

Feladat: Adott egy $D \subset \mathbb{R} \times \mathbb{R}$ tanító adatbázis amely n db. példát tartalmaz. A következőképpen szeretnénk erre modellt illeszteni optimalizálás segítségével:

1. $H = \{h_{a,b,c}(x) = ax^2 + bx + c : a, b, c \in \mathbb{R}\}$
2. $\ell(x, y, h_{a,b,c}) = (h_{a,b,c}(x) - y)^2$
3. Az optimalizáló algoritmus a gradiens módszer

Add meg a frissítési szabályt az a paraméter tanulásához.

Megoldás: A veszteségfüggvénybe behelyettesítve a hipotézist, egy (x, y) példa vesztesége

$$\ell(x, y, h_{a,b,c}) = (h_{a,b,c}(x) - y)^2 = (ax^2 + bx + c - y)^2.$$

A teljes n elemű adatbázis vesztesége lesz a minimalizálandó függvény:

$$J(a, b, c) = \sum_{i=1}^n (ax_i^2 + bx_i + c - y_i)^2.$$

A frissítési szabály ennek a szerinti parciális deriváltjából adódik, azaz

$$a_{t+1} = a_t - \gamma \left. \frac{\partial J(a, b, c)}{\partial a} \right|_{a_t, b_t, c_t} = a_t - \gamma \sum_{i=1}^n 2(a_t x_i^2 + b_t x_i + c_t - y_i) x_i^2.$$

Megjegyzés: Lehet variálni az adatok típusát, pl. lehet $D \subset \mathbb{R}^2 \times \mathbb{R}$, lehet variálni a h hipotézist, pl. lehet akár trigonometrikus, pl. $h_{a,b}(x) = \sin(ax + b)$, illetve lehet variálni a veszteségfüggvényt is, pl. lehet $\ell(x, y, h_{a,b}) = |h_{a,b}(x) - y|$ (ekkor persze a derivált két alesetre bomlik, ezért kissé nehezebb). Rá lehet kérdezni bármelyik paraméterre, ha több is van.