M.Atlas Malik   HomeTask#12   24i-8020

Protein Intake	Exercise	Hydration	Muscle mass %
80	5	3	45
100	7	3.5	50
90	6	3.2	48
110	8	4	55
95	6.5	3.8	50

Multivariate linear Regression:

$$\eta = 0.0001$$

Initial weights: $w_1 = w_2 = w_3 = b = 0$

Iteration 1:

$$\hat{y}_i = w_1x_1 + w_2x_2 + w_3x_3 + b$$

$$\hat{y}_1 = \hat{y}_2 = \hat{y}_3 = \hat{y}_4 = \hat{y}_5 = 0$$

Error $= y_i - \hat{y}_i = [45, 50, 48, 55, 50]$.

Update b:

$$\frac{\partial L}{\partial b} = \frac{-2}{5}\sum_{i=1}^{5}(y_i - \hat{y}_i)$$

$$= -\frac{2}{5}[45+50+48+55+50].$$

$$\frac{\partial L}{\partial b} = -99.2$$

$$b_{new} = b_{old} - \eta\frac{\partial L}{\partial b} = 0 - (0.0001)(-99.2)$$

$$\boxed{b_{new} = 0.00992}$$

Update weights: $\frac{\partial L}{\partial w_i} = \frac{-2}{m}\sum x_i(y_i - \hat{y}_i)$.

$\underline{w_1}$: $\frac{\partial L}{\partial w_1} = -\frac{2}{5}[3600 + 5000 + 4320 + 6050 + 4750]$

$$= -\frac{2}{5}(23720) = -9488$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_1}$$

$$= 0 - (0.0001)(\cdot 9488)$$

$$\boxed{w_1 = 0.9488}$$

w_2 :

$$\frac{\partial L}{\partial w_2} = -\frac{2}{5}[1628]$$

$$= -651.2$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_2}$$

$$= 0 - (0.0001)(-651.2)$$

$$\boxed{w_2 = 0.06512}$$

w_3 :

$$\frac{\partial L}{\partial w_3} = -\frac{2}{5}(873.6)$$

$$= -349.44$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_3}$$

$$= 0 - (0.0001)(-349.44)$$

$$\boxed{w_3 = 0.034944}$$

MSE :

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

$$= \frac{1}{5}\left[(50)^2 + (45)^2 + (48)^2 + (55)^2 + (50)^2\right]$$

$$\boxed{MSE = 2470.8}$$

Predicted y:

$$\hat{y}_1 = x_{11}w_1 + x_{12}w_2 + x_{13}w_3 + b$$

$$= 80(0.9488) + 5(0.06512) + 3(0.034944) + 0.00992$$

$$\boxed{\hat{y}_1 = 76.3443}$$

$$\hat{y}_2 = 100(0.9488) + 7(0.06512) + 3.5(0.034944) + 0.00992$$

$$= 95.468$$

$$\hat{y}_3 = 90(0.9488) + 6(0.06512) + 3.2(0.034944) + 0.00992$$

$$= 85.9044$$

$$\hat{y}_4 = 110(0.9488) + 8(0.06512) + 4(0.034944) + 0.0092$$

$$= 105.0386$$

$$\hat{y}_5 = 95(0.9488) + 6.5(0.06512) + 3.8(0.034944) + 0.0092$$

$$= 90.7019$$

Iteration 2:

$$\text{Error} = y_i - \hat{y}_i = [-31.3443, -45.468, -37.9044, -50.0386, -40.7019]$$

Update b:

$$\frac{\partial L}{\partial b} = -\frac{2}{5}[-31.3443, -45.468 -37.9044 -50.0386 -40.7019]$$

$$\frac{\partial L}{\partial b} = 82.18$$

$$b_{new} = b_{old} - \eta \frac{\partial L}{\partial b}$$

$$= 0.0092 - (0.0001)(82.18)$$

$$b_{new} = 0.00172$$

Update weights:

$$\frac{\partial L}{\partial w_i} = -\frac{2}{m}\sum_{i=1}^{m} x_{ij}(y_i - \hat{y}_i)$$

$$w_1 : -\frac{2}{5}[80(-31.3448) + 100(-45.468) + 90(-37.9044)$$

$$+ 110(-50.0386) + 95(-40.7019)]$$

$$\frac{\partial L}{\partial w_1} = 7934.6826$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_1}$$

$$= 0.9488 - (0.0001)(7934.6826)$$

$$\boxed{w_1 = 0.1553}$$

w_2 : $$\frac{\partial L}{\partial w_2} = -\frac{2}{5}\Big[5(-31.3443) + 7(-45.468) + 6(-37.9044) + 8(-50.0386) + 6.5(-40.7019)\Big].$$

$$= 546.91$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_2}$$

$$= 0.06512 - (0.0001)(546.91)$$

$$\boxed{w_2 = 0.01042}$$

w_3 : $$\frac{\partial L}{\partial w_3} = 291.71$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_3}$$

$$= 0.034944 - (0.0001)(291.71)$$

$$\boxed{w_3 = 0.005773}$$

MSE : $$MSE = \frac{1}{m}\sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

$$= \frac{1}{5}\Big[(-31.3443)^2 + (-45.468)^2 + (-37.904)^2 + (-50.0386)^2 + (-40.7019)^2\Big]$$

$$MSE = 1729.41$$

Predicted y:

$\hat{y}_1 = 129.49$

$\hat{y}_2 = 15.62$

$\hat{y}_3 = 14.059$

$\hat{y}_4 = 17.191$

$\hat{y}_5 = 14.844$

Iteration 3:

$\text{Error} = y_i - \hat{y}_i = [32.50, 34.375, 33.940, 37.808, 35.175]$.

Update b:

$$\frac{\partial L}{\partial b} = \frac{-2}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i).$$

$$= -\frac{2}{5}[3.50 + 34.375 + 33.940 + 37.808 + 35.155].$$

$$\frac{\partial L}{\partial b} = -69.51$$

$$b_{new} = b_{old} - \eta\frac{\partial L}{\partial b}$$

$$= 0.001702 - (0.0001)(-69.51)$$

$$\boxed{b_{new} = 0.00865}$$

Update weights:

$$\frac{\partial L}{\partial w_j} = -\frac{2}{m}\sum_{i=1}^{m} x_{ij}(y_i - \hat{y}_i)$$

$\underline{w_1}$:

$$\frac{\partial L}{\partial w_1} = -\frac{2}{5}[80(32.50) + 100(34.375) + 90(33.94) + 110(37.8089) + 95(35.155)]$$

$$= -6636.41.$$

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial w_1}$$

$$= 0.1553 - (0.0001)(-6636.41)$$

$$\boxed{w_1 = 0.8189}$$

w_2 :

$$\frac{\partial L}{\partial w_2} = -455.01$$

$$W_{new} = 0.01042 - (0.0001)(-455.01)$$

$$\boxed{w_2 = 0.0559}$$

w_3 :

$$\frac{\partial L}{\partial w_3} = -244.40$$

$$W_{new} = 0.00575 - (0.0001)(-244.40)$$

$$\boxed{w_3 = 0.0302}$$

MSE :

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2$$

$$= \frac{1}{5}\left[(32.50)^2 + (34.3752)^2 + (33.94)^2 + (37.808)^2 + (35.155)^2\right]$$

$MSE = 1211.09$

predicted y :

$$\hat{y}_1 = 65.8908$$

$$\hat{y}_2 = 82.3957$$

$$\hat{y}_3 = 74.1417$$

$$\hat{y}_4 = 90.6557$$

$$\hat{y}_5 = 78.2823.$$

$\Rightarrow \quad x_1 = 105 \quad ; \quad x_2 = 7.5 \quad ; \quad x_3 = 3.6$

$\hat{y} = x_1 w_1 + x_2 w_2 + x_3 w_3 + b$

$= 105(0.8189) + 7.5(0.0559) + 3.6(0.0302) + 0.00855$

$\hat{y} = 86.5231$

The predicted muscle mass percentage for an individual consuming 105 g of protein, exercising 7.5 hours per week, and hydrating with 3.6L per day is 86.52%.

Pros and cons of using Gradient Descent:

Pros:

* scalability
* Flexibility
* Efficiency for high-dimensional data
* supports regularization.

Cons:

* Slow convergence
* Sensitive to learning rate η
* May get stuck in local minima.
* Requires feature scaling
* Hyperparameter tuning needed.

Impact of choosing a very large or small learning rate :

1. Large learning rate :

* converges if chosen correctly
* The gradient updates are too large, leading to oscillations or divergence.
* May overshoot the optimal solution.

2. Small learning rate :

* Guarantees convergence (if training time is long enough).
* Stable updates, no oscillation.
* Extremely slow convergence.
* Might get stuck in a plateau, requiring any iterations to reach the optimal solution.

Ridge Regression :

$\alpha = 0.1$

$\eta = 0.0001$

Initial weights $= w_1 = w_2 = w_3 = 0$

$b = 0$.

Iteration 1 :

$$\hat{y}_1 = \hat{y}_2 = \hat{y}_3 = \hat{y}_4 = \hat{y}_5 = 0$$

Error $= y_i - \hat{y}_i$

$= [45, 50, 48, 55, 50]$.

Update b: $\frac{\partial L}{\partial b} = \frac{-2}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)$

$$= \frac{-2}{5}[45 + 50 + 48 + 55 + 50]$$

$$\frac{\partial L}{\partial b} = -99.2$$

$$b_{new} = b_{old} - \eta \frac{\partial L}{\partial b}$$

$$= 0 - (0.0001)(-99.2)$$

$$= 0.00992$$

Update weights:

$$L = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 + \alpha \sum_{j=1}^{i} w_j^2$$

$$\frac{\partial L}{\partial w_j} = \frac{-2}{m} \sum_{i=1}^{m} x_j (y_i - \hat{y}_i) + 2\alpha w_j$$

$\underline{\underline{w_1}}$: $\frac{\partial L}{\partial w_1} = \frac{-2}{5}\Big[(80)(45) + (100)(50) + (90)(48) + (110)(55)$

$$+ 2(0.1)(0) + (95)(50)\Big].$$

$$= -9488$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_1}$$

$$= 0 - (0.0001)(-9488)$$

$$\boxed{w_1 = 0.9488}$$

$\underline{\underline{w_2}}$: $\frac{\partial L}{\partial w_2} = \frac{-2}{5}\Big[(5)(45) + (7)(50) + (6)(48) + (8)(55) +$

$$(6.5)(50)\Big] + 2(0.1)(0).$$

$= -651.2$

$w_2 = 0 - (0.0001)(-651.2)$

$\boxed{w_2 = 0.06512}$

W_3 : $\frac{\partial L}{\partial w_3} = -\frac{2}{5}\left[(3)(45) + (3.5)(50) + (3.2)(48) + (4)(55) + (3.8)(50)\right] + 2(0.1)(0)$

$= -349.44$

$w_3 = 0 - (0.0001)(-349.44)$

$\boxed{w_3 = 0.03494}$

MSE : $MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2$

$= \frac{1}{5}(45^2 + 50^2 + 48^2 + 55^2 + 50^2)$

$= 2470.8$

Predicted y :

$\hat{y}_i = w_1 x_{i1} + w_2 x_{i2} + w_3 x_{i3} + b$

$\hat{y}_1 = 76.344$

$\hat{y}_2 = \cancel{85.904}\ 45.468$

$\hat{y}_3 = 85.904$

$\hat{y}_4 = 105.038$

$\hat{y}_5 = 90.702$

Iteration 2 : Error = $y_i - \hat{y}_i$

$= [-31.34432, -45.468, -37.904, -50.038, -40.702]$.

Update b :

$$\frac{\partial L}{\partial b} = \frac{-2}{5}[-31.344, -45.468, -37.904, -50.038, -45.7]$$

$= 82.182$

$$b_{new} = b_{old} - \eta \frac{\partial L}{\partial b}$$

$= 0.00992 - (0.0001)(82.182)$

$\boxed{b = 0.0017}$

Update weights:

w_1 :

$$\frac{\partial L}{\partial w_1} = \frac{-2}{5}\Big[(80)(-31.344) + (100)(-45.468) + (90)(-37.90) + (110)(-50.038) + (95)(-40.702)\Big] + 2(0.1)(0.9488).$$

$= 7934.809$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_1}$$

$= 0.9488 - (0.0001)(7934.809)$

$\boxed{w_1 = 0.1553}$

w_2 :

$$\frac{\partial L}{\partial w_2} = \frac{-2}{5}\Big[(5)(-31.344) + (7)(-45.468) + (6)(-37.844) + (8)(-50.038) + (6.5)(-40.702)\Big] + 2(0.1)(0.0652)$$

$= 546.927$.

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial w_2}$$

$$= 0.0612 - (0.0001)(546.927)$$

$$\boxed{w_2 = 0.0104}$$

$\underline{\underline{w_3}}$: $$\frac{\partial L}{\partial w_3} = \frac{-2}{5}\Big[(3)(-31.344) + (3.5)(-45.468) + (3.2)(-37.904) + (4)(-50.038) + (3.8)(-40.702)\Big] + 2(0.1)(0.03$$

$$= 291.72$$

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial w_3}$$

$$= 0.03494 - (0.0001)(291.72)$$

$$\boxed{w_3 = 0.00576}$$

MSE :

$$MSE = \frac{1}{5}\Big[(-31.344)^2 + (-45.468)^2 + (-37.9044)^2 + (-50.038)^2 + (-40.701)^2\Big].$$

$$= 1729.41$$

Predicted y :

$$\hat{y}_i = w_1 x_{i1} + x_{i2} w_2 + x_{i3} w_3 + b$$

$$\hat{y}_1 = 12.49$$

$$\hat{y}_2 = 15.62$$

$$\hat{y}_3 = 14.06$$

$\hat{y}_4 = 17.19$

$\hat{y}_5 = 14.84$

Iteration 3: Error $= y_i - \hat{y}_i$

Error $= [32.503, 34.373, 33.939, 37.807, 35.153]$.

Update b:

$$\frac{\partial L}{\partial b} = \frac{-2}{5}[32.503 + 34.373 + 33.939 + 37.807 + 35.153].$$

$$\frac{\partial L}{\partial b} = -69.51$$

$$b_{new} = b_{old} - \eta \frac{\partial L}{\partial b}$$

$$= 0.001702 - (0.0001)(-69.51)$$

$$\boxed{b = 0.00865}$$

Update weights:

$$\frac{\partial L}{\partial w_j} = \frac{-2}{m} \sum_{i=1}^{m} x_{ij} (y_i - \hat{y}_i) + 2\alpha w_j$$

$\underline{\underline{w_1}}$:

$$\frac{\partial L}{\partial w_1} = -6636.41$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_1}$$

$$\boxed{w_1 = 0.8189}$$

$\underline{\underline{w_2}}$:

$$\frac{\partial L}{\partial w_2} = -455.089$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_2}$$

$$\boxed{w_2 = 0.0559}$$

$\underline{\underline{w_3}}$:

$$\frac{\partial L}{\partial w_3} = -244.4943$$

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w_3}$$

$$\boxed{w_3 = 0.0302}$$

MSE :

$$MSE = \frac{1}{n} \Sigma (y_i - \hat{y}_i)^2$$

$$MSE = 1211.02$$

$\Rightarrow x_1 = 105 \; ; \; x_2 = 7.5 \; ; \; x_3 = 3.6$

$$\hat{y} = x_1 w_1 + x_2 w_2 + x_3 w_3 + b$$

$$= (105)(0.8189) + (0.0559)(7.5) + (3.6)(0.0302) + 0.00865$$

$$\hat{y} = 86.52$$

$\Rightarrow$ **Difference :**

L1 Regularization	L2 Regularization
* L1 (Lasso) can shrink some weights to exact zero, effictively selecting important features.	* L2 (Ridge) shrinks weights closer to zero but does not eliminate any features.

* L1 tends to select only one of the correlated features while reducing others to zero.

* L2 is better at handling multicollinearity by distributing weights among correlated features.

How regularization helps prevent overfitting:

Overfitting occurs when a model learns patterns from noise rather than general trends.

Regularization helps prevents this by:

* Reducing model complexity.
* Preventing large coefficients.
* Improving generalization.