Iteratively Re-weighted least squares

Weighted Least Squares: Minimize

$$Q = \sum_{i=1}^{n} w_i \left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)^2$$

Iterative method for fitting generalized linear models

Let $\widehat{\boldsymbol{\beta}}$ be the current estimate of $\boldsymbol{\beta}$. For $i = 1, \ldots, n$,

1. Calculate $\widehat{\eta}_i = \mathbf{x}_i'\widehat{\boldsymbol{\beta}}$.
2. Obtain $\widehat{\mu}_i$ from the link function: $\eta = g(\mu)$ so $\widehat{\mu}_i = g^{-1}(\widehat{\eta}_i)$.
3. Form the "dependent variable" $z_i = \widehat{\eta}_i + (y_i - \widehat{\mu}_i) \left(\frac{\partial \eta}{\partial \mu}\right)\Big|_{\mu=\widehat{\mu}_i}$
4. Form the weight $w_i^{-1} = \left(\frac{\partial \eta}{\partial \mu}\right)^2\Big|_{\mu=\widehat{\mu}_i} V(\widehat{\mu}_i)$
5. Perform a weighted least squares regression minimizing $Q = \sum_{i=1}^{n} w_i \left(z_i - \mathbf{x}_i'\boldsymbol{\beta}\right)^2$, obtaining a new $\widehat{\boldsymbol{\beta}}$.
6. Go to Step 1 and repeat.

For logistic regression, which has a logit link and $V(\mu) = \mu(1-\mu)$,

1. $\widehat{\eta}_i = \mathbf{x}_i'\widehat{\boldsymbol{\beta}}$.
2. $\widehat{\mu}_i = \frac{e^{\widehat{\eta}_i}}{1+e^{\widehat{\eta}_i}}$
3. $z_i = \widehat{\eta}_i + \frac{y_i - \widehat{\mu}_i}{\widehat{\mu}_i(1-\widehat{\mu}_i)}$
4. $w_i = \widehat{\mu}_i(1-\widehat{\mu}_i)$
5. Minimize $Q = \sum_{i=1}^n w_i \left(z_i - \mathbf{x}_i'\boldsymbol{\beta}\right)^2$ to obtain a new $\widehat{\boldsymbol{\beta}}$.
6. Go to Step 1 and repeat.

```
> datta <-
read.table("http://www.utstat.toronto.edu/~brunner/2201s11/data/logreg.data")
> x1 <- datta$x1; x2 <- datta$x2; x3 <- datta$x3; y <- datta$y
> # Get the answer with glm
> goal = glm(y~x1+x2+x3,family=binomial); goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
>
> # Start with mu near y
> epsilon = 0.1
> mu0 = abs(y-epsilon)
> eta0 = log(mu0/(1-mu0))
> z0 = eta0 + (y-mu0)/(mu0*(1-mu0))
> w0 = mu0*(1-mu0)
> mod = lm(z0~x1+x2+x3,weights=w0)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
  0.1230650   0.9133944  -0.2982430   0.6549709
>
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.07956336  0.57567370 -0.18788979  0.41258949
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
```

```
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08513501  0.62673467 -0.20363130  0.44737877
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08527758  0.62806971 -0.20403243  0.44826726
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
```

```
> # Try another way to start
> mod = lm(y~x1+x2+x3)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.51859923  0.13804439 -0.04507448  0.09898797
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08395495  0.58334777 -0.18768575  0.41312850
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
  0.0852311   0.6270387  -0.2036804   0.4475001
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08527765  0.62807005 -0.20403251  0.44826745
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
```

```
> ################ Iterate this ##########################
> eta = b0 + b1*x1 + b2*x2 + b3*x3
> mu = exp(eta)/(1+exp(eta))
> z = eta + (y-mu)/(mu*(1-mu))
> w = mu*(1-mu)
> mod = lm(z~x1+x2+x3,weights=w)
> b0 = mod$coefficients[1]; b1 = mod$coefficients[2]
> b2 = mod$coefficients[3]; b3 = mod$coefficients[4]
> c(b0,b1,b2,b3)
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> goal$coefficients
(Intercept)          x1          x2          x3
 0.08527768  0.62807061 -0.20403269  0.44826785
> ########################################################
```