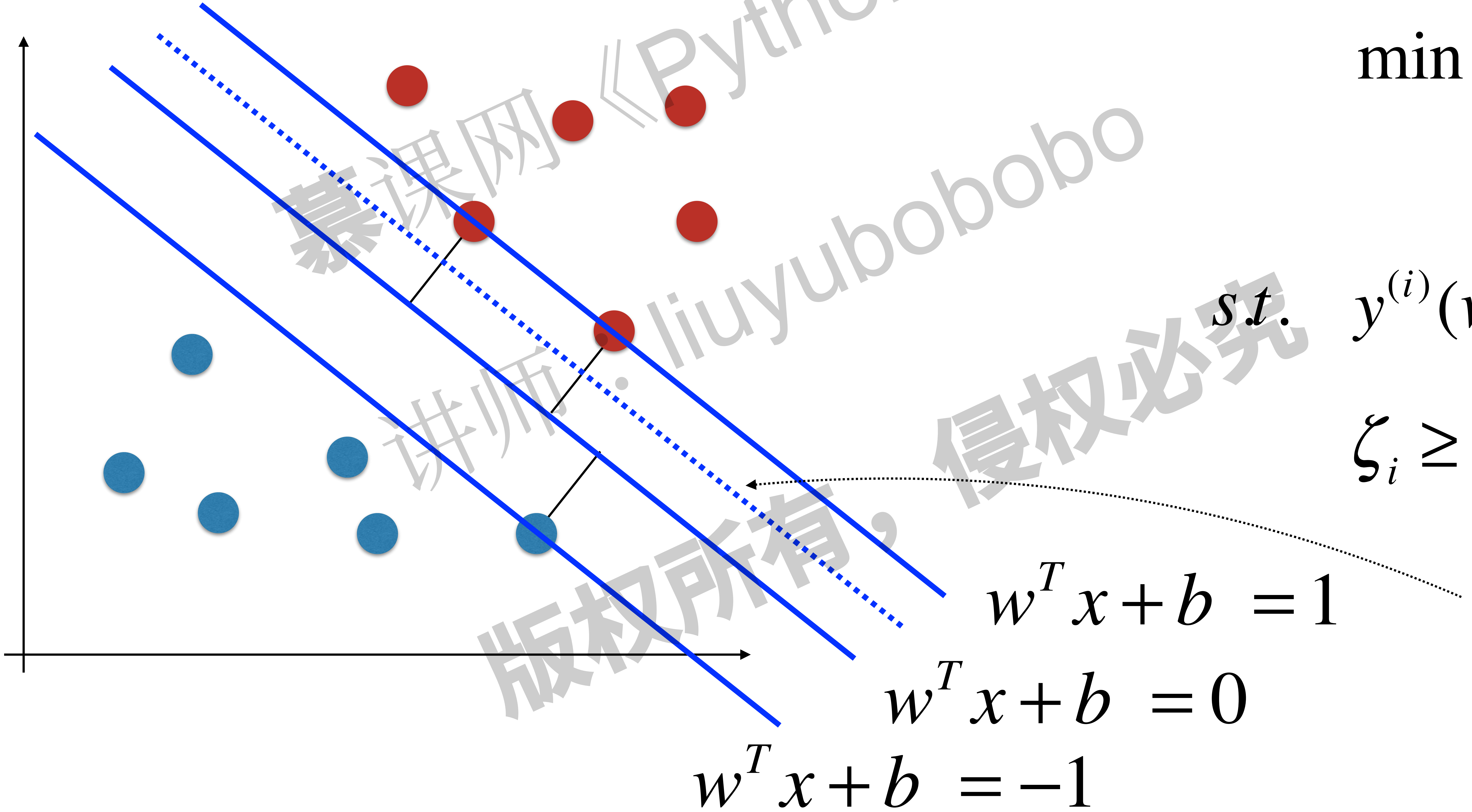
$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \geq 1 - \zeta_i$$

Soft Margin SVM

$$\min \frac{1}{2}\|w\|^2 + \sum_{i=1}^{m} \zeta_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \ge 1 - \zeta_i$$

$$\zeta_i \ge 0$$



Soft Margin SVM

$$\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\zeta_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \geq 1 - \zeta_i$$

$$\zeta_i \geq 0$$

Soft Margin SVM

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\zeta_i$$

$$s.t.\quad y^{(i)}(w^T x^{(i)} + b) \ge 1 - \zeta_i$$

$$\zeta_i \ge 0$$

L1正则

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\zeta_i^2$$

$$s.t.\quad y^{(i)}(w^T x^{(i)} + b) \ge 1 - \zeta_i$$

$$\zeta_i \ge 0$$

L2正则

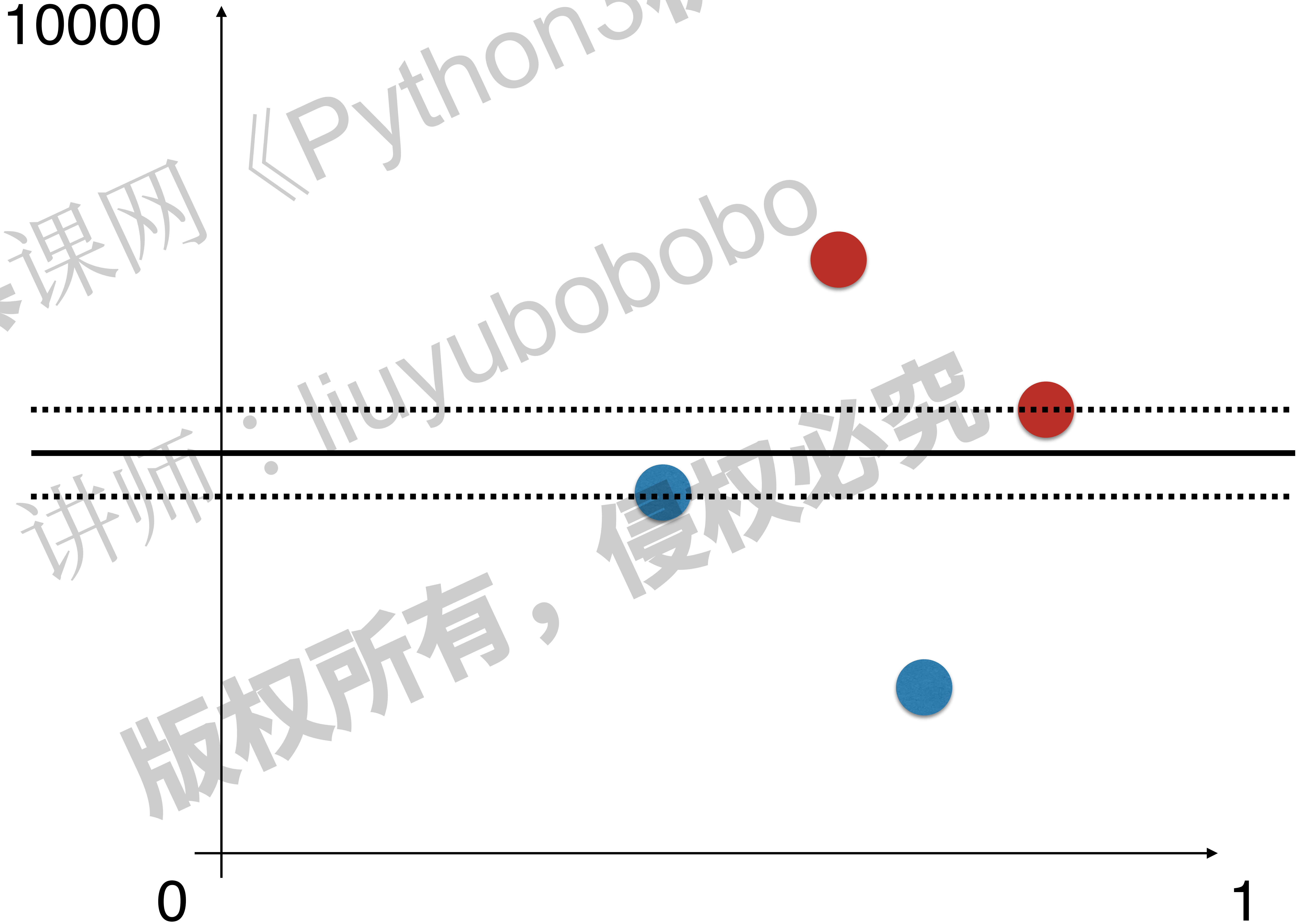
scikit-learn中的SVM

实际使用SVM

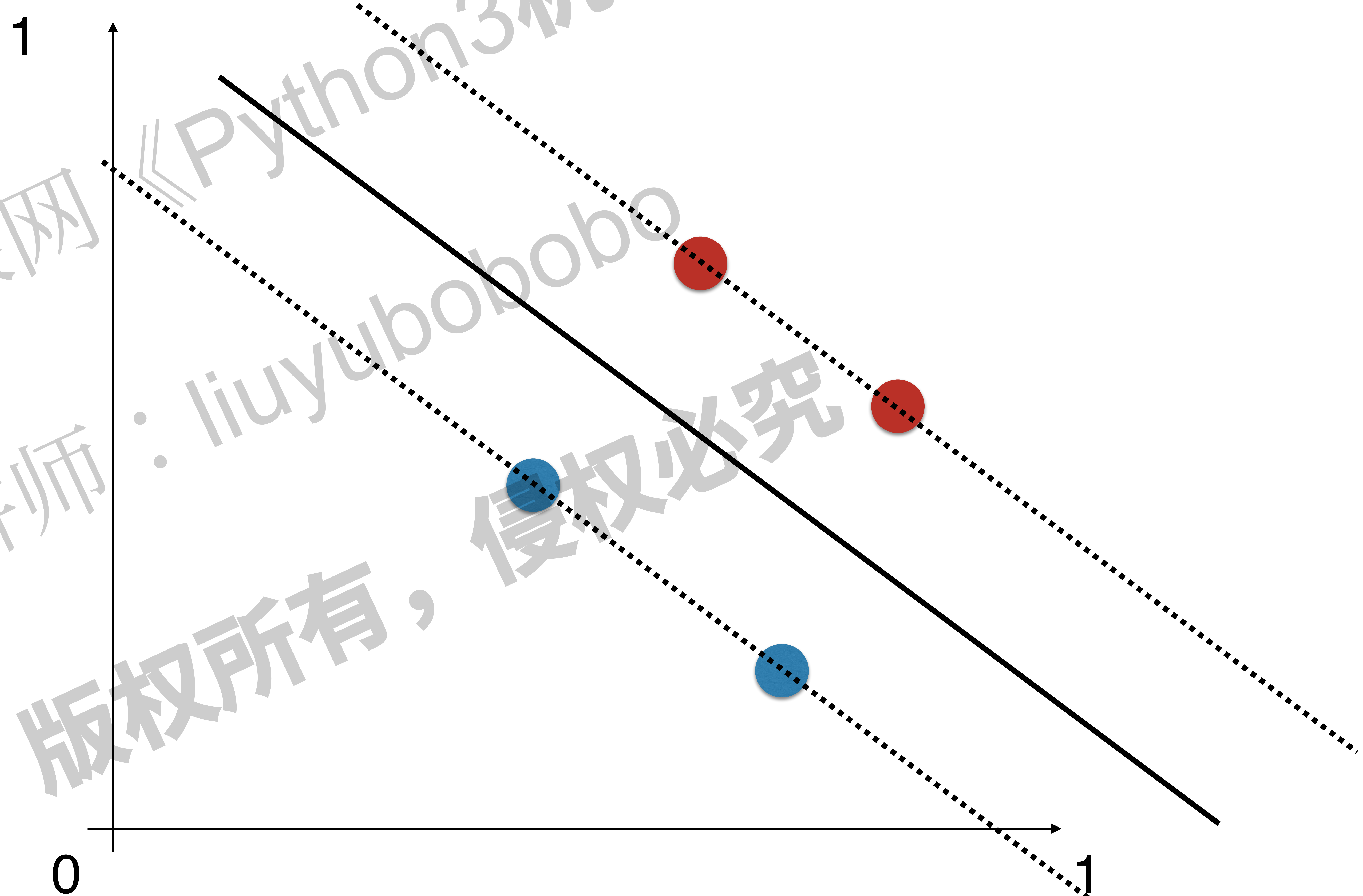
和kNN一样，要做数据标准化处理！

涉及距离！

实际使用SVM



实际使用SVM



实践：scikit-learn中的linearSVC

SVM中使用多项式特征

实践：SVM使用多项式特征

什么是核函数

SVM

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\zeta_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \ge 1 - \zeta_i$$

$$\zeta_i \ge 0$$

SVM

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\zeta_i$$

$$s.t. \quad y^{(i)}(w^T x^{(i)} + b) \geq 1 - \zeta_i$$

$$\zeta_i \geq 0$$

$$\max \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j x_i x_j$$

$$s.t. \quad 0 \leq \alpha_i \leq C$$

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

SVM

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \boxed{x_i x_j}$$

$$s.t. \quad 0 \le \alpha_i \le C$$

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

核函数

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \boxed{x_i x_j}$$

$$s.t. \quad 0 \le \alpha_i \le C$$

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

$$x^{(i)} \rightarrow x'^{(i)}$$

$$x^{(j)} \rightarrow x'^{(j)}$$

$$x'^{(i)} x'^{(j)}$$

核函数

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \boxed{x_i x_j}$$

$$x'^{(i)} x'^{(j)}$$

$$s.t. \quad 0 \le \alpha_i \le C$$

$$K(x^{(i)}, x^{(j)}) = x'^{(i)} x'^{(j)}$$

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

核函数

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \boxed{x_i x_j}$$

$$s.t. \quad 0 \le \alpha_i \le C$$

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

核函数

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$s.t. \quad 0 \le \alpha_i \le C$$

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

多项式核函数

$$K(x,y) = (x \cdot y + 1)^2$$

$$K(x,y) = (\sum_{i=1}^{n} x_i y_i + 1)^2$$

$$= \sum_{i=1}^{n} (x_i^2)(y_i^2) + \sum_{i=2}^{n} \sum_{j=1}^{i-1} (\sqrt{2} x_i x_j)(\sqrt{2} y_i y_j) + \sum_{i=1}^{n} (\sqrt{2} x_i)(\sqrt{2} y_i) + 1$$

$$x' = (x_n^2, \dots, x_1^2, \sqrt{2} x_n x_{n-1}, \dots, \sqrt{2} x_n, \dots \sqrt{2} x_1, 1) \qquad = x' \cdot y'$$

多项式核函数

$$\max \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$s.t. \quad 0 \le \alpha_i \le C$$

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

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^2$$

核函数

多项式核函数 $K(x,y) = (x \cdot y + c)^d$

线性核函数 $K(x,y) = x \cdot y$

高斯核函数

高斯核函数

K(x,y) 表示x和y的点乘

$$K(x,y) = e^{-\gamma\|x-y\|^2}$$

高斯函数 $g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$

高斯核函数

$$K(x,y) = e^{-\gamma \|x-y\|^2}$$

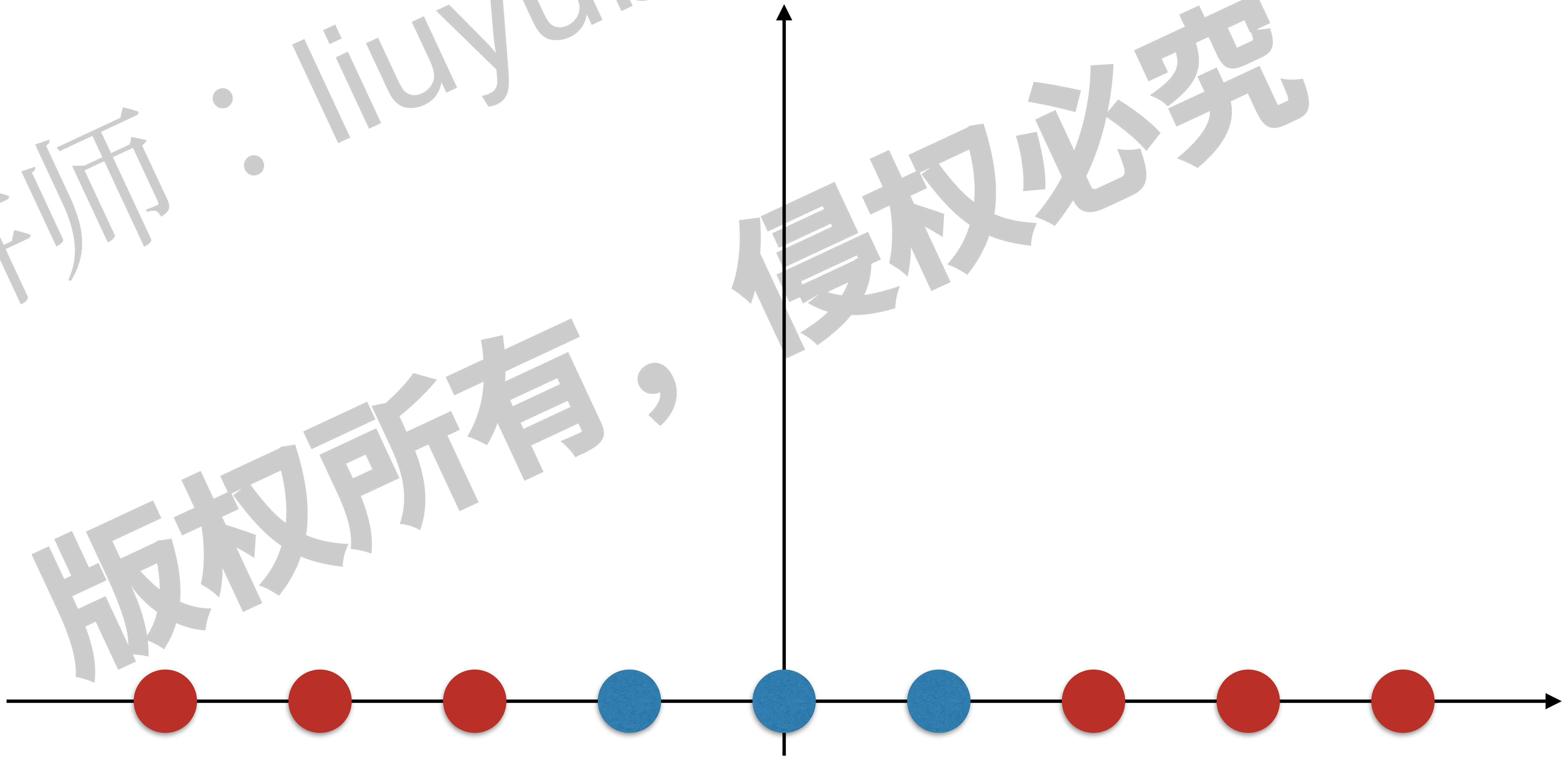
RBF核 Radial Basis Function Kernel

将每一个样本点映射到一个无穷维的特征空间

多项式特征

依靠升维使得原本线性不可分的数据线性可分

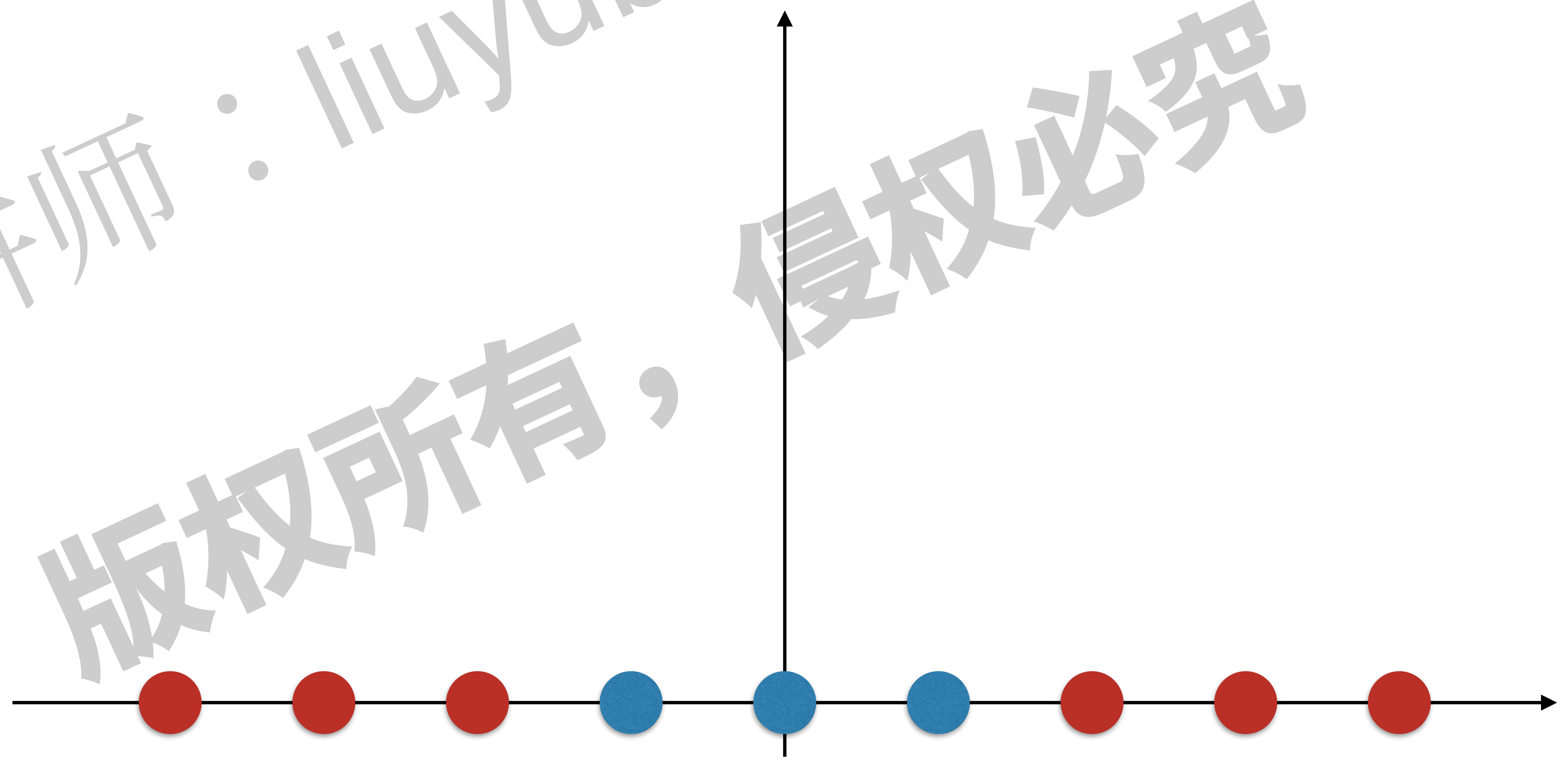
多项式特征



多项式特征

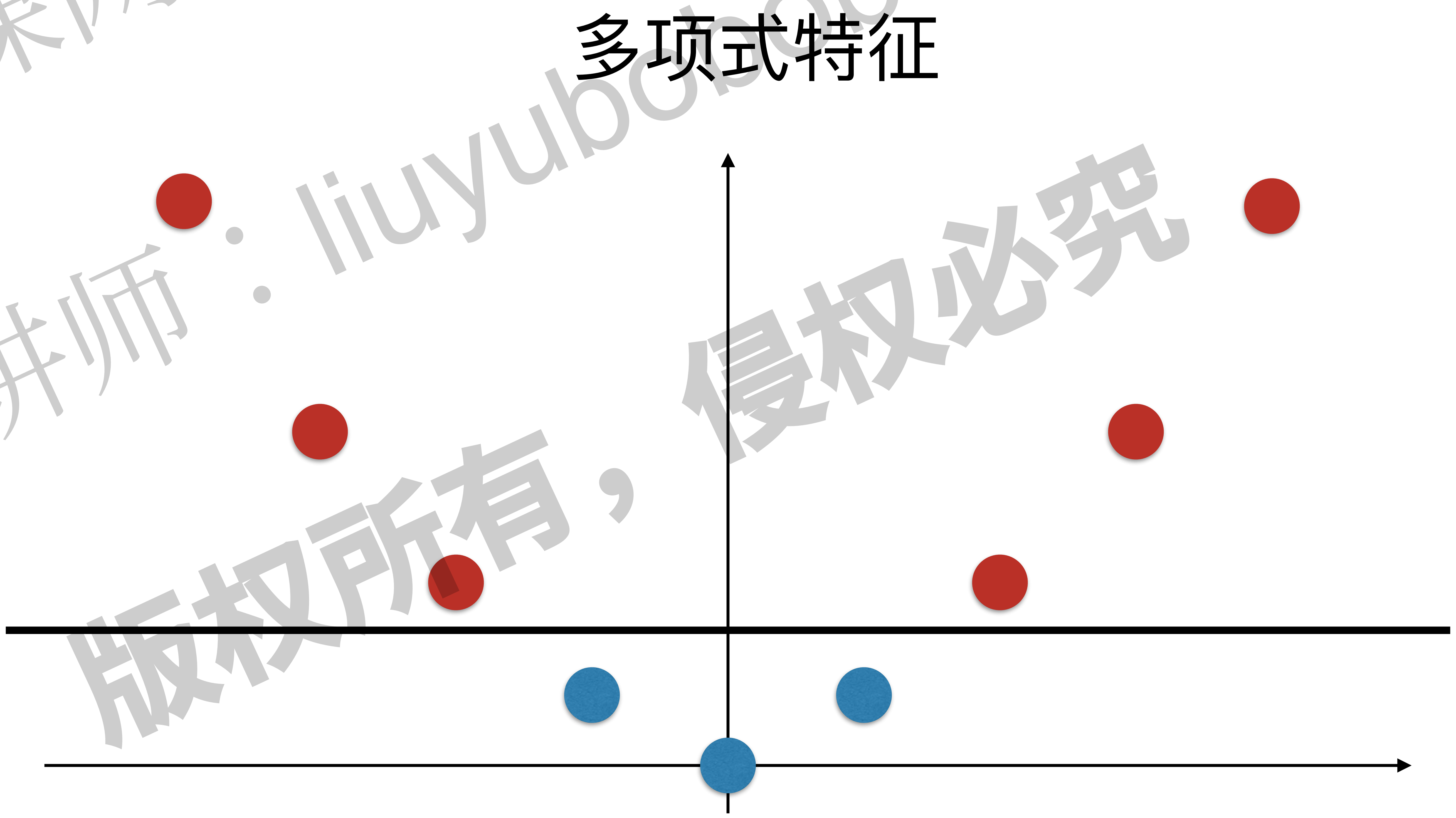
依靠升维使得原本线性不可分的数据线性可分

多项式特征



多项式特征

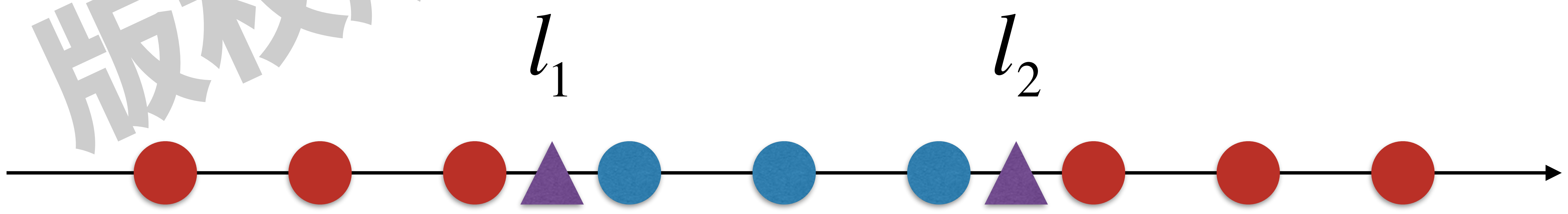
依靠升维使得原本线性不可分的数据线性可分



高斯核

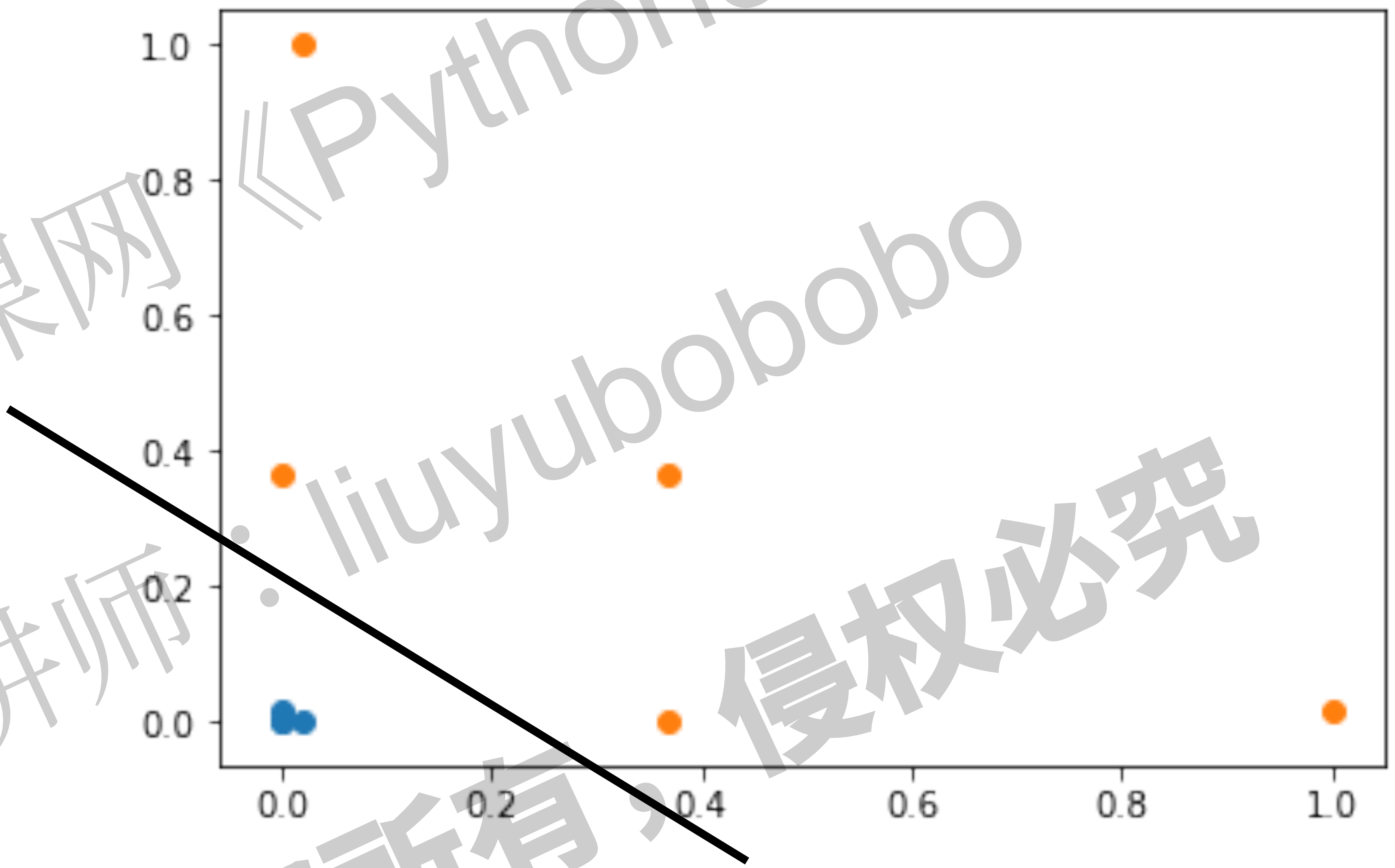
$$K(x,y) = e^{-\gamma\|x-y\|^2}$$

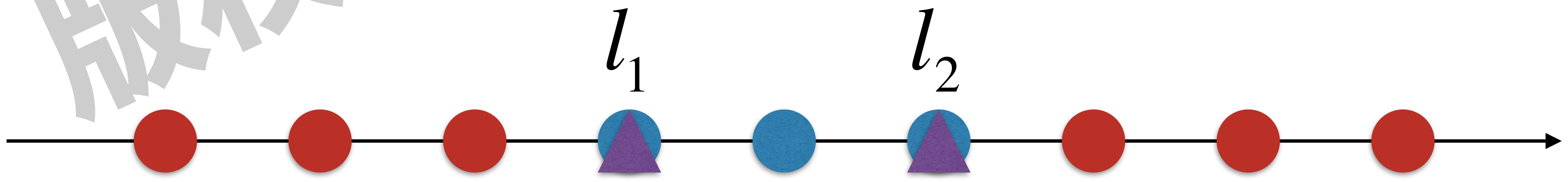
$$x \rightarrow (e^{-\gamma\|x-l_1\|^2}, e^{-\gamma\|x-l_2\|^2})$$



实践：模拟高斯核函数

高斯核





高斯核

$$x \rightarrow (e^{-\gamma\|x-l_1\|^2}, e^{-\gamma\|x-l_2\|^2})$$

$$K(x,y) = e^{-\gamma\|x-y\|^2}$$

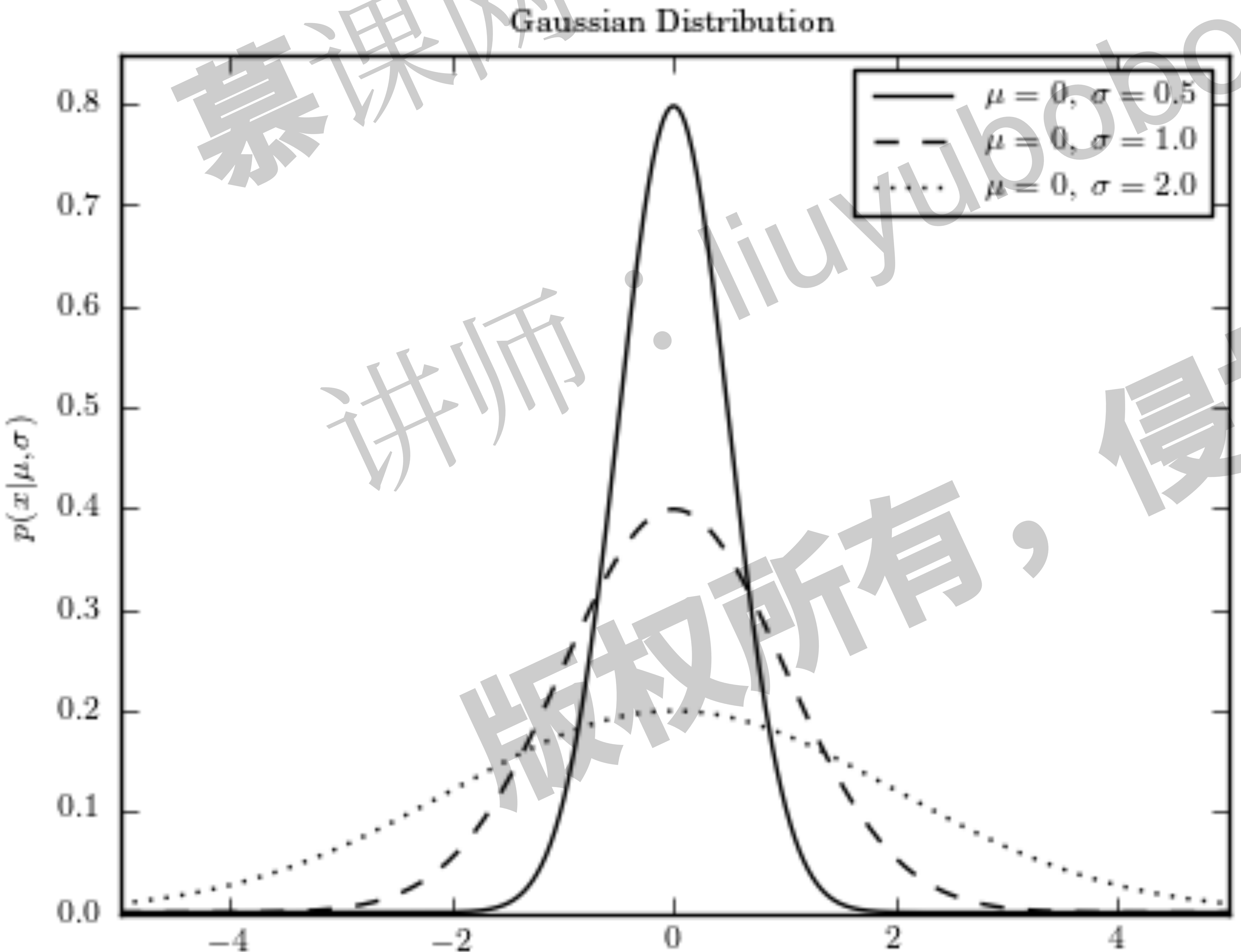
高斯核：对于每一个数据点都是landmark

m*n的数据映射成了m*m的数据

scikit-learn中的高斯核函数

高斯核

$$K(x,y) = e^{-\gamma\|x-y\|^2}$$



高斯函数 $$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

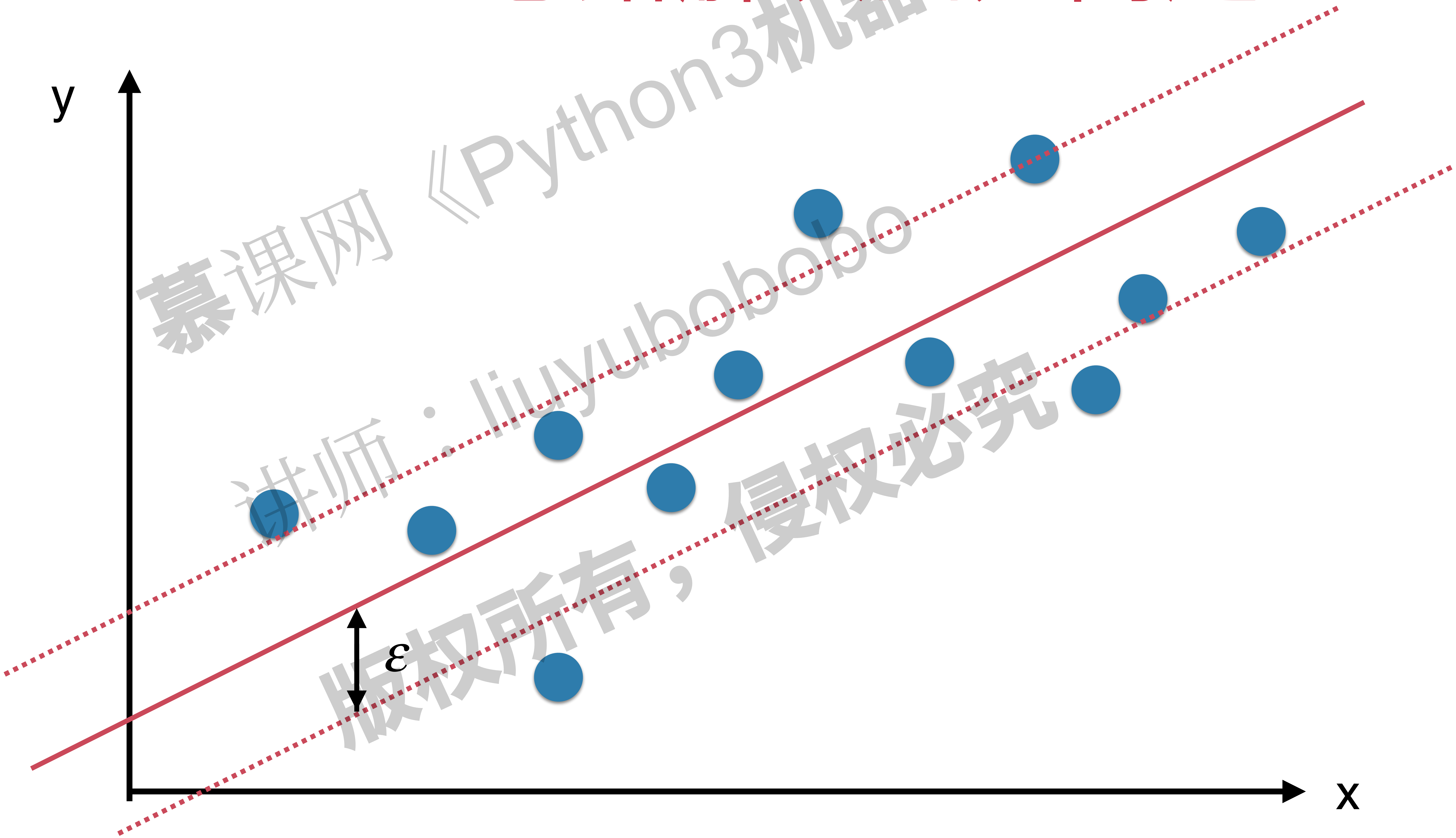
gamma越大，高斯分布越窄；

gamma越小，高斯分布越宽。

实践：scikit-learn中的高斯核函数

SVM思想解决回归问题

SVM思路解决回归问题



实践：SVM解决回归问题

其他

欢迎大家关注我的个人公众号：是不是很酷

Python 3 玩儿转机器学习

liuyubobobo