

2025 模式识别 作业三

人工智能学院 221300079 王俊童

2025.4.15

221300079 王俊童, 人工智能学院

1 问题一

a. 用 python 完成, 结果如下:

10 维样本的 12 范数统计:
均值: 2.8871
最小值: 2.0136
最大值: 4.0669

b. 同样, 用 python 可以得到:

100 维样本的 12 范数统计:
均值: 9.9181
最小值: 8.8613
最大值: 11.0844

1000 维样本的 12 范数统计:
均值: 31.7088
最小值: 30.6157
最大值: 32.9135

10000 维样本的 12 范数统计:
均值: 99.9484
最小值: 98.8400
最大值: 101.5054

100000 维样本的 12 范数统计:
均值: 316.4926
最小值: 315.0430
最大值: 318.7029

c. 可以提出一个猜想是: 随着维度 d 变高, 样本的 12 范数更加接近 \sqrt{d} 的取值。他们的相对波动减少
可以理解为: 维度增加, 对于 d 维度的标准高斯分布, 12 范数越来越集中于 \sqrt{d} 附近。

相对波动:

$$relative\ variance = \frac{\max - \min}{mean}$$

100 维样本的 l2 范数统计：
波动：0.2241

1000 维样本的 l2 范数统计：
波动：0.0725

10000 维样本的 l2 范数统计：
波动：0.0267

100000 维样本的 l2 范数统计：
波动：0.0116

d. 如果用严谨的数学语言描述这个现象，可以解释为：

令 $X = (X_1, X_2, \dots, X_d)^\top \sim \mathcal{N}(0, I_d)$ ，且其 l2 范数满足 $\|X\|_2 = \sqrt{\sum_{i=1}^d X_i^2}$ ，其 l2 范数的平方服从自由度为 d 的 χ^2 分布。那么有以下形式的式子在概率意义下成立：

$$\frac{\|X\|_2}{\sqrt{d}} \xrightarrow{P} 1, \quad \text{when } d \rightarrow \infty$$

当然这是一个服从大数定律形式的，或者可以有一个 concentration 的形式：

$$P(|\frac{\|X\|_2}{\sqrt{d}} - 1| \geq \epsilon) \rightarrow 0, \quad \text{when } d \rightarrow \infty$$

e. 虽然集中不等式和大数定律忘的差不多了，但是可以知道这个东西服从一定的集中不等式，或者从大数定律来解释也是可以的，最后会有一个逼近性。

2 问题二

- 对于核函数的半正定性，有： $y^\top K_1 y \geq 0, y^\top K_2 y \geq 0$ ，对于 $K = K_1 + K_2$ ，有 $y^\top K y = y^\top K_1 y + y^\top K_2 y \geq 0$ ，所以合法
- 同理 $K = K_1 - K_2$ ，有 $y^\top K y = y^\top K_1 y - y^\top K_2 y$ ，但这个不是合法的核函数，可以有一个例子是如果 $K = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ 那么对于一个 $y = (1, 0)^\top$ ，有 $K_1 = 0, K_2 = x^\top y$ ，这个是不合法的，不正定。
- 对于 $y^\top K_1 y \geq 0$ ，且 $\alpha \geq 0$ ，所以 $\alpha y^\top K_1 y \geq 0$ 成立，所以合法。
- 同上 $\alpha y^\top K_1 y \geq 0$ 成立，但是 $-\alpha y^\top K_1 y \leq 0$ 成立，所以不合法。
- 对于 $K = K_1 K_2$ ，这个是逐元素相乘，所以可以得到我们需要证明 $K = K_1 K_2 \rightarrow y^\top K_1 K_2 y \geq 0$ ，所以我们需要证明矩阵 K_1, K_2 是半正定的，我们知道他们分别是半正定，所以我们可以让 $K_1 = U^\top U$ 。所以 $y^\top K_1 K_2 y = \sum_{i=1}^n \sum_{j=1}^n y_i^\top K_1 K_2 y_j = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n (u_{ki} y_i) K_{ij} (u_{kj} y_j) \geq 0$ ，所以合法。
- 因为 ϕ 是一个合法的从 $d \rightarrow d'$ 的合法映射，所以对于 $\{x_1, x_2, \dots, x_n\}$ ，可以有 $\{\phi(x_1), \phi(x_2), \dots, \phi(x_n)\}$ 对于原核函数也是在核函数处理范围内，所以对于原来的 $K = \kappa_3(\phi(x), \phi(y))$ 可以得到 $K_3 = \kappa_3(\phi(x), \phi(y))$ ， $K = K_3$ ，由于 K_3 一定合法，所以 K 合法。

3 问题三

a. 因为要施加一个 k ，根据题目描述，可以得到 k 对于样本 i 有：

$$k_i = \begin{cases} 1, & y_i = 1 \\ k, & y_i = -1 \end{cases}$$

若 $f(x) = \mathbf{w}^\top \mathbf{x}_i + b$ 所以优化问题变为:

$$\begin{aligned} \min_{w,b} & \frac{1}{2} \mathbf{w}^\top \mathbf{w} + C \sum_{i=1}^n k_i \xi_i \\ s.t. & y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1 - \xi_i, \quad 1 \leq i \leq n \\ & \xi_i \geq 0, \quad 1 \leq i \leq n \end{aligned}$$

b. 把这个问题写成 lagrange 形式:

$$L = \frac{1}{2} \mathbf{w}^\top \mathbf{w} + C \sum_{i=1}^n k_i \xi_i + \sum_{i=1}^n \alpha_i (1 - \xi_i - y_i(\mathbf{w}^\top \mathbf{x}_i + b)) - \sum_{i=1}^n \mu_i \xi_i$$

求 \mathbf{w}, b, ξ_i 的导数:

$$\begin{aligned} \mathbf{w} &= \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \\ 0 &= \sum_{i=1}^n \alpha_i y_i \\ k_i C &= \alpha_i + \mu_i \end{aligned}$$

所以带入可以得到:

$$L = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^\top \mathbf{x}_j$$

所以对偶问题为:

$$\begin{aligned} \max_{\alpha} & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^\top \mathbf{x}_j \\ s.t. & \sum_{i=1}^n \alpha_i y_i = 0, \quad 1 \leq i \leq n, \\ & 0 \leq \alpha_i \leq k_i C, \quad 1 \leq i \leq n, \end{aligned}$$

KKT 条件为:

$$\begin{cases} \mu_i \geq 0, \alpha_i \geq 0 \\ (1 - \xi_i - y_i(\mathbf{w}^\top \mathbf{x}_i + b)) \geq 0 \\ \alpha_i (1 - \xi_i - y_i(\mathbf{w}^\top \mathbf{x}_i + b)) = 0 \\ \xi_i \geq 0 \\ \mu_i \xi_i = 0 \end{cases}$$

4 问题四

a. 1

5 问题五

a. 1