

Weighted Least Squares Regression

#Math Equation for the gradient of the loss with respect to the weight:

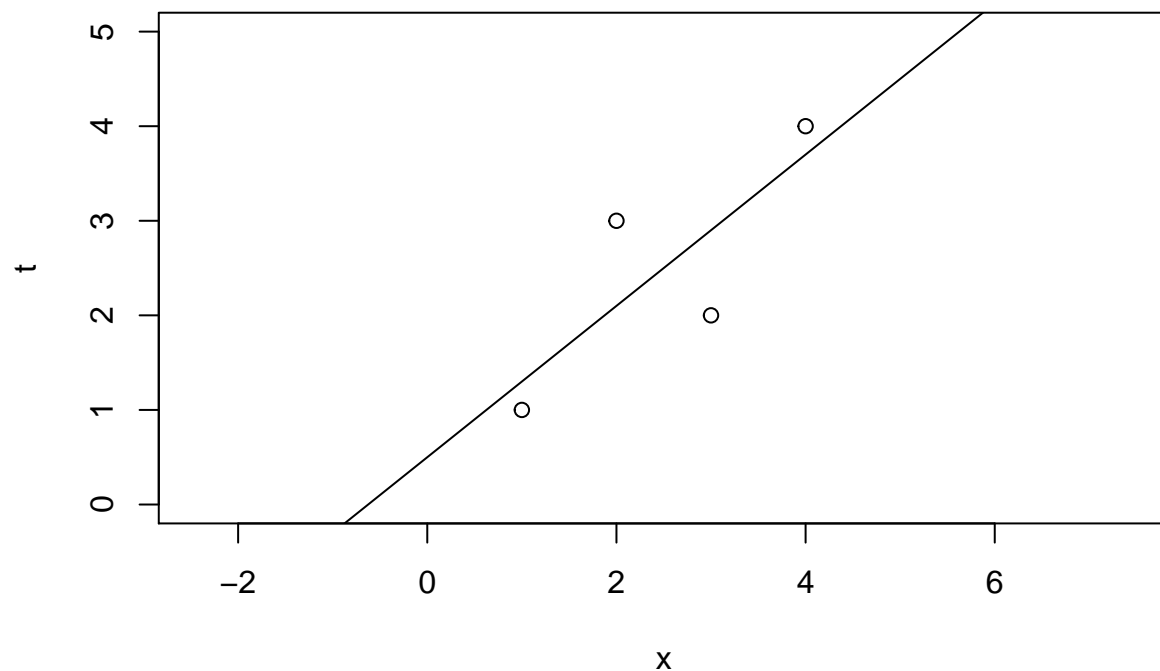
$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} = -\frac{2}{N} \mathbf{X}^T \mathbf{A} \mathbf{t} + \frac{2}{N} \mathbf{X}^T \mathbf{A} \mathbf{X} \mathbf{w}$$

#Solution for the weight:

$$\hat{\mathbf{w}} = (\mathbf{X}^T \mathbf{A} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{A} \mathbf{t}$$

```
#some random data
x <- c(1, 2, 3, 4)
t <- c(1, 3, 2, 4)

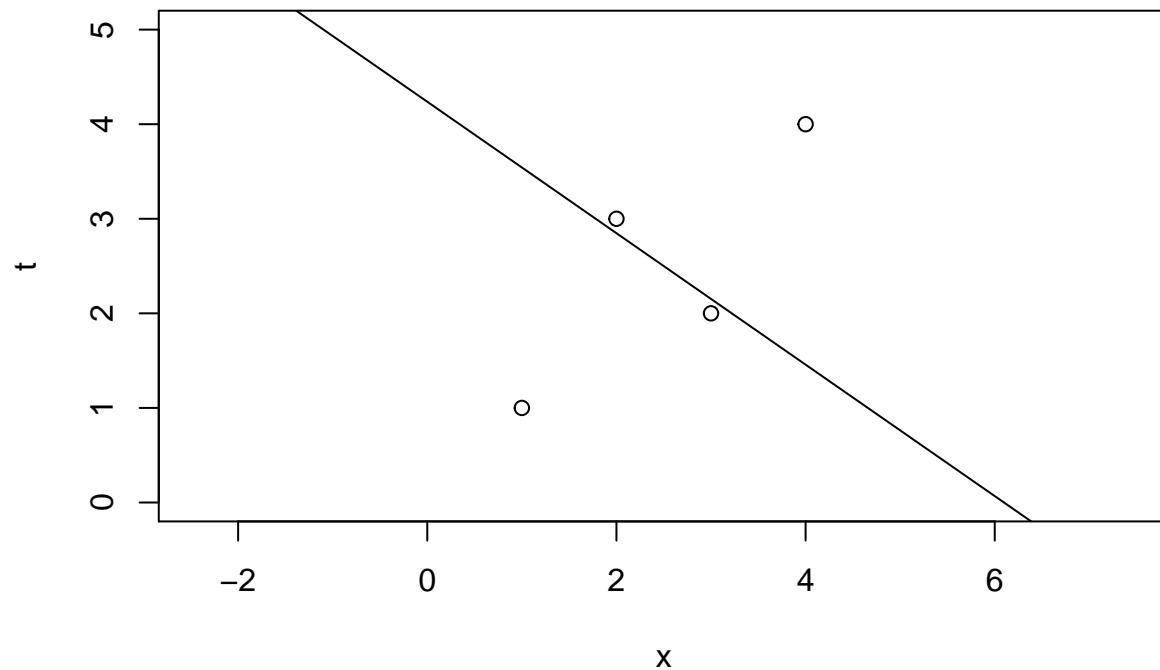
#lm with weights
a1 <- c(1, 1, 1, 1)
model1 <- lm(t ~ x, weights = a1)
plot(x, t, xlim = c(0, 5), ylim = c(0, 5), asp = 1)
abline(model1)
```



```
print(model1$coefficients)
```

```
## (Intercept)      x
##      0.5      0.8
```

```
#Effects of altering the weights on the two inner points (x=2, x= 3) and the outer points (x=1, x=4)
a2 <- c(0.1, 5, 5, 0.1)
model2 <- lm(t ~ x, weights = a2)
plot(x,t,xlim = c(0,5), ylim = c(0,5), asp=1)
abline(model2)
```



```
print(model2$coefficients)
```

```
## (Intercept)      x
##  4.2372881 -0.6949153
```

Results

Weighted Least Squares is used to maximize the efficiency of parameter estimation. It gives each data point its proper amount of influence over the parameter estimates. The size of the weight points out the precision of the information held in the associated observation. In other words putting a large weight on a point would make the point to have large amount of influence over the parameter estimates. On the other hand putting a small weight on a point would make the point to have small amount of influence over the parameter estimates.