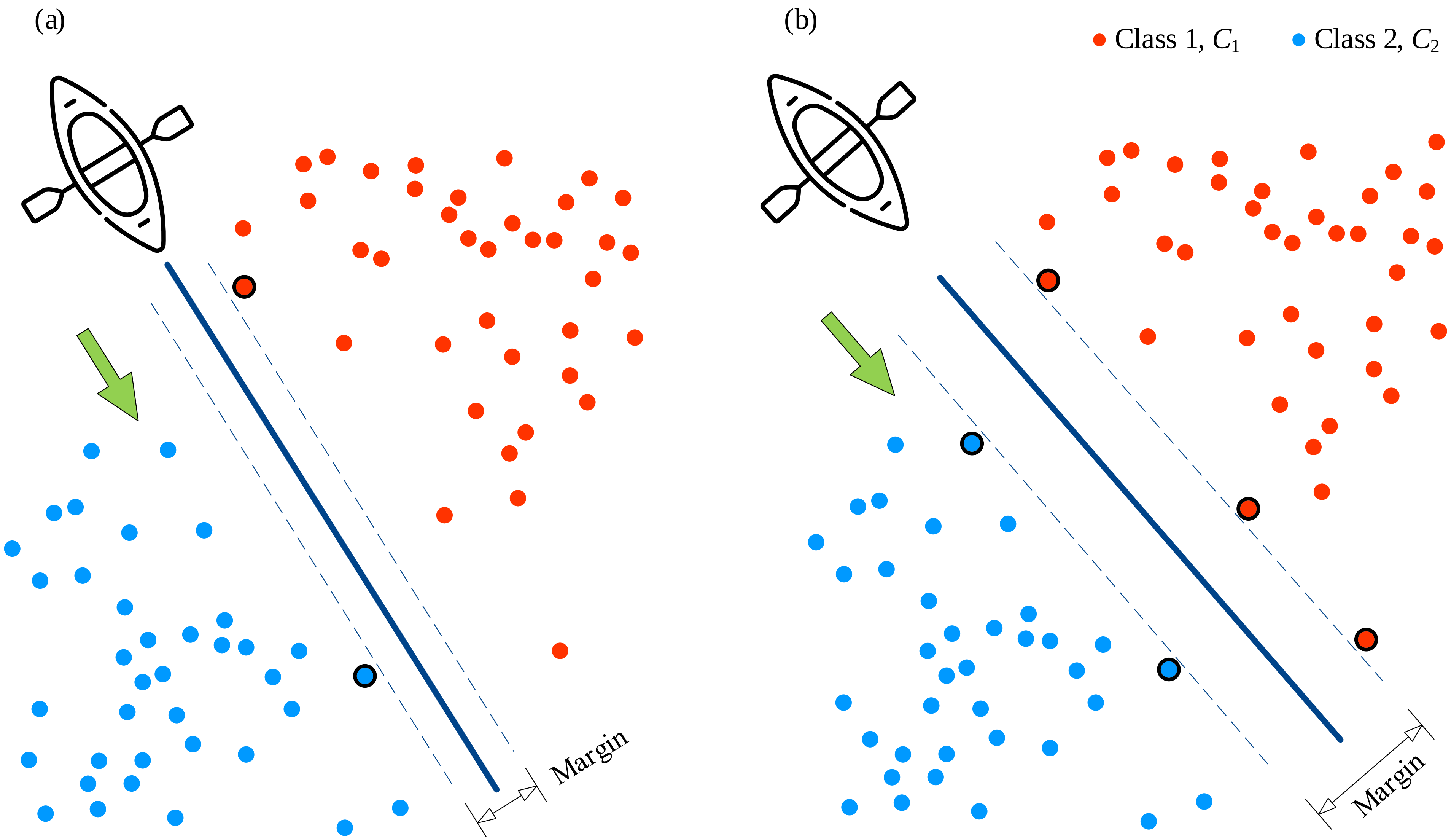
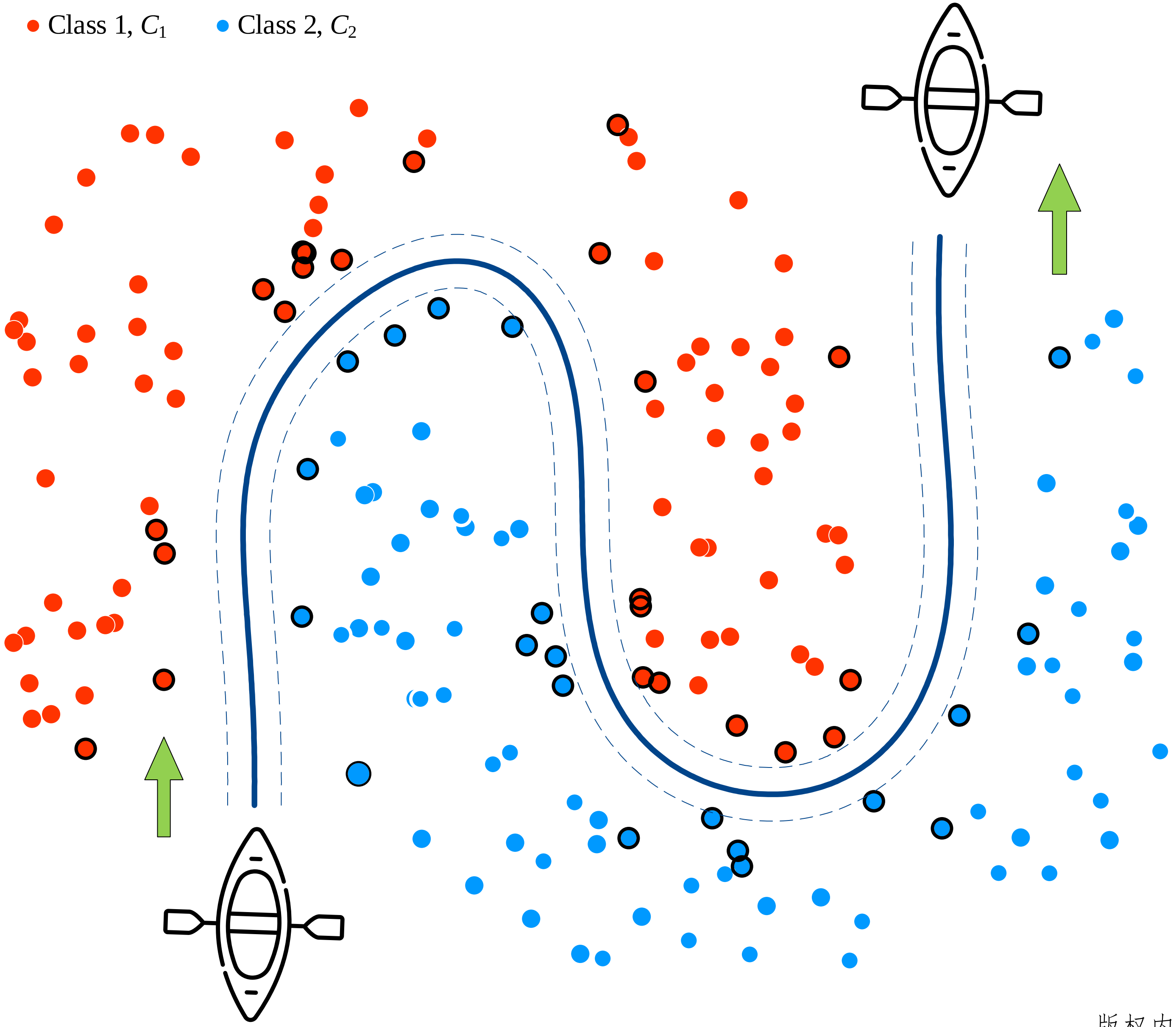
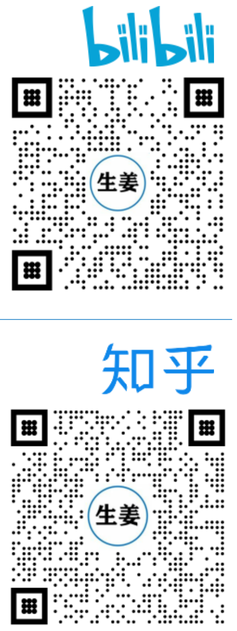
支持向量机原理



bilibili
知乎
生姜
版权内容，请勿商用；转载引用，请注明出处

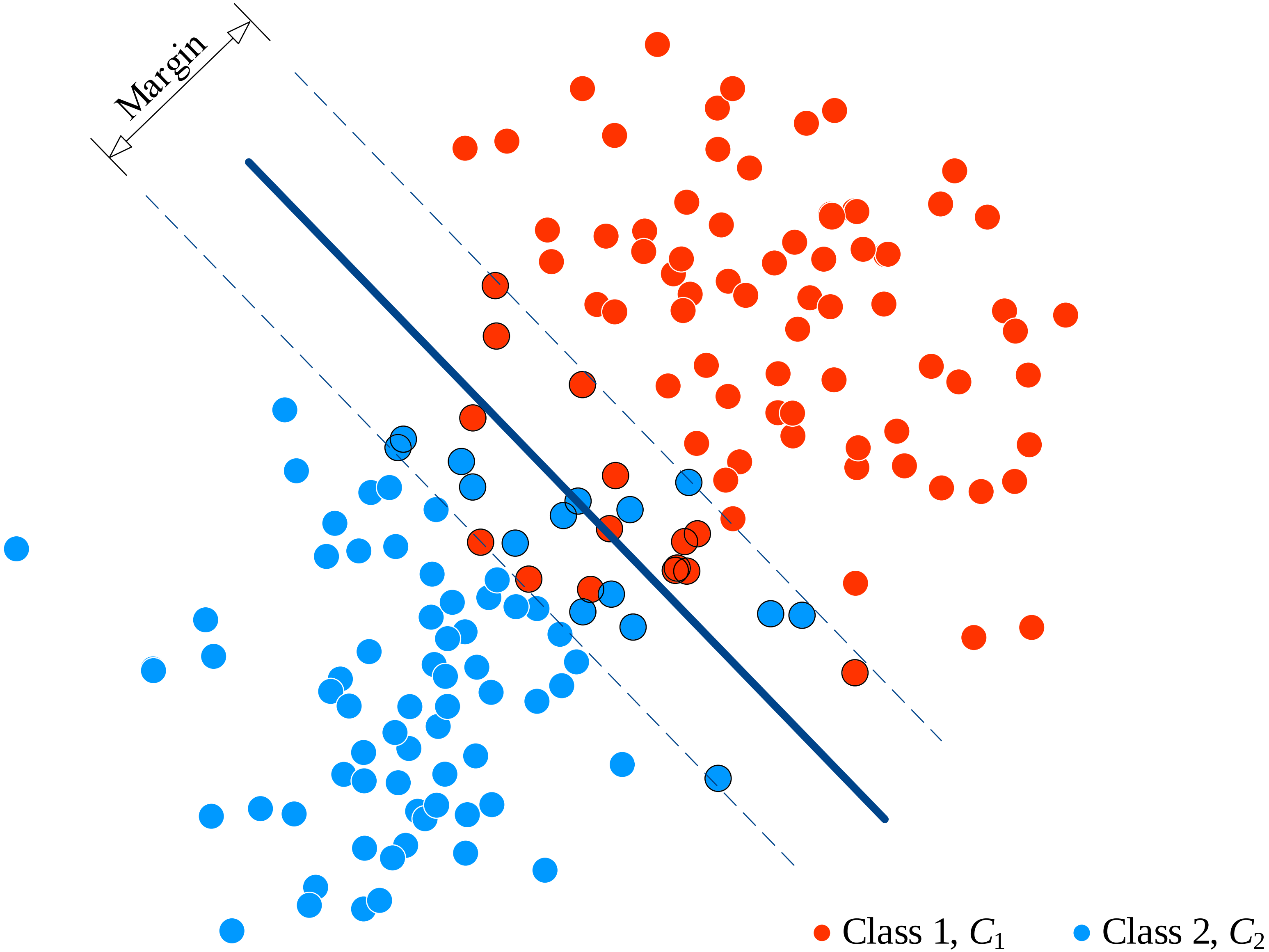
线性不可分数据





版权内容，请勿商用；转载引用，请注明出处

软间隔

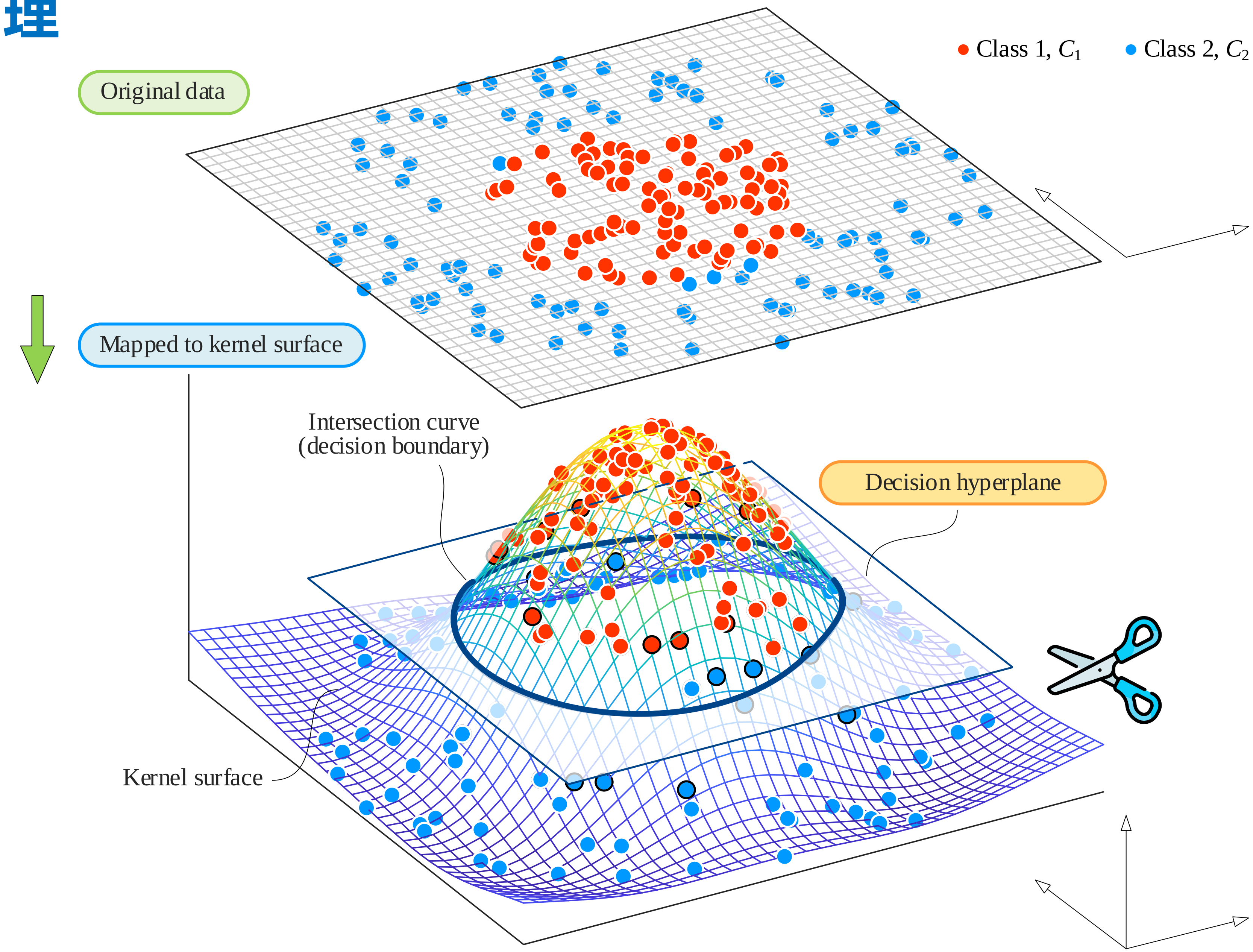


bilibili

知乎

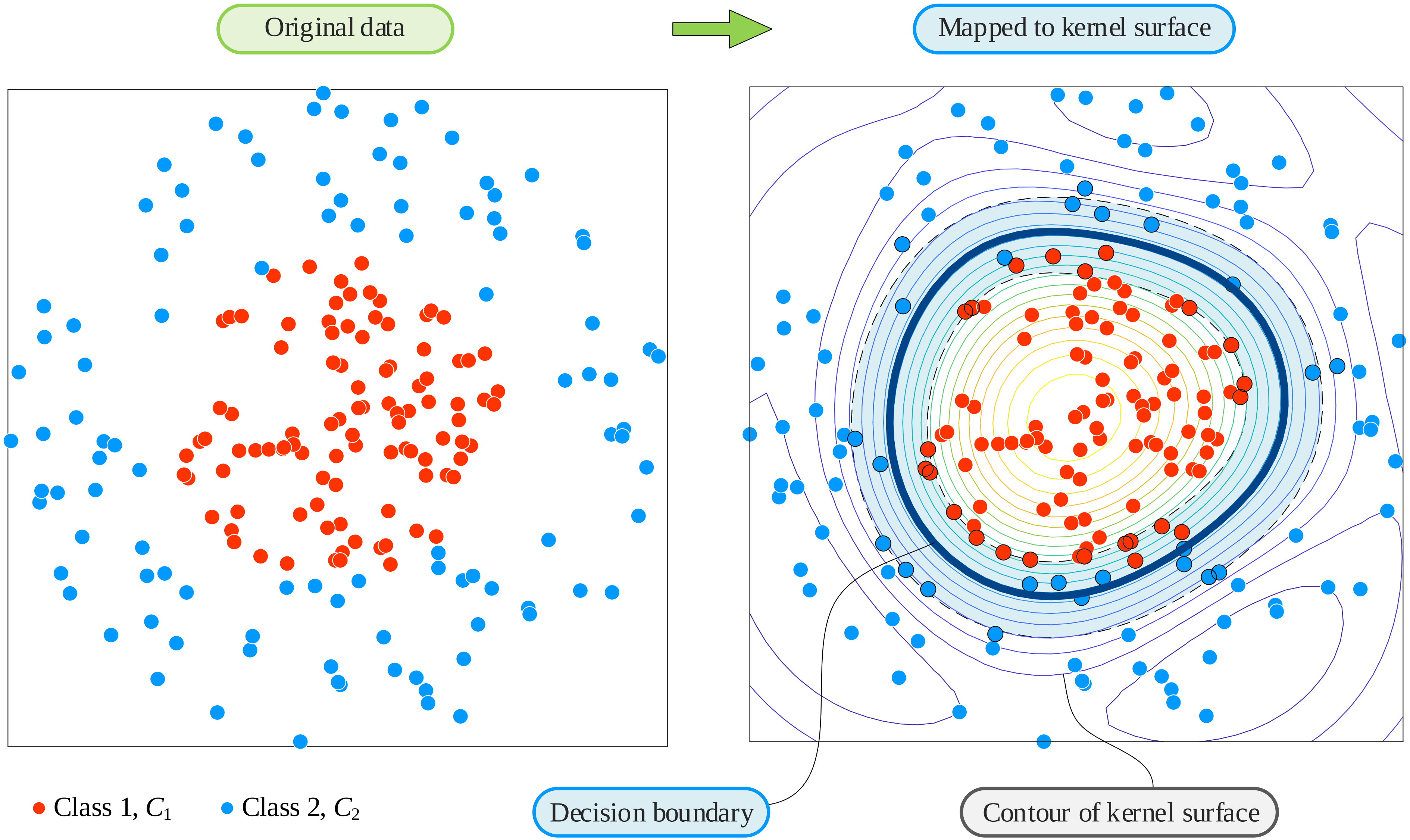
版权内容，请勿商用；转载引用，请注明出处

核技巧原理



bilibili
生姜
知乎
生姜
版权内容，请勿商用；转载引用，请注明出处

核技巧配合软间隔

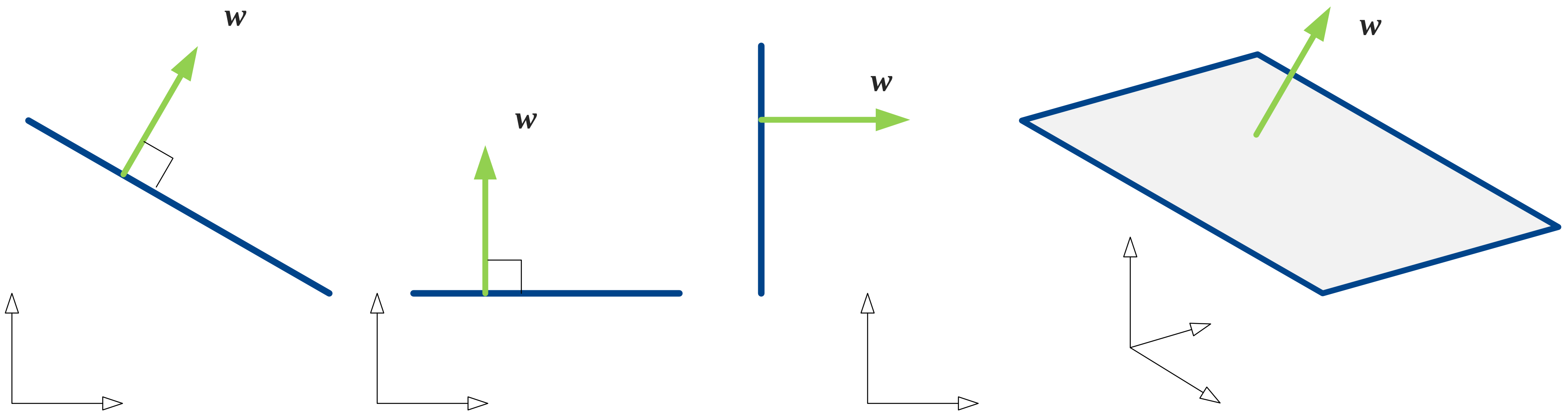


bilibili
生姜
知乎
生姜
版权内容，请勿商用；转载引用，请注明出处

硬间隔

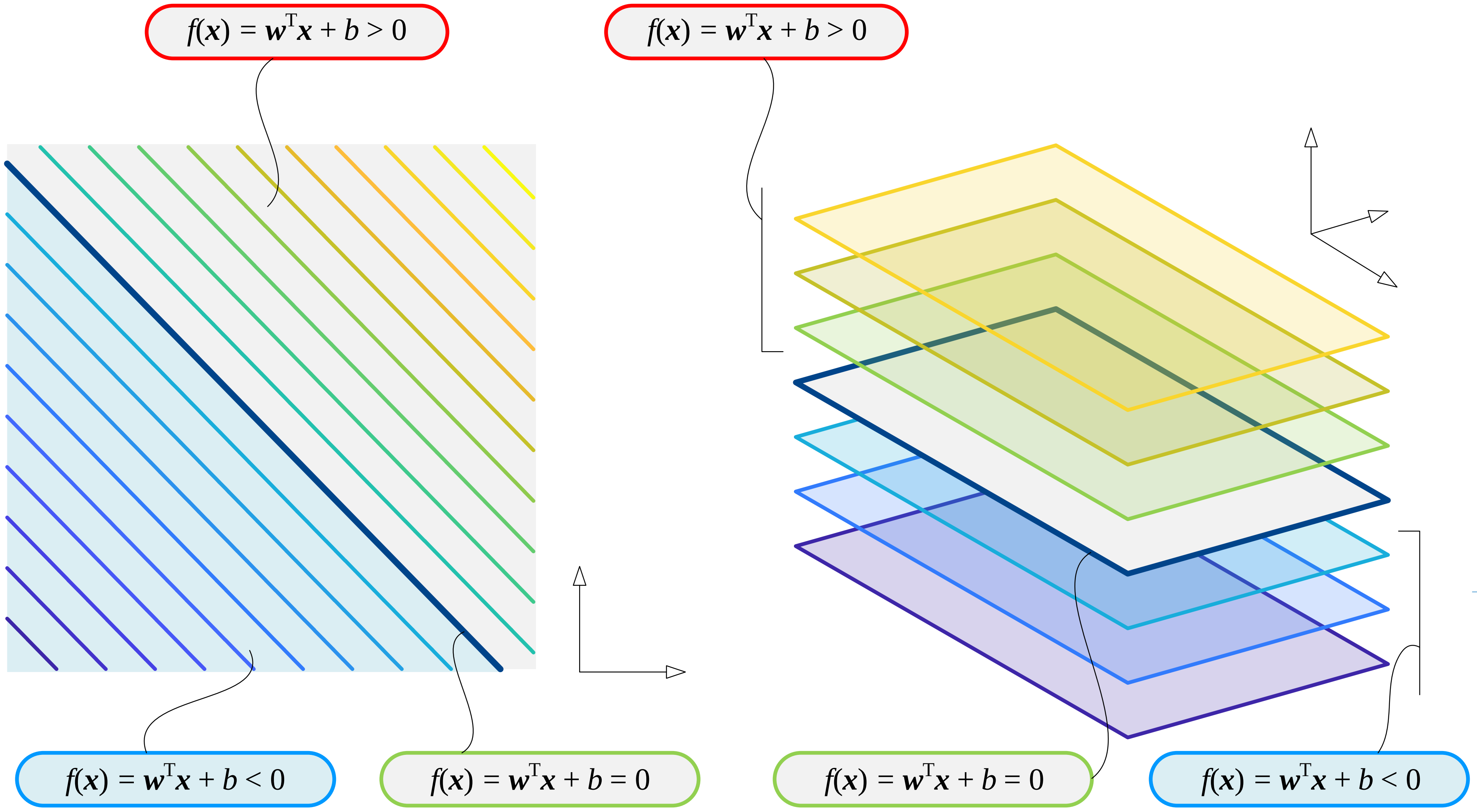
$$f(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0$$

$$w_1 x_1 + w_2 x_2 + b = 0$$



bilibili
知乎
版权内容，请勿商用；转载引用，请注明出处

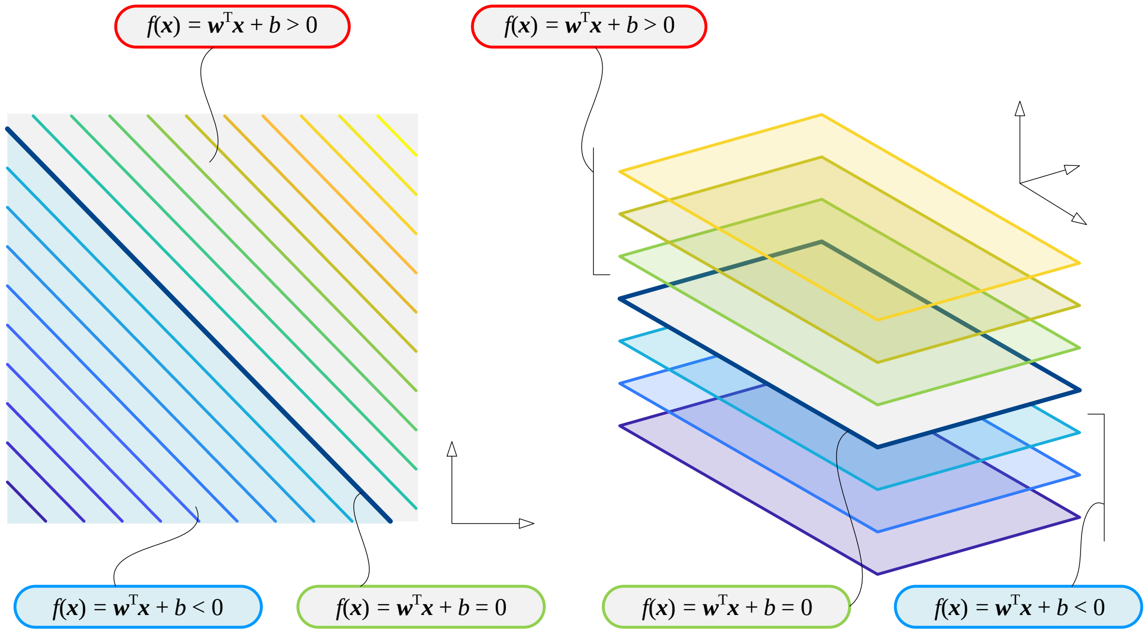
决策边界分割空间



bilibili
知乎
版权内容，请勿商用；转载引用，请注明出处

二分类决策函数

$$p(\boldsymbol{q}) = \operatorname{sign}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q} + b\right)$$



bilibili

生姜

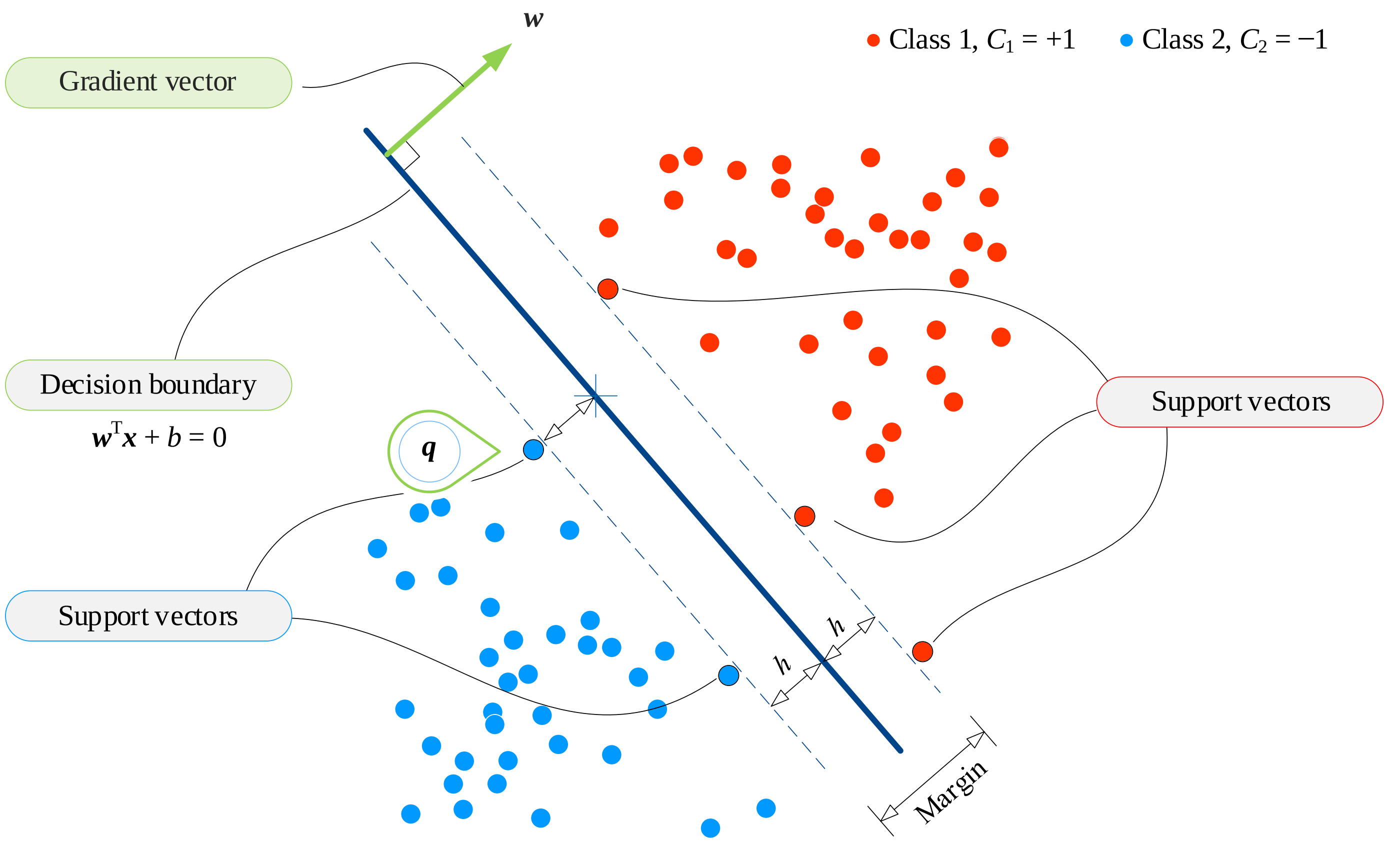
知乎

生姜

版权内容，请勿商用；转载引用，请注明出处

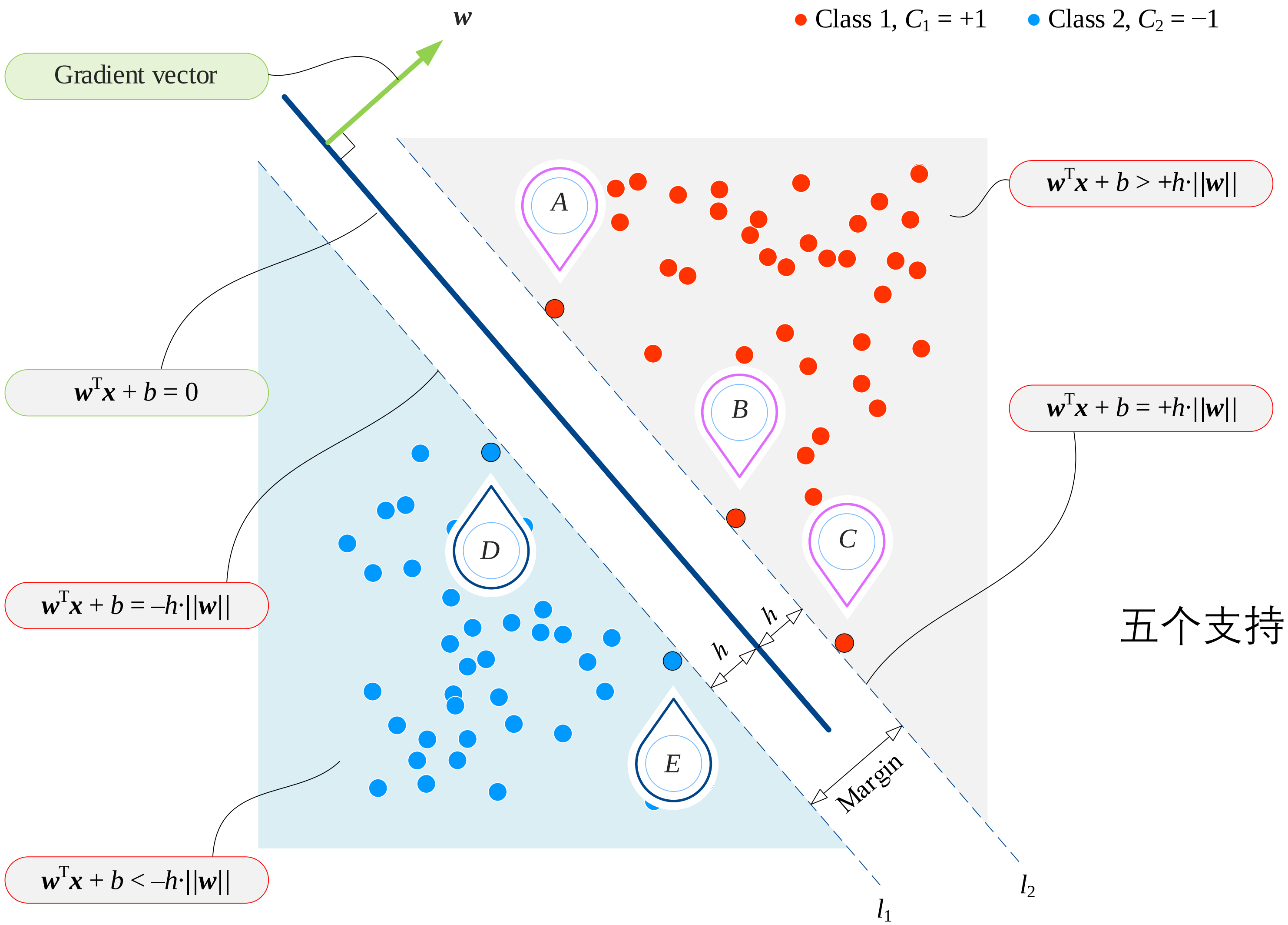
硬间隔SVM处理二分类问题

$$d=\frac{\left|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}+b\right|}{\|\boldsymbol{w}\|}=\frac{\left|\boldsymbol{w}\cdot\boldsymbol{q}+b\right|}{\|\boldsymbol{w}\|}\qquad d=\frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{q}+b}{\|\boldsymbol{w}\|}=\frac{\boldsymbol{w}\cdot\boldsymbol{q}+b}{\|\boldsymbol{w}\|}$$



bilibili
生姜
知乎
生姜
版权内容，请勿商用；转载引用，请注明出处

硬间隔、决策边界和支持向量之间关系



$$\begin{cases} \dfrac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b}{\|\boldsymbol{w}\|h} \geq +1, & y=+1 \\ \dfrac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b}{\|\boldsymbol{w}\|h} \leq -1, & y=-1 \end{cases}$$

$$\frac{\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b\right)y}{\|\boldsymbol{w}\|h} \geq 1$$

五个支持向量点 (A、B、C、D和E)，满足

$$\frac{\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b\right)y}{\|\boldsymbol{w}\|h} = 1$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

简化运算

$$\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b\right)y\geq 1$$

$$\left(\boldsymbol{w}\cdot\boldsymbol{x}+b\right)y\geq 1$$

$$\left\|\boldsymbol{w}\right\|h=1$$

间隔上下边界的解析式

$$\begin{cases}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b=+1\\ \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b=-1\end{cases}$$

间隔宽度2h可以用$\boldsymbol{w}$表达

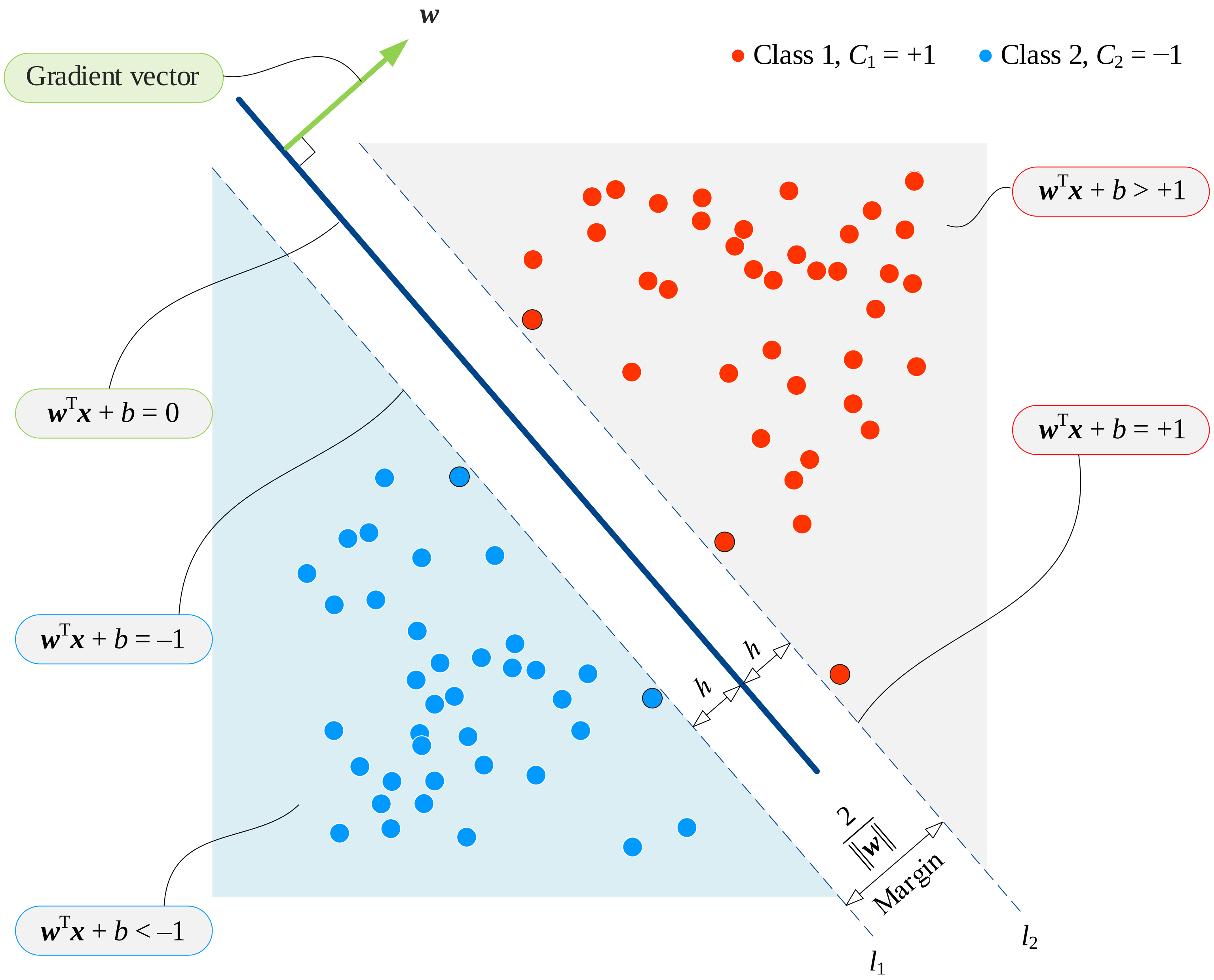
$$2h=\frac{2}{\left\|\boldsymbol{w}\right\|}$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

硬间隔、决策边界和支持向量之间关系



$$2h = \frac{2}{\|\mathbf{w}\|}$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

构造优化问题

$$\arg\max_{\boldsymbol{w},b} \frac{2}{\|\boldsymbol{w}\|}$$

$$\text{subject to} \quad \left(\boldsymbol{w}\cdot\boldsymbol{x}^{(i)}+b\right)y^{(i)} \geq 1, \quad i=1,2,3,\ldots,n$$

$$\arg\min_{\boldsymbol{w},b} \frac{\|\boldsymbol{w}\|^2}{2} = \frac{\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}}{2} = \frac{\boldsymbol{w}\cdot\boldsymbol{w}}{2}$$

$$\text{subject to} \quad \left(\boldsymbol{w}\cdot\boldsymbol{x}^{(i)}+b\right)y^{(i)} \geq 1, \quad i=1,2,3,\ldots,n$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

构造拉格朗日函数

$$L(\boldsymbol{w}, b, \lambda) = \frac{\boldsymbol{w} \cdot \boldsymbol{w}}{2} + \sum_{i=1}^{n} \lambda_i \left(1 - y^{(i)}\left(\boldsymbol{w} \cdot \boldsymbol{x}^{(i)} + b\right)\right)$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

偏导

$$\begin{cases} \dfrac{\partial L(\boldsymbol{w}, b, \lambda)}{\partial \boldsymbol{w}} = \boldsymbol{w} - \sum_{i=1}^{n} \lambda_i y^{(i)} \boldsymbol{x}^{(i)} = 0 \\ \dfrac{\partial L(\boldsymbol{w}, b, \lambda)}{\partial b} = \sum_{i=1}^{n} \lambda_i y^{(i)} = 0 \end{cases}$$

$$\begin{cases} \boldsymbol{w} = \sum_{i=1}^{n} \lambda_i y^{(i)} \boldsymbol{x}^{(i)} \\ \sum_{i=1}^{n} \lambda_i y^{(i)} = 0 \end{cases}$$

bilibili

生姜

知乎

生姜

版权内容，请勿商用；转载引用，请注明出处

简化拉格朗日函数

$$L(\lambda)=\sum_{i=1}^{n}\lambda_i-\frac{\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_i\lambda_j y^{(i)}y^{(j)}\left(\boldsymbol{x}^{(i)}\cdot\boldsymbol{x}^{(j)}\right)}{2}$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

拉格朗日对偶问题

$$\underset{\lambda}{\arg\min}\ \sum_{i=1}^{n}\lambda_i-\frac{\sum_{j=1}^{n}\sum_{i=1}^{n}\lambda_i\lambda_j y^{(i)}y^{(j)}\left(\boldsymbol{x}^{(i)}\cdot\boldsymbol{x}^{(j)}\right)}{2}$$

$$\text{subject to}\ \begin{cases}\sum_{i=1}^{n}\lambda_i y^{(i)}=0\\ \lambda_i\geq 0,\quad i,j=1,2,3,\dots,n\end{cases}$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

$$f(\boldsymbol{x}) = \boldsymbol{w} \cdot \boldsymbol{x} + b = \underbrace{\left(\sum_{i=1}^{n} \lambda_i y^{(i)} \boldsymbol{x}^{(i)} \right)}_{\text{Coefficients}} \cdot \boldsymbol{x} + b = 0$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

分类决策函数

$$p(\boldsymbol{x}) = \operatorname{sign}(\boldsymbol{w}\cdot\boldsymbol{x}+b) = \operatorname{sign}\left(\underbrace{\left(\sum_{i=1}^{n}\lambda_i y^{(i)}\boldsymbol{x}^{(i)}\right)}_{\text{Coefficients}}\cdot\boldsymbol{x}+b\right)$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

支持向量

$$\begin{bmatrix} w_1 & w_2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \end{bmatrix}^{\mathrm{T}} + b = 0 \quad \Rightarrow \quad w_1 x_1 + w_2 x_2 + b = 0$$

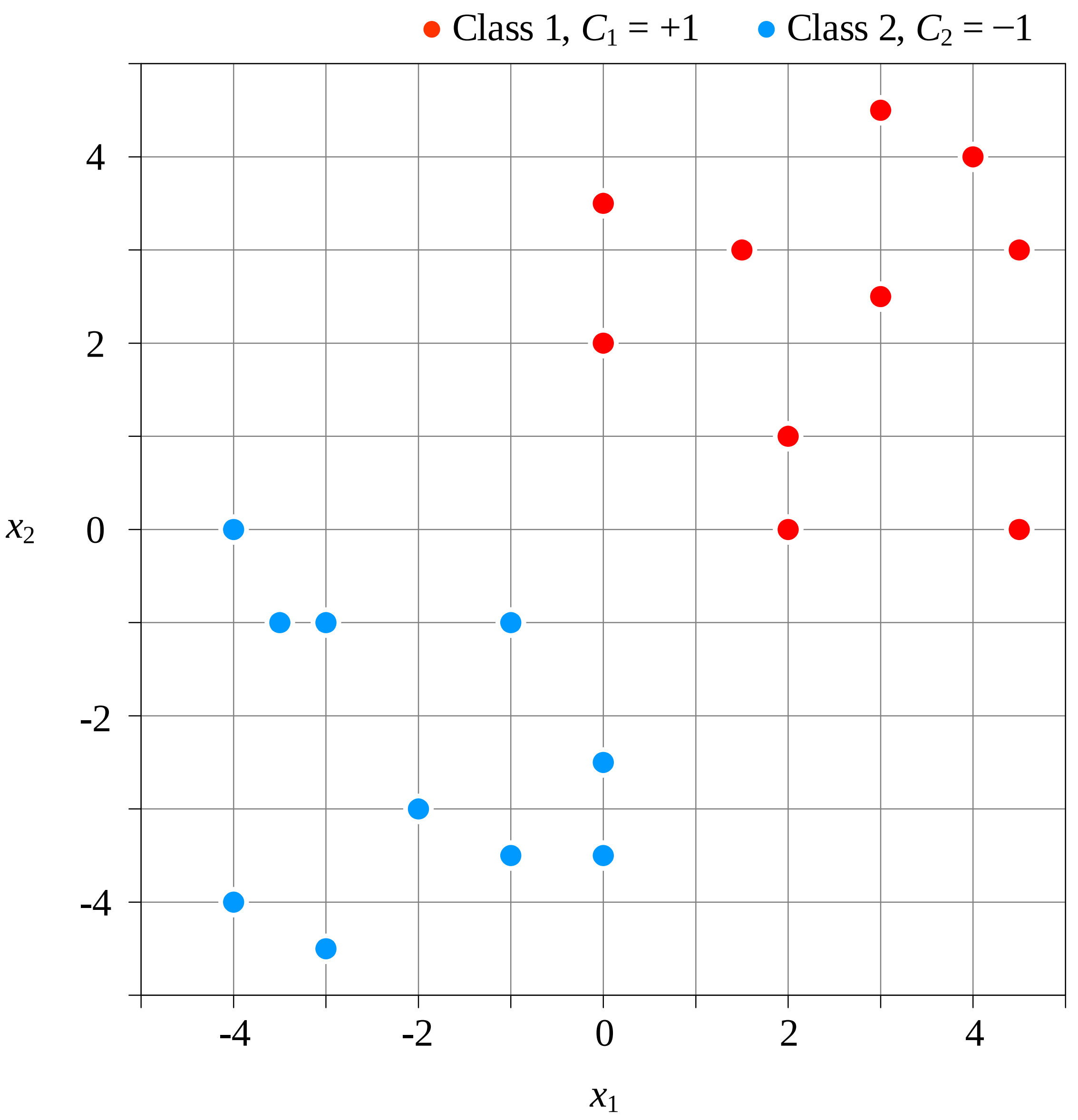
$$w_1 x_1 + w_2 x_2 + b = 0 \quad \Rightarrow \quad x_2 = -\frac{w_1}{w_2} x_1 - \frac{b}{w_2}$$

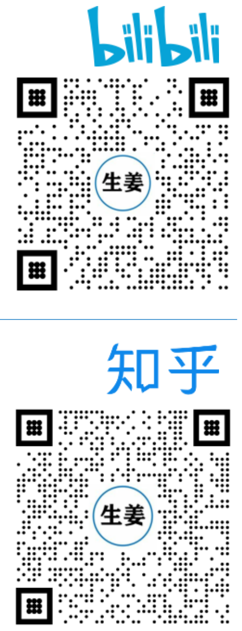
bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

样本数据点平面位置





版权内容，请勿商用；转载引用，请注明出处

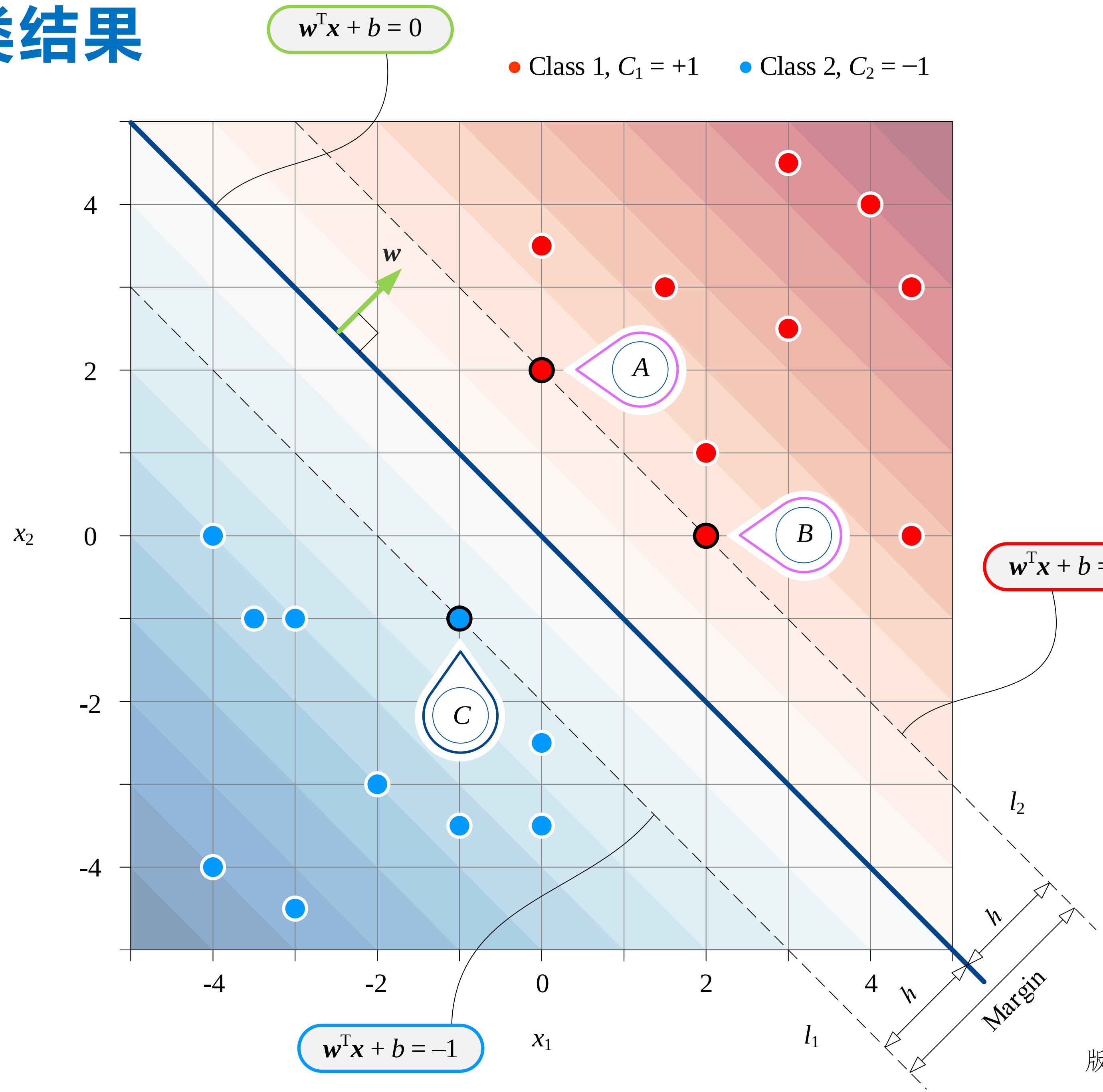
间隔上下界

$$w_1x_1 + w_2x_2 + b = 1 \quad \Rightarrow \quad x_2 = -\frac{w_1}{w_2}x_1 - \frac{b-1}{w_2}$$

$$w_1x_1 + w_2x_2 + b = -1 \quad \Rightarrow \quad x_2 = -\frac{w_1}{w_2}x_1 - \frac{b+1}{w_2}$$

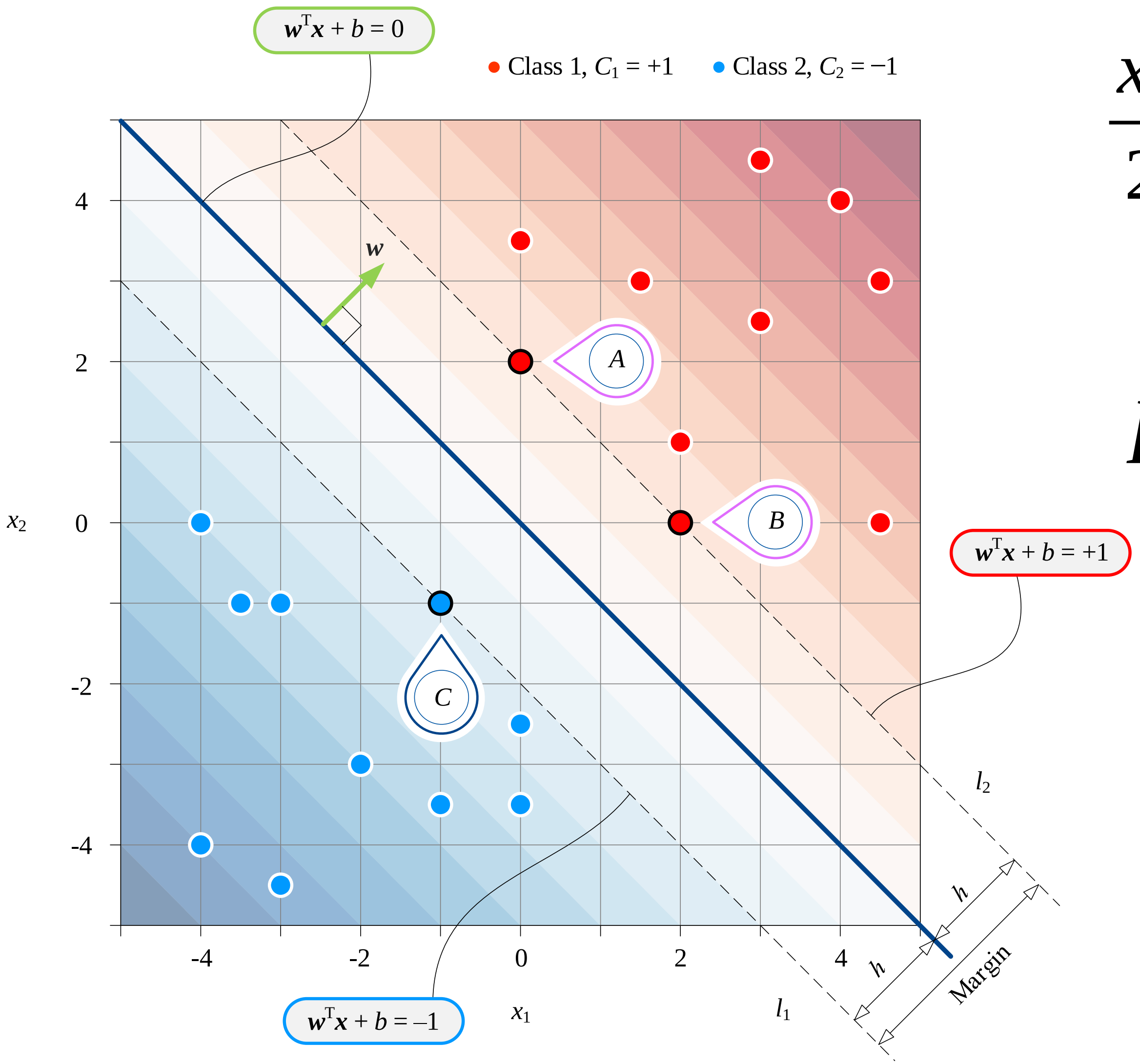
bilibili
知乎
版权内容，请勿商用；转载引用，请注明出处

硬间隔分类结果



版权内容，请勿商用；转载引用，请注明出处

决策边界



$$\frac{x_1}{2} + \frac{x_2}{2} = 0 \quad \Rightarrow \quad x_1 + x_2 = 0 \quad \Rightarrow \quad x_2 = -x_1$$

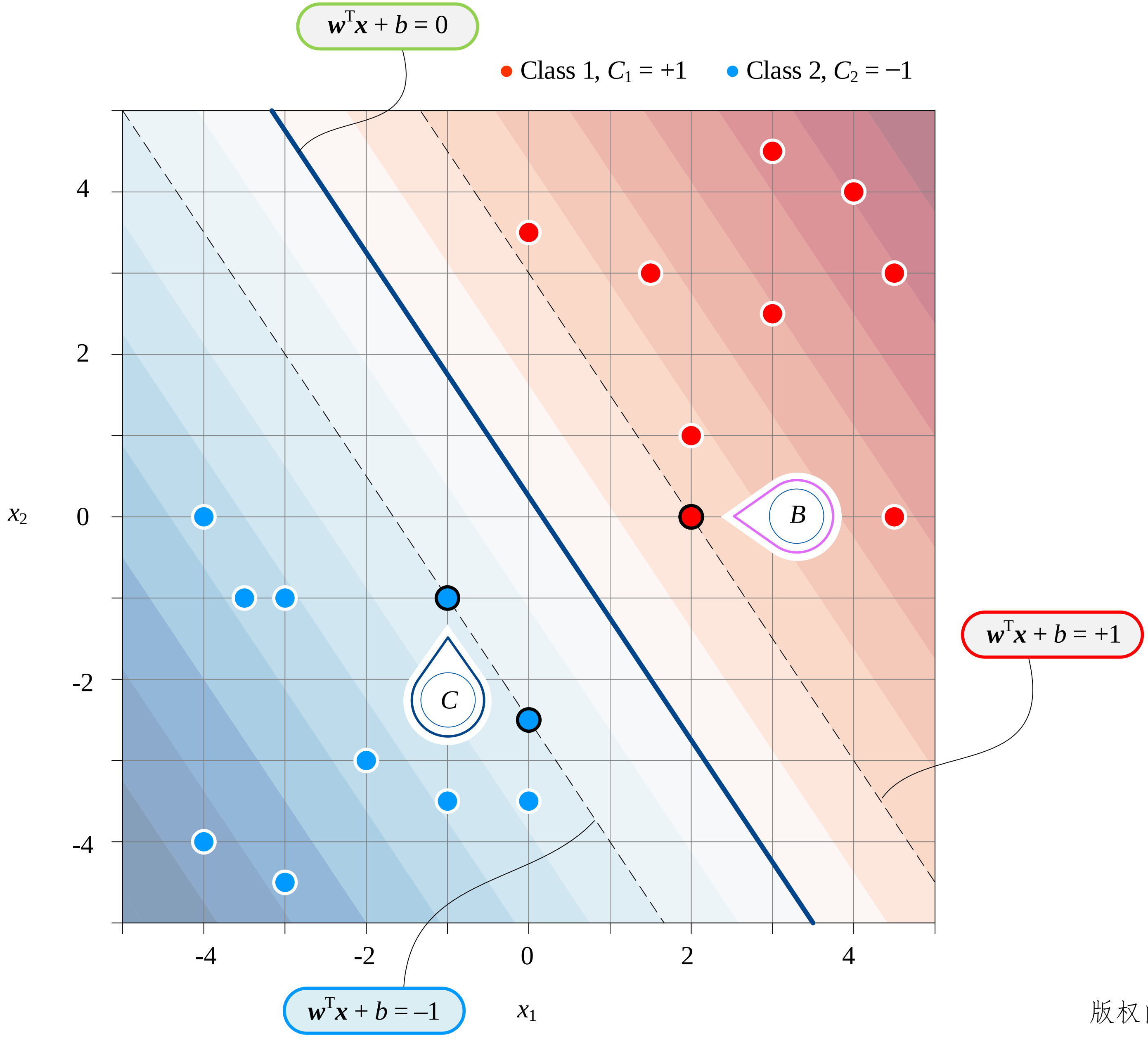
$$p(x_1, x_2) = \mathrm{sign}(x_1 + x_2)$$

bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

删除点A后硬间隔SVM分类结果

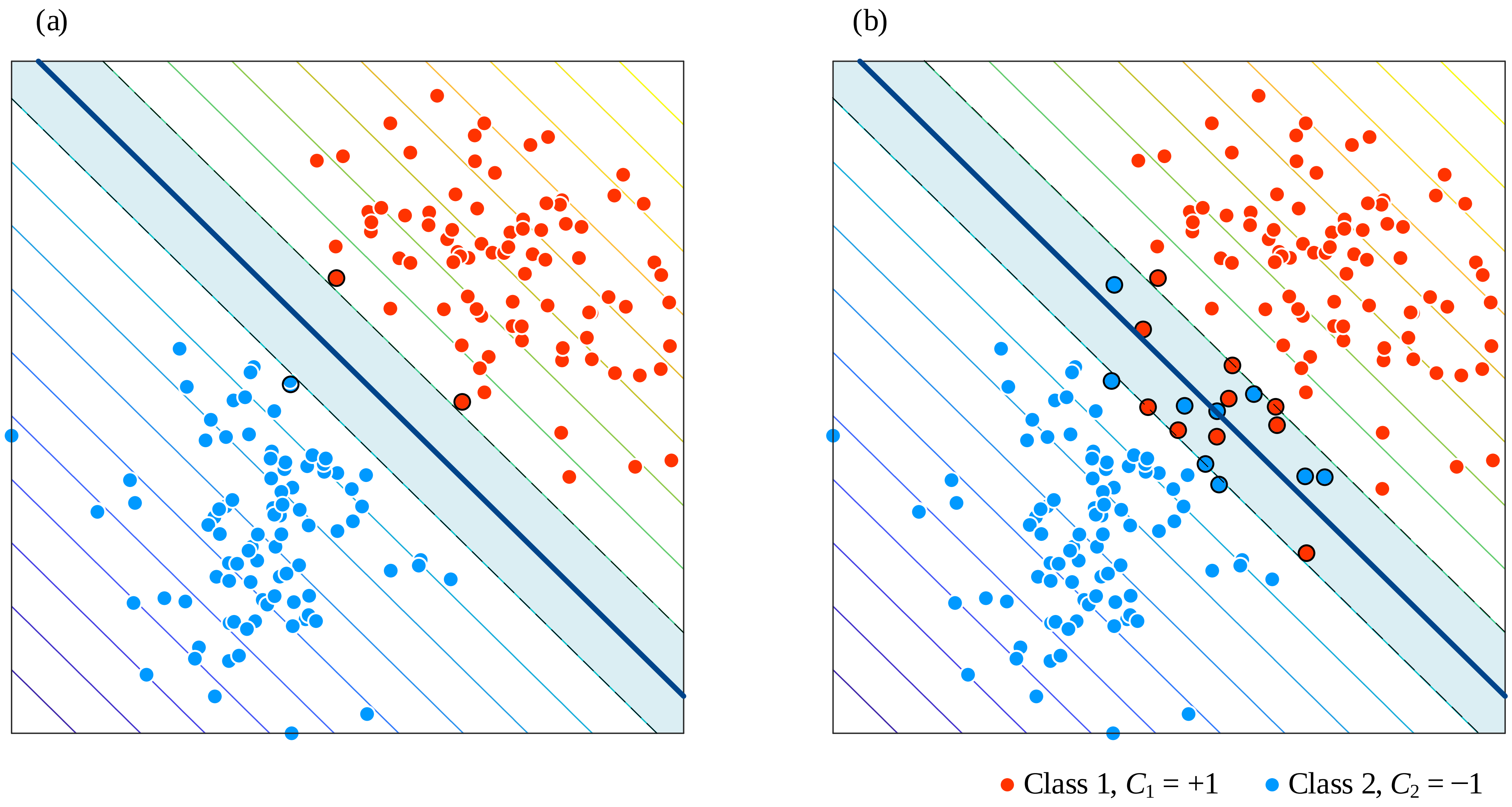


bilibili

知乎

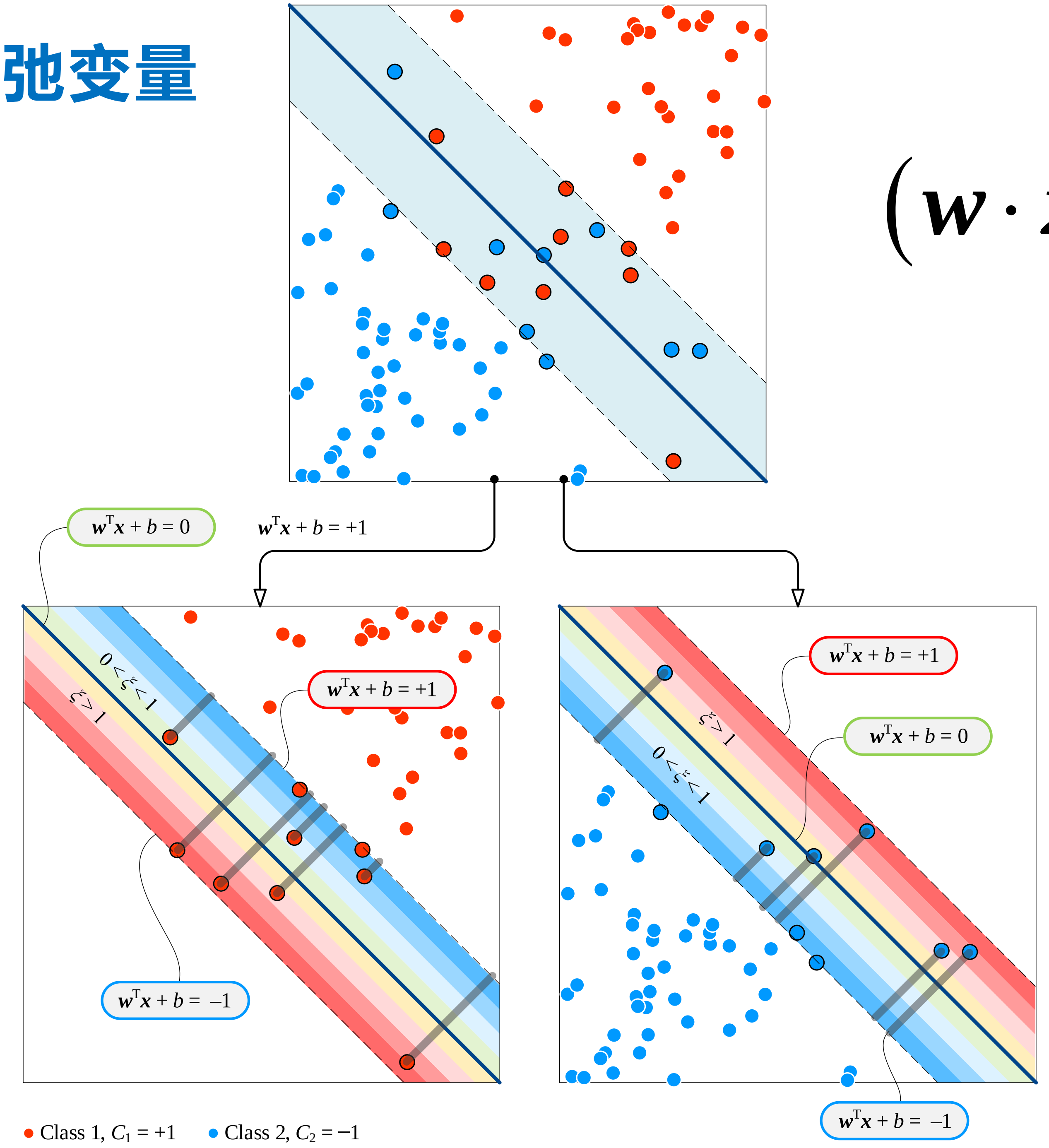
版权内容，请勿商用；转载引用，请注明出处

比较硬间隔和软间隔



bilibili
知乎
生姜
生姜
版权内容，请勿商用；转载引用，请注明出处

引入松弛变量



$$(\boldsymbol{w} \cdot \boldsymbol{x} + b) y \geq 1 - \xi$$

bilibili
知乎
版权内容，请勿商用；转载引用，请注明出处

构造软间隔SVM优化问题

$$\underset{\boldsymbol{w},b}{\arg\min} \quad \frac{\boldsymbol{w}\cdot\boldsymbol{w}}{2} + C\sum_{i=1}^{n}\xi_i$$

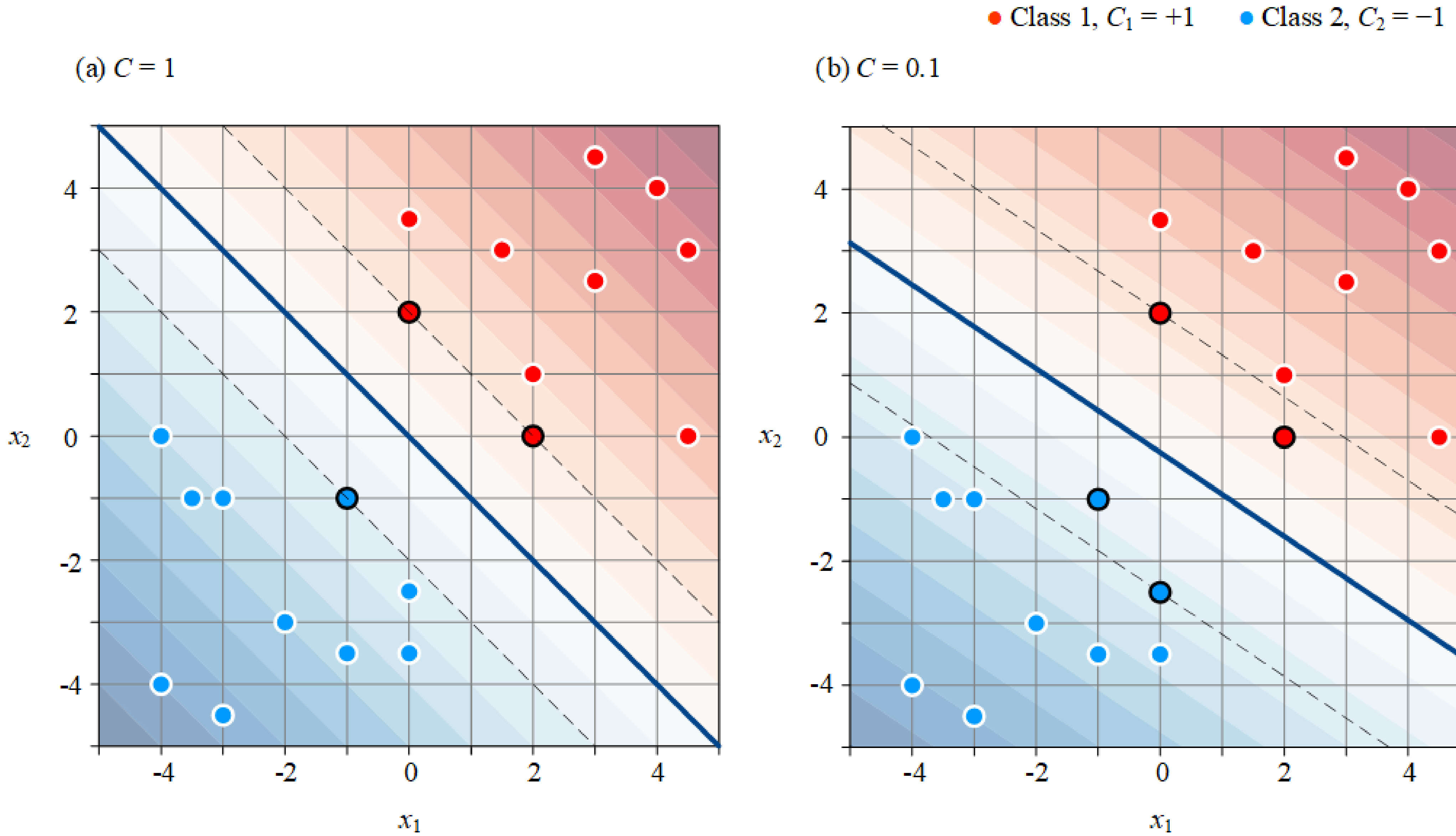
$$\text{subject to} \quad \begin{cases} y^{(i)}\left(\boldsymbol{w}\cdot\boldsymbol{x}^{(i)}+b\right) \geq 1-\xi_i, \quad i=1,2,3,\ldots,n \\ \xi_i \geq 0 \end{cases}$$

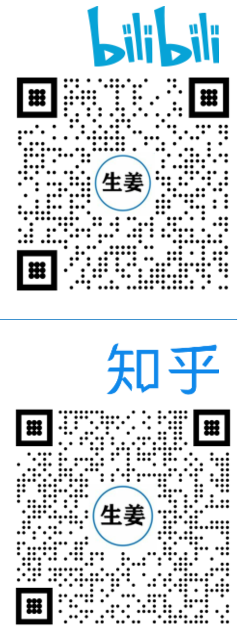
bilibili

知乎

版权内容，请勿商用；转载引用，请注明出处

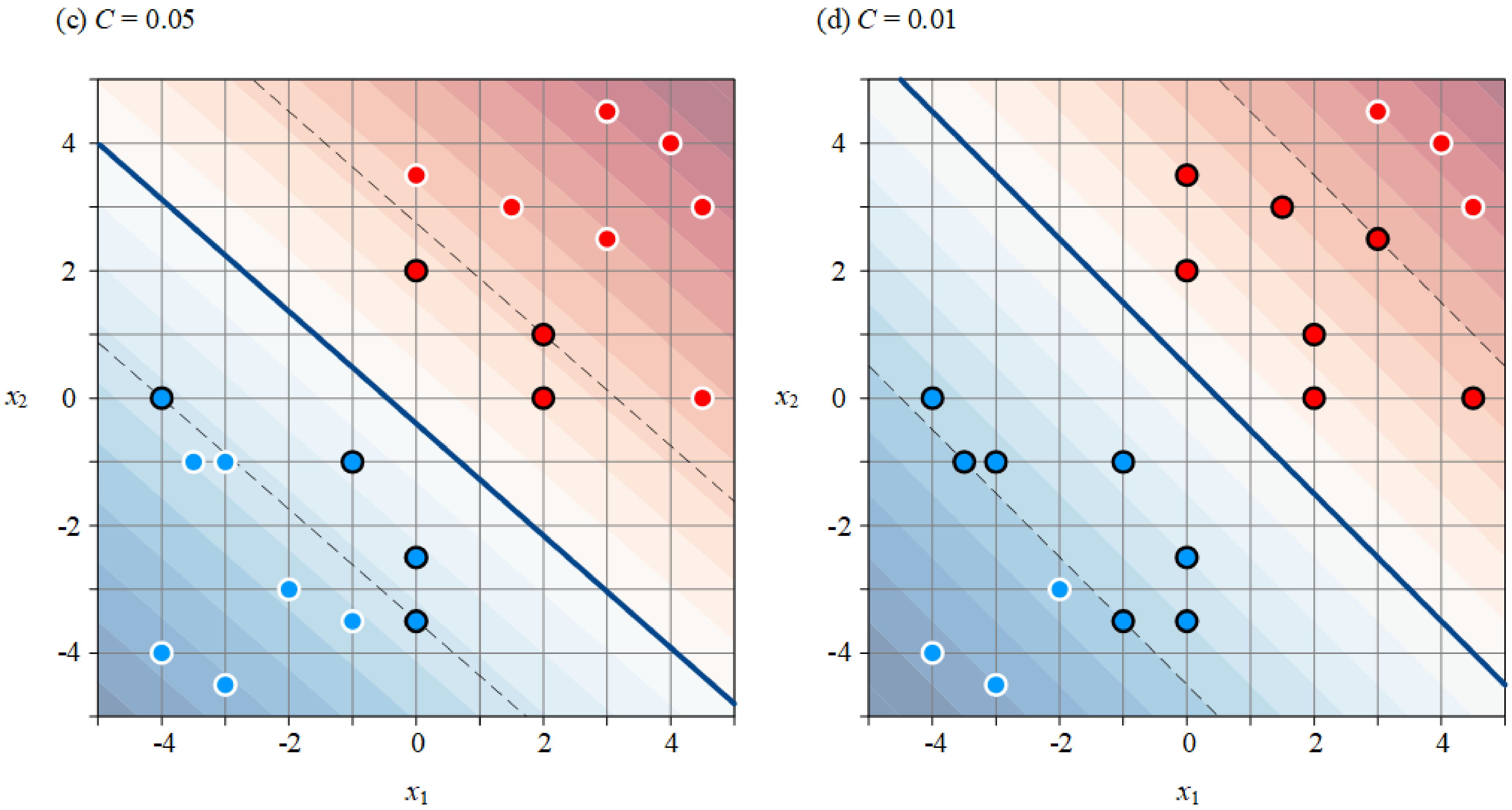
惩罚因子对软间隔宽度和决策边界影响





版权内容，请勿商用；转载引用，请注明出处

惩罚因子对软间隔宽度和决策边界影响



bilibili
知乎
版权内容，请勿商用；转载引用，请注明出处