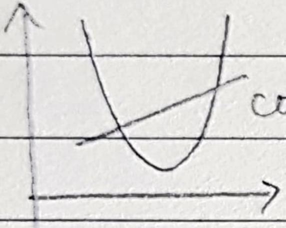
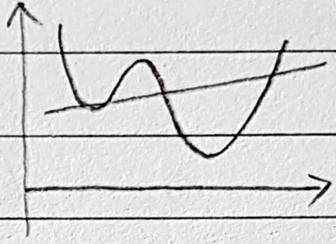
Convex Function

When any line segment drawn between two points on the function lies above or on the function and the min. value occurs at a single point (one global minima)



convex



not convex

e.g $f(x) = 1/x$ $\quad x > 0$ $\qquad y = \theta_0 + \theta_1 x$

$f(x) = -\sqrt{x}$ $\quad x \geq 0$

x	y
0.80581147	2.27515993
0.38348503	-0.15308204
0.6652413	1.43590601
0.64155897	0.79686399
0.24070017	-0.45275524
0.35429554	2.03862963
0.70827991	1.16148089
0.32378987	2.68814558
0.8708779	0.7010376
0.22902348	0.04881045

$$h_\theta(x) = \theta_0 + \theta_1 x$$

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x_{1i}) - y_i \right)^2$$

$m = len(x)$

$m = 10$

$$J(\theta_0, \theta_1) = \frac{1}{2m}\Big[(\theta_0 + 0.805811\theta_1 - 2.27515)^2 + (\theta_0 + 0.383485\theta_1 + 0.15308)^2$$

(convex function)

$$+ (\theta_0 + 0.66524\theta_1 - 1.43590)^2 + (\theta_0 + 0.641559\theta_1 - 0.79686)^2$$

$$+ (\theta_0 + 0.240700\theta_1 + 0.45275)^2 + (\theta_0 + 0.354290\theta_1 - 2.03863)^2$$

$$+ (\theta_0 + 0.708280\theta_1 - 1.16148)^2 + (\theta_0 + 0.323790\theta_1 - 2.68814)^2$$

$$+ (\theta_0 + 0.870870\theta_1 - 0.701103)^2 + (\theta_0 + 0.229020\theta_1 - 0.048810)^2 \Big]$$

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} -\left(h_\theta \left(x_{1i} \right) - y_i \right)$$

$$\theta_0 = \theta_0 - \alpha \, \partial J / \partial \theta_0$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} -x_i \left(h_\theta \left(x_{1i} \right) - y_i \right)$$

$$\theta_1 = \theta_1 - \alpha \, \partial J / \partial \theta_0$$

$$\frac{\partial}{\partial \theta_0} J = \frac{1}{10}\Big[-(\theta_0 + \theta_1(0.805811) - 2.27575) - (\theta_0 + 0.38348\theta_1 + 0.15308)$$
$$- (\theta_0 + 0.665240\theta_1 - 1.43590) - (\theta_0 + 0.641559\theta_1 - 0.79686)$$
$$- (\theta_0 + 0.24070\theta_1 + 0.45275) - (\theta_0 + 0.354290\theta_1 - 2.03863)$$
$$- (\theta_0 + 0.708280\theta_1 - 1.16148) - (\theta_0 + 0.323790\theta_1 - 2.68814)$$
$$- (\theta_0 + 0.870870\theta_1 - 0.701103) - (\theta_0 + 0.229020\theta_1 - 0.048810) \Big]$$

$$\frac{\partial}{\partial \theta_1} J = \frac{1}{10}\Big[-(0.805811(\theta_0 + \theta_1(0.805811) - 2.27575)) - (0.38348(\theta_0 + 0.38348\theta_1 + 0.15308))$$
$$- (0.66524(\theta_0 + 0.66524\theta_1 - 1.43590)) - (0.641559(\theta_0 + 0.641559\theta_1 - 0.79682))$$
$$- (0.24070(\theta_0 + 0.24070\theta_1 + 0.45275)) - (0.35429(\theta_0 + 0.35429\theta - 2.03863))$$
$$- (0.70828(\theta_1 + 0.70828\theta_1 - 1.16148)) - (0.32379(\theta_0 + 0.32379\theta_1 - 2.68814))$$
$$- (0.87087(\theta_0 + 0.87087\theta_1 - 0.701103)) - (0.22902(\theta_0 + 0.22902\theta_1 - 0.048810)) \Big]$$

#	tetha0	tetha1	J
1	0	0	1.066335
2	0.737818	0.431844	0.491165
3	0.801275	0.495645	0.484387
4	0.796986	0.521724	0.483450

First iteration

initially,

$\theta_0 = 0$

$\theta_1 = 0.$

$m = len(x)$

$m = 10.$

$J(0, 0) = 1.066335.$

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} -(h_\theta(x_{1i}) - y_i)$$

at $(0, 0)$,

$$\frac{\partial J}{\partial \theta_0} = -1.054026$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} -x_i(h_\theta(x_{1i}) - y_i)$$

at $(0, 0)$,

$$\frac{\partial J}{\partial \theta_1} = -0.616920.$$

$$\text{new } \theta_0 = \theta_0 - \alpha * \frac{\partial J}{\partial \theta_0}$$

$$\theta_0 = 0 - 0.7 * -1.054026.$$

$$\theta_0 = 0.737818$$

$$\text{new } \theta_1 = \theta_1 - \alpha * \frac{\partial J}{\partial \theta_1}$$

$$\theta_1 = 0 - 0.7 * -0.616920.$$

$$\theta_1 = 0.431844.$$

$$J(0.737818, 0.431844) = 0.491165.$$

Second iteration

initially,

$\theta_0 = 0.737818$.

$\theta_1 = 0.431844$

at which the cost is 0.491165.

at $(0.737818, 0.431844)$,

$$\frac{\partial J}{\partial \theta_0} = -0.090652$$

$$\frac{\partial J}{\partial \theta_1} = -0.091144$$

new $\theta_0 = \theta_0 - \alpha * \frac{\partial J}{\partial \theta_0}$

$\theta_0 = 0.737818 - 0.7 * (-0.090652)$.

$\theta_0 = 0.801275$

new $\theta_1 = \theta_1 - \alpha * \frac{\partial J}{\partial \theta_1}$

$\theta_1 = 0.431844 - 0.7 * (-0.091144)$

$\theta_1 = 0.495645$.

$J(0.801275, 0.495645) = 0.484387$.

Third iteration

initially,

$\theta_0 = 0.801275.$

$\theta_1 = 0.495645.$

at which the cost is 0.484387.

at $(0.801275, 0.495645)$,

$$\frac{\partial J}{\partial \theta_0} = 0.0061279.$$

$$\frac{\partial J}{\partial \theta_1} = -0.037256.$$

new $\theta_0 = \theta_0 - \alpha * \frac{\partial J}{\partial \theta_0}$

$\theta_0 = 0.801275 - 0.7 * (0.0061279)$

$\theta_0 = 0.796985.$

new $\theta_1 = \theta_1 - \alpha \frac{\partial J}{\partial \theta_1}$

$\theta_1 = 0.495645 - 0.7 * (-0.037256)$

$\theta_1 = 0.521724.$

$J(0.796985, 0.521724) = 0.483450.$