Жеглов Дмитрий 431 группа

Задание 1.

Жеглов Дмитрий А 431

3 1 $\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{\ell}\left(1-M_i(w,b)\right)_+ \to \min_{w,b}$

$M_i(w,b) = y_i(w^T x_i + b)$

Из формулы видно, что чем меньше C тем больше мы штрафуем за большее $\|w\|^2$

(1) $C=10 \Longrightarrow f$

(2) $C=1 \Longrightarrow b$

(3) $C=0.1 \Longrightarrow c$

Т.к. ширина полосы $\frac{2}{\|w\|}$, значит при маленьких C ширина полосы увеличивается а при больших C ширина полосы уменьшается!

$K(x,x') = \exp(-\gamma\|x-x'\|^2)$ Ядро

$sgn(\sum \lambda_i y_i K(x_i,x) - w_0)$ Классификатор

Из вида ф-ции ядра видно, что при больших γ уменьшается влияние дальних точек т.к. $\|x-x'\|$ растет и ядро принимает их к нулю $\Rightarrow$ классификатор подстраивается под ближайшие точки $\Rightarrow$ более сложное разбиение пространства

(4) $\gamma=1$, $C=3 \Rightarrow e$

(5) $\gamma=10$, $C=1 \Rightarrow a$

(6) $\gamma=0.1$, $C=15 \Rightarrow d$

Параметр C так же как и описывалось выше влияет на ширину полос, что подтверждает наш выбор.

Задание 2.

$\boxed{3\ 2.}$ Наша оптимизационная задача имеет вид

1) $$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum_{n=1}^{N}\xi_n \to \min\limits_{w,\xi} \\ y_n w^T \cancel{X}x_n \geq 1-\xi_n \\ \xi_n \geq 0 \qquad n=1\ldots N \end{cases}$$

~~[illegible]~~ $\Rightarrow \begin{cases} M_n \geq 1-\xi_n \\ \xi_n \geq 0 \end{cases}$

где $M_n = y_n w^T x_n$ отступ от разделяющей границы

на лекции говорилось о существовании хотя бы одного равенства в двойственной задаче

$\Rightarrow \xi_n = \max(1-M_n, 0)$

$\Rightarrow$ Функция потерь имеет вид $\boxed{Loss(M) = \begin{cases} 1-M, & M<1 \\ 0, & M\geq 1 \end{cases}}$

2) Построим задачу квадратичного программирования для функции потерь вида

$$\mathcal{L}(M) = \begin{cases} (M-1)^2, & M<1 \\ 0, & M\geq 1 \end{cases}$$

Запишем оптимизационную задачу

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\sum_{n=1}^{N}\varphi_n^2 \to \min\limits_{w,\varphi} \\ y_n w x_n \geq 1-\varphi_n \\ \varphi_n \geq 0 \quad, n=1\ldots N \end{cases}$$

Запишем Лагранжиан

$$L = \frac{1}{2}\|w\|^2 + C\sum_{n=1}^{N}\varphi_n^2 + \sum_{n=1}^{N}\alpha_n(1-\varphi_n - M_n) - \sum_{n=1}^{N}\beta_n\varphi_n$$

$\alpha_n, \beta_n \geq 0$ не равны одновременно нулю все множители Лагранжа

$$\begin{cases} \dfrac{\partial L}{\partial w} = w - \sum\limits_{n=1}^{N} \alpha_n y_n x_n = 0 \\ \dfrac{\partial L}{\partial \varphi_n} = 2C\varphi_n - \alpha_n - \beta_n = 0, \; n = 1..N \end{cases}$$

$$\Rightarrow \begin{cases} w = \sum\limits_{n=1}^{N} \alpha_n y_n x_n \\ \varphi_n = \dfrac{\alpha_n + \beta_n}{2C}, \; n = 1..N \end{cases}$$

$$\begin{cases} \beta_n \varphi_n = 0 \\ \alpha_n (1 - \varphi_n - M_n) \end{cases} \quad n = 1..N$$

Условия нежесткости из Th Куна-Таккера

Сведем двойственну задачу

$$\Rightarrow \begin{cases} L(\varphi, w, \alpha, \beta) \to \min\limits_{\alpha,\beta} \\ \beta_n \varphi_n = 0 \\ \alpha_n (1 - \varphi_n - M_n) = 0 \quad n = 1..N \\ \alpha_n \geq 0, \beta_n \geq 0 \quad \text{Неравные одновременно нулю} \end{cases}$$

$$\Rightarrow \begin{cases} \frac{1}{2} \left\| \sum\limits_{n=1}^{N} \alpha_n y_n x_n \right\|^2 + \sum\limits_{n=1}^{N} \dfrac{(\alpha_n + \beta_n)^2}{4C} \to \min\limits_{\alpha,\beta} \\ \alpha_n, \beta_n \geq 0, \; n = 1..N \end{cases}$$

Получим такой вид двойственной задачи

Задание 3.

$\boxed{3.3.}$ $K(x,y) = \exp\left(-\frac{\|x-y\|^2}{2}\right)$

$\varphi : X \to H$ спрямляющее

Возьмем произвольные точки x, y

$x_\varphi = \varphi(x)$

$y_\varphi = \varphi(y)$

метод 1 ближайшего соседа изменит свое качество, если изменится порядок близости между двумя точками

Рассмотрим расстояние

$$\|x_\varphi - y_\varphi\|_H^2 = K(x-y, x-y) =$$

$$= K(x,x) - 2K(x,y) + K(y,y) = 2(1-K(x,y)) =$$

$$= 2\left(1 - \exp\left(-\frac{\|x-y\|^2}{2}\right)\right)$$

Это монотонная функция, а значит порядок близости точек не изменится

А значит качество тоже не изменится

Т.к. метод 1 ближайшего соседа ставит класс точки, которая является ближайшей.

Задание 4.

$\boxed{3\ 4}$ $\forall$ непрерывной $\varphi: X \to H$, $X \subseteq \mathbb{R}^n$

H - гильбертово

Выполнено $\int\limits_X \int\limits_X \langle \varphi(x), \varphi(x') \rangle g(x)\, g(x')\, dx\, dx' \geq 0$

$\forall g: X \to \mathbb{R}$

$\square$

интеграл $\int\limits_X \int\limits_X K(x, x')\, g(x)\, g(x')$ существует

т.к. $K(x, x') = \langle \varphi(x), \varphi(x') \rangle$ непрерывен и все подинтегральные выражение непрерывно как композиция непрерывных функций

т.к наш интеграл является пределом интегральных сумм

$$\sum_i \sum_j K(x_i, x_j)\, g(x_i) g(x_j)\, \mu_i \mu_j \geq 0$$

~~~~ ~~~~ сумма будет неотрицательна по Теореме Мерсера, которая определяет необходимые и достат. условия, которыми должна обладать функция $K(x, x')$ для того, чтобы являться ядром

Если введем обозначения $g(x_i)\mu_i = c_i \in \mathbb{R}$

$g(x_j) \cdot \mu_j = c_j \in \mathbb{R}$

$\Rightarrow$ ~~получим~~ применим теорему Мерсера (о характеризации)

$\blacksquare$
~~~~

Задание 5.

35 Th: если $K_i : X \times X \to \mathbb{R}$, $i \in \mathbb{N}$ удовлетворяет конечной неотрицательной определенности $\Rightarrow$

1) $\alpha K_1 + \beta K_2$, $\alpha, \beta \geq 0$
2) $K_1 \cdot K_2$
3) $\lim\limits_{n \to \infty} K_n$

Тоже удовлетворяет

$\square$ 1) Т.к. K_1 и K_2 неотрицательно определены то умножение на неотрицательные константы и последующие суммирование не изменит свойства неотрицательной определенности

2) $\exists\ A = U \Sigma U^* = K_1$, $B = K_2 = \left(K(x_i, x_j) \right)_{\substack{i = 1..n \\ j = 1..n}}$

U унитарная матрица

$\Sigma = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix}$ $\lambda_i \geq 0$

$U^* = \bar{U}^T$

$\forall c_1 \ldots c_n \in \mathbb{C}$

$\Rightarrow K_1 \cdot K_2 = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} c_i \bar{c}_j A_{ij} B_{ij} =$

$= \sum\limits_{\alpha=1}^{n} \lambda_\alpha c_i \bar{c}_j u_i^\alpha \bar{u}_j^\alpha B_{ij} = \sum\limits_{\alpha=1}^{n} \lambda_\alpha \xi^{\alpha T} B \bar{\xi}^\alpha \geq 0$

где $\xi^\alpha = (c_1 u_1^\alpha, \ldots, c_n u_n^\alpha)^T \in \mathbb{C}^n$

Т.к. $\xi^{\alpha T} B \bar{\xi}^\alpha$ и λ_α неотрицательны $\forall \alpha \Rightarrow$

и вся сумма неотрицательна

3) $0 \leq K_n \Rightarrow \lim\limits_{n \to \infty} K_n \geq 0$
$\forall n \in \mathbb{N}$

$\blacksquare$

Т.к. $f(K_0(x, x'))$ представляется в виде сходящегося степенного ряда с неотрицательными коэффициентами и K_0 это ядро $\Rightarrow$ по утверждению теоремы выходит что и $f(K_0)$ ядро

Задание 6.

З 6 $H = \{ \sum_{i=1}^{\ell} \alpha_i K(x_i, \cdot) \mid \ell \in \mathbb{N}, x_i \in X, \alpha_i \in \mathbb{R} \}$

$$\langle f, g \rangle = \sum_{i=1}^{\ell} \sum_{j=1}^{n} \alpha_i \beta_j K(x_i, z_j) = \sum_{i=1}^{\ell} \alpha_i g(x_i) = \sum_{j=1}^{n} \beta_j f(z_j)$$

$$f(x) = \sum_{i=1}^{\ell} \alpha_i K(x_i, x)$$

$$g(x) = \sum_{j=1}^{n} \beta_j K(z_j, x)$$

при непрерывном ядре $K(x, z)$ сепарабельно

$\square$ покажем что внутри H есть счетное всюду плотное множество. Построим по индукции

База: $f(x) = \alpha_1 K(x_1, x)$, $g(x) = \beta_1 K(x_1', x)$

$\alpha_1 \in \mathbb{R}$ приблизим x_1 рациональным x_1' $\Rightarrow \exists \beta_1 \in \mathbb{Q}$, $\exists x_1' \in \mathbb{Q}^n \cap X$

$x_1 \in X \subseteq \mathbb{R}^n$ в силу сепарабельности $\mathbb{R}^n$

т.е. для некоторого $\varepsilon > 0$ в силу непрерывности функции K найти $g(x)$ такую что

$$|K(x_1, x_1) - K(x_1', x_1)| < \frac{\varepsilon}{2}$$

$$|K(x_1', x_1') - K(x_1, x_1')| < \frac{\varepsilon}{2}$$

$$|\alpha_1 - \beta_1| < \varepsilon$$

Оценим норму

$$\| \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \|_H^2 =$$

$$= \alpha_1^2 K(x_1, x_1) - \alpha_1 \beta_1 K(x_1', x_1) - \alpha_1 \beta_1 K(x_1, x_1') + \beta_1^2 K(x_1', x_1') \le$$

$$\le \alpha_1^2 K(x_1, x_1) + \alpha_1 \beta_1 \left(\frac{\varepsilon}{2} - K(x_1, x_1)\right) - \alpha_1 \beta_1 \left(K(x_1, x_1) - \frac{\varepsilon}{2}\right) +$$

$$+ \beta_1^2 \left(\frac{\varepsilon}{2} + K(x_1, x_1)\right) = K(x_1, x_1)(\alpha_1 - \beta_1)^2 + \beta_1^2 \frac{\varepsilon}{2} + \alpha_1 \beta_1 \left(\frac{\varepsilon}{2} + \frac{\varepsilon}{2}\right) \le$$

$$\le \varepsilon \left(\alpha_1 \beta_1 + \frac{1}{2} \beta_1^2 + K(x_1, x_1)\right)$$, при $\varepsilon \to 0$ имеем сходимость

Шаг: $\exists\ f(x) = f'(x) + \alpha_1 K(x_1, x)$

$$g(x) = g'(x) + \beta_1 K(x_1', x)$$

по индукции $\exists\ g'(x), \beta_1, x_1'$:

$$\| f'(x) - g'(x) \|^2 < \varepsilon$$

$$\| \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \|^2 < \varepsilon$$

$$\langle f'(x) - g'(x), \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \rangle \leq$$

$$\leq \| f'(x) - g'(x) \| \cdot \| \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \| \leq \sqrt{\varepsilon} \cdot \sqrt{\varepsilon} = \varepsilon$$

по Коши-Буняковскому-Шварцу

$$\Rightarrow \| f(x) - g(x) \|^2 = \| f'(x) - g'(x) \|^2 +$$

$$+ 2 \langle f'(x) - g'(x), \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \rangle +$$

$$+ \| \alpha_1 K(x_1, x) - \beta_1 K(x_1', x) \|^2 < 3\varepsilon$$

$\Rightarrow$ доказали

$\Rightarrow$ построили

$$H' = \left\{ \sum_{i=1}^{l} \alpha_i K(x_i', \cdot) \;\middle|\; l \in \mathbb{N},\ x_i' \in \mathbb{Q}^n \cap X,\ \alpha_i \in \mathbb{Q} \right\}$$

всюду плотное в H

счетно т.к. $l \in \mathbb{N}$, $\alpha_i \in \mathbb{Q}$, $x_i' \in \mathbb{Q}^n \cap X$

а они все счетные