

Logistická regrese

To samé, ale zvolíme si nějakou hranici, která určí, jakou třídu máme přiřadit.

MLE

Maximální věrohodnost. Náš model predikuje jaká je distribuce tříd pro každé dané $x \rightarrow p_{model}(x, w)$. Model má modelovat distribuci daných dat. Věrohodnost je pak pravděpodobnost tohom že daný model vygeneruje naše data. Tj

$$L(w) = p_{model}(X, w) = \prod_{i=1}^N p_{model}(x_i, w)$$

Maximal likelihood estimation MLE je pak:

$$\begin{aligned} w_{MLE} &= \operatorname{argmax}_w p_{model}(X, w) = \operatorname{argmin}_w \sum_i -\log(p_{model}(x_i, w)) \\ &= \operatorname{argmin}_w E_{x \in p_{data}} -\log(p_{model}(x, w)) \\ &= \operatorname{argmin}_w H(p_{data}(x), p_{model}(x, w)) \end{aligned}$$

Takže hledám cross entropii mezi datama a modelem.

Sigmoid

Z predikce chceme vyrobiť pravděpodobnost:

$$\begin{aligned} \sigma(x) &= \frac{1}{1 + e^{-x}} \\ \sigma(x)' &= \sigma(x)(1 - \sigma(x)) \end{aligned}$$

Trénování

Naše chybová funkce je

$$\begin{aligned} E(w) &= \frac{1}{N} \sum_i -\log(p(C_{t_i}|x_i, w)) \\ g &= \frac{1}{|b|} \sum_{i \in b} \nabla w - \log(p(C_{t_i}|x_i, w)) \\ w &= w - \alpha g \end{aligned}$$