Теоретические задачи

3.2 Вероятностный смысл регуляризаторов Рассмотрим вектор весов

$$w \in \mathbb{R}^n$$

Положим, чо веса распределены по нормальному закону со средним 0 и дисперсией σ^2

$$f_w = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{||w||^2}{2\sigma}) \quad (*)$$

Тогда максимизация правдоподобия $(*)$ будет эквивалентна минимизации

$$-ln(f_w) = \frac{1}{2\sigma}||w||_2 + ln(...)$$

Что в свою очередь эквивалентно минимищации

$$||w||_2 \to min$$

это и есть L_2 регуляризация по определению

Заметим, что распределение лапласа имеет такой же вид с точностью до констант, а единственным отличием является норма => проделав то же самое получаем

$$||w||_1 \to min$$

а это L_1 регуляризация по определению

3.3 SVM и максимизация разделяющей полосы Классификатор

$$a(x) = sign(< w, x > -w_0)$$

Подгоним веса так чтобы:

$$min_{i=1..l} y_i(< w, x_i > -w_0) = 1$$

В случае линейно разделимой выборки мы имеем некотрые точки x_+, x_- которые лежат на границе полосы, тогда ширина полосы будет равна

$$< x_+ - x_-, \frac{w}{||w||} >=$$

$$= \frac{< x_+, w > - < x_-, w >}{||w||} = \frac{2}{||w||}$$

В случае линейно неразделимой выборки получаем

$$\begin{cases} ||w||_2 \to min \\ y_i(< w, x_i > -w_0) \geq 1, i = 1..l \end{cases}$$

$$\begin{cases} \frac{1}{2}||w||_2 + C\sum_{i=1}^{l} \xi_i \to min_{w,w_0,\xi} \\ y_i(< w, x_i > -w_0) \geq 1 - \xi_i, i = 1..l \\ \xi_i \geq 0, i = 1..l \end{cases}$$

Заметим, что

$$y_i(< w, x_i > -w_0) = M_i$$

$$\xi_i \geq 0$$

$$\xi \geq 1 - M_i$$

Тогда получаем, что $\xi_i = (1 - M_i)_+$

Следовательно безусловная задача оптимизации:

$$Q(w, w_0) = \sum_{i=1}^{l}(1 - M_i(w, w_0))_+ + \frac{1}{2C}||w||^2 \rightarrow min_{w,w_0}$$

3.4 Kenel trick

$$K(x, y) = x_1^2 y_1^2 + x_2^2 y_2^2 + 1$$

$$\varphi(x) \rightarrow (x_1^2, x_2^2, 1)$$

$$dimV = 3$$

$$(w_1, w_2, w_3)\varphi(x_1, x_2, x_3) + w_0 = |_{(w_1,w_2,w_3)=(1,2,0),w_0=-3} = x_1^2 + 2x_2^2 - 3 = 0$$

3.6 Повторение метрики качества

accacy - доля правильных ответов при классификации

precision - $\frac{tp}{tp+fp}$

recall - $\frac{tp}{tp+fn}$