Сеинов Н.

3.1

① $a(x) = \text{sign}(\langle w, x\rangle) = \text{sign}(\sum_{i=0}^{\ell} w_i x_i)$

$x_0 = 1$

② $M_i = y_i \langle w, x^i \rangle$, где $y_i$ – ответ

$y_i = \begin{cases} 1, & \langle w, x^i\rangle > 0 \\ -1, & \langle w, x^i\rangle < 0 \end{cases}$

Если $M_i > 0 \Rightarrow$
объект классифицирован правильно,
т.к. $y_i$ и $\langle w, x^i\rangle$ одного знака

$M_i < 0 \Rightarrow$ $y_i$ и $\langle w, x^i\rangle$ разных знаков $\Rightarrow$ объект класс-н неверно.

③ $a(x) = \text{sign}(\langle w, x\rangle - w_0)$

добавляем $x_0 = 1 \Rightarrow a(x) = \text{sign}(\langle w, x\rangle) = \text{sign}(\sum_{i=0}^{\ell} w_i x_i)$

④ $Q(x, M) = \frac{1}{\ell} \sum_{i=1}^{\ell} I[M_i \le 0]$

для наилучшего алгоритма $Q(M) = 0$

⑤ возьмем $w = \overline{0} \Rightarrow \forall\, M = 0 \Rightarrow Q(M) = 0$

⑥ $Q(M) = \frac{1}{\ell} \sum_{i=1}^{\ell} I[M_i \le 0] \le \frac{1}{\ell} \sum_{i=1}^{\ell} L(M) = \tilde{Q}(M)$

⑦ Функция потерь показывает ошибку алгоритма на объекте $x$. Минимизируя её мы улучшаем качество алгоритма

⑧ $Q(x) = I[M_i \le 0]$

(9) Регуляризация – наложение штрафа на искомые параметры, что позволяет избавиться от переобучения

$l_1 = \lambda \sum_{i=0}^{k} |w_i|$, $l_2 = \lambda \sum_{i=0}^{l} w_i^2$ регуляризаторы

$\lambda$ – гиперпараметр

(10) Алгоритм способны к обобщению, если качество на тестовой выборке существенно не отличается от обучающей. Для переобучения характерно очень хорошее качество второго и плохие результаты тестов

Регуляризация штрафует за большие веса => препятствует усложнению модели (переобучению) и улучшает обобщ. сп-ть

(11) В таких случаях при попадании в окрестность точки будем получать большие значения, из-за чего возникнет переобучение.

(12) Увеличивается с приближением к границам регуляризатора

(13) С регуляризатором не меньше, т.к. он дает неотрицательную добавку в сумму функционала.

(14) Если тестовая выборка несильно отличалась от обучающей, то результат без регуляризации окажется лучше.

Если же данные различаются, то функционал с регур-м даст лучший результат.

(15) Accuracy = $\frac{\#\text{верных ответов}}{\#\text{ответов}}$

Precision = $\frac{tp}{tp+fp}$ = $\frac{\#\text{попаданий}}{\#\text{выстрелов}}$

Recall = $\frac{tp}{tp+fn}$ = $\frac{\#\text{попаданий}}{\#\text{мишеней}}$

(16) ROC-AUC – вероятность встретить объект целевого класса раньше, чем нецелевого.

$TPR = \frac{TP}{tp + fn}$

$FPR = \frac{fp}{fp + tn}$ $ROC = TPR(FPR)$.


AUC – площадь под ROC кривой

(17) $\{(FPR_i, TPR_i)\}_{i=0}^{\ell}$

- сортируем выборку $X^\ell$ по значениям $f(x_i, w)$
- $(FPR_0, TPR_0) = (0, 0)$

```
for i in range(1, ℓ):
    if (y_i == -1):
        (FPR_i, TPR_i) = (FPR_{i-1} + 1/ℓ_-, TPR_{i-1})
    else:
        (FPR_i, TPR_i) = (FPR_{i-1}, TPR_{i-1} + 1/ℓ_+),
```

где $\ell_+$ – # [illegible] в выборке
$\ell_-$ – # остальных слов

$\boxed{3.4}$. Квадратное ядро:

$$K(x, y) = \langle x, y\rangle^2 = (x_1 y_1 + x_2 y_2)^2 = \langle (x_1^2, x_2^2, \sqrt{2}x_1x_2), (y_1^2, y_2^2, \sqrt{2}y_1y_2)\rangle.$$

получаем отображение в спрямляющее пр-во:

$$\varphi(x_1, x_2) \to (x_1^2, x_2^2, \sqrt{2}x_1x_2)$$

линейная пов-ть будет иметь вид

$$\langle (x_1^2, x_2^2, \sqrt{2}x_1x_2), (w_1, w_2, w_3)\rangle + w_0 = w_1x_1^2 + w_2x_2^2 + w_3x_1x_2\sqrt{2}$$

$\Rightarrow w = (1, 2, 0),\ w_0 = -3 \Rightarrow x_1^2 + 2x_2^2 - 3 = 0$

$\boxed{3.5}$ Теорема Куна-Таккера

$$\begin{cases} f(x) \to \min \\ g(x) \le 0 \end{cases} \Leftrightarrow \quad \begin{array}{l} \min L(x) = f(x) + \lambda g(x) \to \min \\ \begin{cases} \lambda_i g_i(x) = 0 \\ \lambda_i \geq 0 \end{cases} \end{array}$$

Ограничение $l_1$-нормы вектора весов чисел

$$\begin{cases} Q(w) = \|Fw - y\|^2 \to \min\limits_{w} \\ \sum\limits_{j=1}^{n} |w_j| \le \tau \end{cases} \Rightarrow \begin{cases} Q(w) = \|Fw - y\|^2 \to \min\limits_{w} \\ \sum\limits_{j=1}^{n} |w_j| - \tau \le 0 \end{cases}$$

из двух условий следует

$$\to \begin{cases} Q(w) = \|Fw - y\|^2 + \lambda \left(\sum\limits_{j=1}^{n} |w_j| - \tau\right) \\ \lambda \geq 0 \\ \lambda \left(\sum\limits_{j=1}^{n} |w_j| - \tau\right) = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} Q(w) = \|Fw - y\|^2 + \lambda \sum\limits_{j=1}^{n} |w_j| + const \to \min\limits_{w} \\ \lambda \geq 0 \\ \left[\begin{array}{l} \lambda = 0 \ \cup \\ \sum\limits_{j=1}^{n} |w_j| = \tau \end{array}\right. \end{cases} \Rightarrow$$

пришли к задаче с добавлением штрафа с $l_1$-нормой