

Jádrová funkce

Z jednoho data vyrobíme feature space:

$$\varphi(x) = (1, x_1, x_2, \dots, x_1^2 x_2 x_1, \dots)^T$$

Predikce modelu pak vypadají $y = \varphi(x)^T w$

Update provedeme následovně

$$w = w - \frac{\alpha}{|B|} \sum_{i \in B} (\varphi(x)^T w - t_i) \varphi(x)$$

To nás ale může stát hodně času. Každý $\varphi(x)w$ je jenom lineární kombinace vah. Mýsto toho abychom si pamatovali všechny váhy si můžeme pamatovat jenom jejich kombinace β_i

$$w = \sum_{i=1}^N ((\beta_i - (i \in B)) \cdot \frac{\alpha}{|B|} (\varphi(x)^T w - t_i)) \varphi(x)$$
$$\beta_i = \beta_i - \frac{\alpha}{|B|} (\sum_j \beta_j \varphi(x_j)^T \varphi(x_i) - t_i) \varphi(x_i)$$

Čísla $\varphi(x_i)\varphi(x_j)$ budeme potřebovat poměrně často, bude lepší si je předpočítat $K_{i,j} = \varphi(x_i)\varphi(x_j)$