Least Squares Method

$f(x) = ax+b$

$$\left.\begin{matrix} y_1 = ax_1+b \\ y_2 = ax_2+b \\ \vdots \\ y_n = ax_n+b \end{matrix}\right\} Y = Ax \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad x = \begin{bmatrix} a \\ b \end{bmatrix}$$

$E = Y - Ax$

$e_i = |y_i - f(x_i)| \quad (i=1,2,\cdots,n)$

Squared to make positive vaue

$$E = (y_1 - ax_1 - b)^2 + (y_2 - ax_2 - b)^2 + \cdots + (y_n - ax_n - b)^2 = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

find a, b minimize E

$$\frac{\partial E}{\partial a} = 0, \quad \frac{\partial E}{\partial b} = 0$$

$$\frac{\partial E}{\partial a} = \sum_{i=1}^{n} \{(-x_i)(y_i - ax_i - b) + (y_i - ax_i - b)\cdot(-x_i)\}$$

$$= -2\sum x_i(y_i - ax_i - b) = 0$$

$$\Rightarrow a\sum x_i^2 + b\sum x_i = \sum x_i y_i$$

$$\frac{\partial E}{\partial b} = -2\sum(y_i - ax_i - b) = 0$$

$$\Rightarrow a\sum x_i + nb = \sum y_i$$

Let Σx_i is matrix A

$$\begin{cases} a\sum x_i^2 + b\sum x_i = \sum x_i y_i \\ a\sum x_i + nb = \sum y_i \end{cases}$$

$\Rightarrow A^TA\,x = A^TY$

$x = (A^TA)^{-1}A^TY$

$= A^+Y$

A^+: pseudo inverse matrix

$$\underbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{x} = \begin{bmatrix} x_1^2 + x_2^2 + \cdots + x_n^2 & x_1 + x_2 + \cdots + x_n \\ x_1 + x_2 + \cdots + x_n & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

$$= \begin{bmatrix} \overbrace{a(x_1^2 + x_2^2 + \cdots + x_n^2)}^{a\Sigma x_i^2} + \overbrace{b(x_1 + x_2 + \cdots + x_n)}^{b\Sigma x_i} \\ \underbrace{a(x_1 + x_2 + \cdots + x_n)}_{a\Sigma x_i} + \underbrace{nb}_{nb} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix}}_{A^T} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{Y}$$

$$= \begin{bmatrix} \overbrace{x_1y_1 + x_2y_2 + \cdots + x_ny_n}^{\Sigma x_i y_i} \\ \underbrace{y_1 + y_2 + \cdots + y_n}_{\Sigma y_i} \end{bmatrix}$$

16-2

2017년 5월 9일 화요일 오후 4:39

Optimization

parameter $\xrightarrow{\text{function}}$ error $\qquad x^* = \underset{x}{\arg\min}\, f(x)$

- objective f.
- cost f.
- loss f.

Opt. : find parameter minimize function(error)

- Grid Search : put various values in $f(x)$ then choose minimum value

- Numerical Optimization : through trials and errors, find x^*

has 2-algorithms (conditions)
- try x_k (current step) and find x_{k+1} (next step)
- judge minimum value on x_k

$\nabla f = 0$

for univariate f. $\quad \dfrac{df(x)}{dx} = 0$

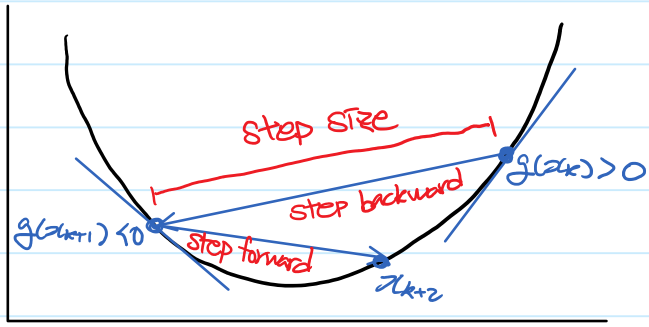
for multivariate f. $\quad \dfrac{\partial f(x_1, x_2, \cdots x_n)}{\partial x_1} = 0, \ \vdots, \ \dfrac{\partial f(x_1, x_2, \cdots x_n)}{\partial x_n} = 0$

SGD (Steepest Gradient Descent)

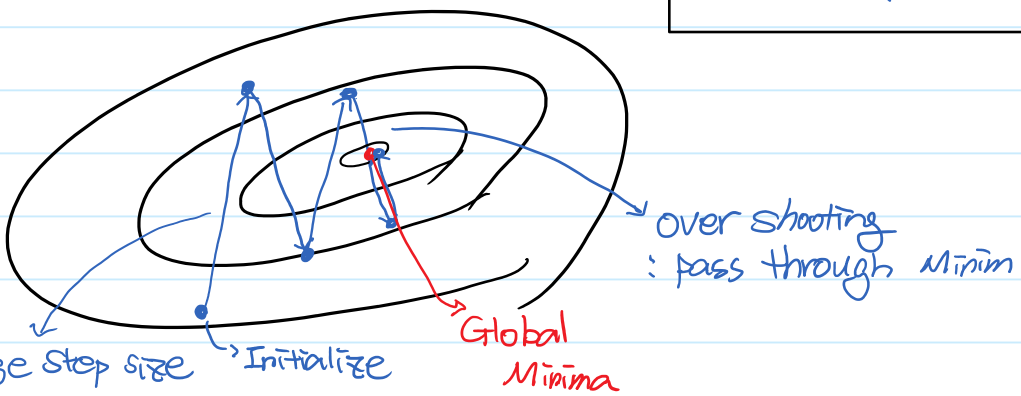
$$\underbrace{x_{k+1}}_{\text{next step}} = \underbrace{x_k}_{\text{current step}} - \underbrace{\mu}_{\text{step size (Learning rate)}} \underbrace{\nabla f(x_k)}_{\text{Slope} = g(x)} = x_k - \mu \cdot g(x_k)$$

if $g(x_k) > 0$, Step backward

$g(x_k) < 0$, Step forward



Large step size (μ) causes over shooting (Slower Convergence)



16-3

2017년 5월 9일 화요일 오후 6:47

Train Classifier

width	length	insect name
3	1	ladybug
1	3	Larva
X	Y	

$Y = AX$

error, prediction: $E = Y - \hat{Y}$ (target: Y)

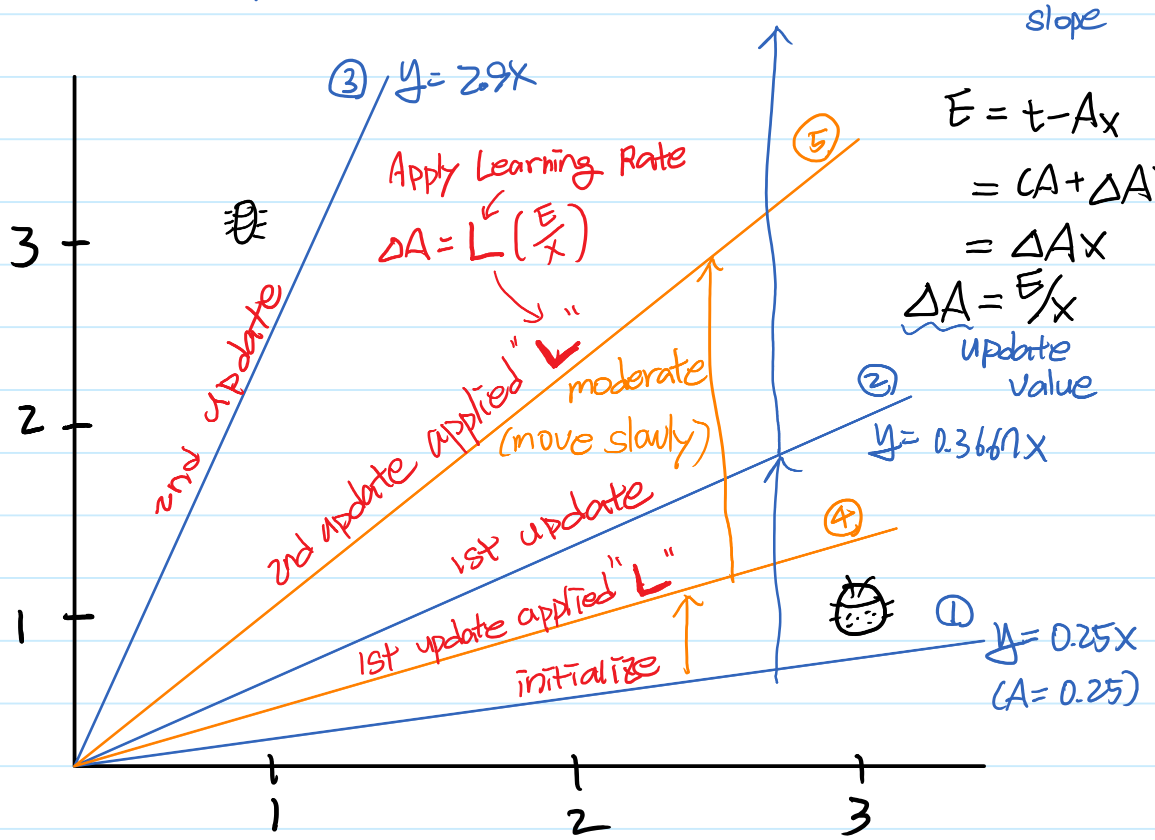
$t = (\underbrace{A}_{\text{current slope}} + \underbrace{\Delta A}_{\text{slope change}})x$

$E = t - Ax$

$= (A + \Delta A)x - Ax$

$= \Delta A x$

$\underbrace{\Delta A}_{\text{update value}} = E/x$



for 1st training data

$(t = 1.1,\ x = 3)$

↳ for proper classification $(1.0 \to 1.1)$

① Initialize: $\hat{y} = 0.25x$

$= 0.25 \cdot 3$

$= 0.75$ (predicted length)

② $E = 1.1 - 0.75 = 0.35$

$\Delta A = E/x = 0.35/3 = 0.1167$ (update)

$A_{new} = A + \Delta A = 0.25 + 0.1167 = 0.3667$

for 2nd training data

$(t = 2.9,\ x = 1)$

③ $\hat{y} = 0.3667 \cdot 1 = 0.3667$

$E = 2.9 - 0.3667 = 2.5333$

$\Delta A = E/x = 2.5333/1 = 2.5333$ (update)

$A_{new} = 0.3667 + 2.5333$

$= 2.9$

16-4

2017년 5월 10일 수요일 오전 12:28