

regression diagnostics

Chris Parrish

December 13, 2016

regression diagnostics

reference:

- hatvalues, Stack Overflow
- model diagnostics for regression, Lindquist, Columbia UNiversity
- regression diagnostics, Quick-R
- regression diagnostics, Fox Extensive notes and R code for a short course in Brazil.

Example from the first reference:

```
fit <- lm(hp ~ cyl + mpg, data=mtcars) #a fake model
```

```
hatvalues(fit)
```

```
##           Mazda RX4           Mazda RX4 Wag           Datsun 710
##           0.03204996           0.03204996           0.12124549
##           Hornet 4 Drive       Hornet Sportabout           Valiant
##           0.03352930           0.11184515           0.05236248
##           Duster 360           Merc 240D           Merc 230
##           0.06555982           0.09237898           0.12124549
##           Merc 280           Merc 280C           Merc 450SE
##           0.03823613           0.05757728           0.07198573
##           Merc 450SL           Merc 450SLC       Cadillac Fleetwood
##           0.08349642           0.06481109           0.12951766
##           Lincoln Continental   Chrysler Imperial           Fiat 128
##           0.12951766           0.06457840           0.19712682
##           Honda Civic           Toyota Corolla           Toyota Corona
##           0.13202106           0.26298312           0.15692651
##           Dodge Challenger           AMC Javelin           Camaro Z28
##           0.06572907           0.06481109           0.07255389
##           Pontiac Firebird           Fiat X1-9           Porsche 914-2
##           0.12505077           0.08238265           0.08011784
##           Lotus Europa           Ford Pantera L           Ferrari Dino
##           0.13202106           0.06723084           0.03440964
##           Maserati Bora           Volvo 142E
##           0.06452342           0.16012526
```

```
hv <- as.data.frame(hatvalues(fit))
mn <- mean(hatvalues(fit))
hv$warn <- ifelse(hv[, 'hatvalues(fit)']>3*mn, 'x3',
  ifelse(hv[, 'hatvalues(fit)']>2*mn, 'x2', '-' ))
```

```
hv
```

```
##           hatvalues(fit) warn
## Mazda RX4           0.03204996 -
## Mazda RX4 Wag       0.03204996 -
## Datsun 710           0.12124549 -
## Hornet 4 Drive       0.03352930 -
```

```

## Hornet Sportabout      0.11184515  -
## Valiant                0.05236248  -
## Duster 360            0.06555982  -
## Merc 240D             0.09237898  -
## Merc 230               0.12124549  -
## Merc 280              0.03823613  -
## Merc 280C             0.05757728  -
## Merc 450SE            0.07198573  -
## Merc 450SL            0.08349642  -
## Merc 450SLC           0.06481109  -
## Cadillac Fleetwood    0.12951766  -
## Lincoln Continental    0.12951766  -
## Chrysler Imperial     0.06457840  -
## Fiat 128               0.19712682  x2
## Honda Civic            0.13202106  -
## Toyota Corolla         0.26298312  x2
## Toyota Corona          0.15692651  -
## Dodge Challenger       0.06572907  -
## AMC Javelin            0.06481109  -
## Camaro Z28             0.07255389  -
## Pontiac Firebird       0.12505077  -
## Fiat X1-9              0.08238265  -
## Porsche 914-2          0.08011784  -
## Lotus Europa           0.13202106  -
## Ford Pantera L         0.06723084  -
## Ferrari Dino           0.03440964  -
## Maserati Bora          0.06452342  -
## Volvo 142E             0.16012526  -

```

```
subset(hv, warn=="x3")
```

```
## [1] hatvalues(fit) warn
## <0 rows> (or 0-length row.names)
```

```
subset(hv, warn%in%c("x2", "x3"))
```

```
##                hatvalues(fit) warn
## Fiat 128          0.1971268    x2
## Toyota Corolla    0.2629831    x2
```

```
hv[order(hv['hatvalues(fit)']), ]
```

```
##                hatvalues(fit) warn
## Mazda RX4          0.03204996  -
## Mazda RX4 Wag      0.03204996  -
## Hornet 4 Drive     0.03352930  -
## Ferrari Dino       0.03440964  -
## Merc 280           0.03823613  -
## Valiant            0.05236248  -
## Merc 280C         0.05757728  -
## Maserati Bora      0.06452342  -
## Chrysler Imperial  0.06457840  -
## Merc 450SLC       0.06481109  -
## AMC Javelin        0.06481109  -
## Duster 360        0.06555982  -
## Dodge Challenger   0.06572907  -

```

```

## Ford Pantera L          0.06723084 -
## Merc 450SE             0.07198573 -
## Camaro Z28             0.07255389 -
## Porsche 914-2         0.08011784 -
## Fiat X1-9              0.08238265 -
## Merc 450SL             0.08349642 -
## Merc 240D              0.09237898 -
## Hornet Sportabout     0.11184515 -
## Merc 230               0.12124549 -
## Datsun 710            0.12124549 -
## Pontiac Firebird      0.12505077 -
## Cadillac Fleetwood    0.12951766 -
## Lincoln Continental    0.12951766 -
## Honda Civic            0.13202106 -
## Lotus Europa           0.13202106 -
## Toyota Corona         0.15692651 -
## Volvo 142E            0.16012526 -
## Fiat 128               0.19712682 x2
## Toyota Corolla        0.26298312 x2

```

```
plot(hatvalues(fit), type = "h")
```

