Why is the concept of “Fix” not defined in R due to singularities?

Authored by

stats writer

May 6, 2024

The concept of "Fix" refers to the number of fixed points in a function, which is the value at which the function does not change. In R, this concept is not defined due to singularities. Singularities are points in a function where the value is undefined or infinite. In such cases, the concept of "Fix" cannot be applied as there is no single value that remains constant. Therefore, the concept of "Fix" is not defined in R to avoid any confusion or inaccuracies caused by singularities.

Fix in R: not defined because of singularities

One error message you may encounter in R is:

Coefficients: (1 not defined because of singularities)

This error message occurs when you fit some model using the glm() function in R and two or more of your predictor variables have an exact linear relationship between them - known as .

To fix this error, you can use the cor() function to identify which variables in your dataset have a perfect correlation with each other and simply drop one of those variables from the regression model.

This tutorial shares how to address this error message in practice.

How to Reproduce the Error

Suppose we fit a to the following data frame in R:

```
#define data
df <- data.frame(y = c(0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1),
x1 = c(3, 3, 4, 4, 3, 2, 5, 8, 9, 9, 9, 8, 9, 9, 9),
x2 = c(6, 6, 8, 8, 6, 4, 10, 16, 18, 18, 18, 16, 18, 18, 18),
x3 = c(4, 7, 7, 3, 8, 9, 9, 8, 7, 8, 9, 4, 9, 10, 13))

#fit logistic regression model
model <- glm(y~x1+x2+x3, data=df, family=binomial)

#view model summary
summary(model)

Call:
glm(formula = y ~ x1 + x2 + x3, family = binomial, data = df)

Deviance Residuals:
Min 1Q Median 3Q Max
-1.372e-05 -2.110e-08 2.110e-08 2.110e-08 1.575e-05

Coefficients: (1 not defined because of singularities)
Estimate Std. Error z value Pr(>|z|)
(Intercept) -75.496 176487.031 0.000 1
x1 14.546 24314.459 0.001 1
```

x2 NA NA NA NA
x3 -2.258 20119.863 0.000 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2.0728e+01 on 14 degrees of freedom
Residual deviance: 5.1523e-10 on 12 degrees of freedom
AIC: 6

Number of Fisher Scoring iterations: 24

Notice that right before the coefficient output, we receive the message:

Coefficients: (1 not defined because of singularities)

This indicates that two or more predictor variables in the model have a perfect linear relationship and thus not every regression coefficient in the model can be estimated.

For example, notice that no coefficient estimate can be made for the x2 predictor variable.

How to Handle the Error

To identify which predictor variables are causing this error, we can use the cor() function to produce a and examine which variables have a correlation of exactly 1 with each other:

#create correlation matrix
cor(df)

y x1 x2 x3
y 1.0000000 0.9675325 0.9675325 0.3610320
x1 0.9675325 1.0000000 1.0000000 0.3872889
x2 0.9675325 1.0000000 1.0000000 0.3872889
x3 0.3610320 0.3872889 0.3872889 1.0000000

From the correlation matrix we can see that the variables x1 and x2 are perfectly correlated.

To resolve this error, we can simply drop one of those two variables from the model since they don't actually provide unique or independent information in the regression model.

#fit logistic regression model

```
model <- glm(y~x1+x3, data=df, family=binomial)

#view model summary
summary(model)

Call:
glm(formula = y ~ x1 + x3, family = binomial, data = df)

Deviance Residuals:
Min 1Q Median 3Q Max
-1.372e-05 -2.110e-08 2.110e-08 2.110e-08 1.575e-05

Coefficients:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -75.496 176487.031 0.000 1
x1 14.546 24314.459 0.001 1
x3 -2.258 20119.863 0.000 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2.0728e+01 on 14 degrees of freedom
Residual deviance: 5.1523e-10 on 12 degrees of freedom
AIC: 6

Number of Fisher Scoring iterations: 24
```

Notice that we don't receive a "not defined because of singularities" error message this time.

Note: It doesn't matter whether we drop x1 or x2. The final model will contain the same coefficient estimate for whichever variable you decide to keep and the overall goodness of fit of the model will be the same.

The following tutorials explain how to handle other errors in R: