How to Perform Stepwise Regression in R: A Complete Guide

Authored by
stats writer

March 3, 2026

The complete guide to performing stepwise regression in R is a comprehensive set of instructions that outlines the step-by-step process for conducting stepwise regression analysis using the statistical software R. It covers the necessary steps for data preparation, model building, and interpretation of results, as well as providing tips and techniques for optimizing the regression model. This guide is designed to assist researchers and data analysts in understanding and implementing this statistical technique in their data analysis. It includes detailed explanations, code snippets, and examples to facilitate a thorough understanding of the process. By following this guide, users can effectively utilize stepwise regression in R to identify the most significant variables and create a parsimonious model for predicting outcomes.

A Complete Guide to Stepwise Regression in R

Stepwise regression is a procedure we can use to build a regression model from a set of predictor variables by entering and removing predictors in a stepwise manner into the model until there is no statistically valid reason to enter or remove any more.

The goal of stepwise regression is to build a regression model that includes all of the predictor variables that are statistically significantly related to the response variable.

This tutorial explains how to perform the following stepwise regression procedures in R:

Forward Stepwise SelectionBackward Stepwise SelectionBoth-Direction Stepwise Selection

For each example we'll use the built-in mtcars dataset:

#view first six rows of *mtcars*
head(mtcars)

mpg cyl disp hp drat wt qsec vs am gear carb
Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 0 1 4 4
Mazda RX4 Wag 21.0 6 160 110 3.90 2.875 17.02 0 1 4 4
Datsun 710 22.8 4 108 93 3.85 2.320 18.61 1 1 4 1
Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 1 0 3 1
Hornet Sportabout 18.7 8 360 175 3.15 3.440 17.02 0 0 3 2
Valiant 18.1 6 225 105 2.76 3.460 20.22 1 0 3 1

We will fit a multiple linear regression model using *mpg* (miles per gallon) as our response variable and all of the other 10 variables in the dataset as potential predictors variables.

For each example will use the built-in step() function from the stats package to perform stepwise selection, which uses the following syntax:

step(intercept-only model, direction, scope)

where:

intercept-only model: the formula for the intercept-only modeldirection: the mode of stepwise search, can be either "both", "backward", or "forward"scope: a formula that specifies which predictors we'd like to attempt to enter into the model

Example 1: Forward Stepwise Selection

The following code shows how to perform forward stepwise selection:

```
#define intercept-only model
intercept_only <- lm(mpg ~ 1, data=mtcars)

#define model with all predictors
all <- lm(mpg ~ ., data=mtcars)

#perform forward stepwise regression
forward <- step(intercept_only, direction='forward', scope=formula(all), trace=0)

#view results of forward stepwise regression
forward$anova

Step Df Deviance Resid. Df Resid. Dev AIC
```

1 NA NA 31 1126.0472 115.94345

2 + wt -1 847.72525 30 278.3219 73.21736

3 + cyl -1 87.14997 29 191.1720 63.19800

4 + hp -1 14.55145 28 176.6205 62.66456

#view final model

forward$coefficients

(Intercept) wt cyl hp

38.7517874 -3.1669731 -0.9416168 -0.0180381

Note: The argument trace=0 tells R not to display the full results of the stepwise selection. This can take up quite a bit of space if there are a large number of predictor variables.

Here is how to interpret the results:

First, we fit the intercept-only model. This model had an AIC of 115.94345.Next, we fit every possible one-predictor model. The model that produced the lowest AIC and also had a statistically significant reduction in AIC compared to the intercept-only model used the predictor *wt*. This model had an AIC of 73.21736.Next, we fit every possible two-predictor model. The model

that produced the lowest AIC and also had a statistically significant reduction in AIC compared to the single-predictor model added the predictor *cyl*. This model had an AIC of 63.19800.Next, we fit every possible three-predictor model. The model that produced the lowest AIC and also had a statistically significant reduction in AIC compared to the two-predictor model added the predictor *hp*. This model had an AIC of 62.66456.Next, we fit every possible four-predictor model. It turned out that none of these models produced a significant reduction in AIC, thus we stopped the procedure.

mpg ~ 38.75 - 3.17*wt - 0.94*cyl - 0.02*hyp

Example 2: Backward Stepwise Selection

The following code shows how to perform backward stepwise selection:

```
#define intercept-only model
intercept_only <- lm(mpg ~ 1, data=mtcars)

#define model with all predictors
all <- lm(mpg ~ ., data=mtcars)
```

```
#perform backward stepwise regression
backward <- step(all, direction='backward', scope=formula(all), trace=0)

#view results of backward stepwise regression
backward$anova

Step Df Deviance Resid. Df Resid. Dev AIC
1 NA NA 21 147.4944 70.89774
2 - cyl 1 0.07987121 22 147.5743 68.91507
3 - vs 1 0.26852280 23 147.8428 66.97324
4 - carb 1 0.68546077 24 148.5283 65.12126
5 - gear 1 1.56497053 25 150.0933 63.45667
6 - drat 1 3.34455117 26 153.4378 62.16190
7 - disp 1 6.62865369 27 160.0665 61.51530
8 - hp 1 9.21946935 28 169.2859 61.30730

#view final model
backward$coefficients

(Intercept) wt qsec am
9.617781 -3.916504 1.225886 2.935837
```

Here is how to interpret the results:

First, we fit a model using all *p* predictors. Define this as Mp.Next, for k = p, p-1, ... 1, we fit all k models that contain all but one of the predictors in Mk, for a total of k-1 predictor variables. Next, pick the best among these k models and call it Mk-1.Lastly, we pick a single best model from among M0...Mp using AIC.

The final model turns out to be:

mpg ~ 9.62 - 3.92*wt + 1.23*qsec + 2.94*am

Example 3: Both-Direction Stepwise Selection

The following code shows how to perform both-direction stepwise selection:

```
#define intercept-only model
intercept_only <- lm(mpg ~ 1, data=mtcars)

#define model with all predictors
all <- lm(mpg ~ ., data=mtcars)

#perform backward stepwise regression
both <- step(intercept_only, direction='both', scope=formula(all), trace=0)

#view results of backward stepwise regression
```

both$anova

Step Df Deviance Resid. Df Resid. Dev AIC
1 NA NA 31 1126.0472 115.94345
2 + wt -1 847.72525 30 278.3219 73.21736
3 + cyl -1 87.14997 29 191.1720 63.19800
4 + hp -1 14.55145 28 176.6205 62.66456

#view final model
both$coefficients

(Intercept) wt cyl hp
38.7517874 -3.1669731 -0.9416168 -0.0180381

Here is how to interpret the results:

First, we fit the intercept-only model.Next, we added predictors to the model sequentially just like we did in forward-stepwise selection. However, after adding each predictor we also removed any predictors that no longer provided an improvement in model fit.We repeated this process until we reached a final model.

The final model turns out to be:

mpg ~ 9.62 - 3.92*wt + 1.23*qsec + 2.94*am

Note that forward stepwise selection and both-direction stepwise selection produced the same final model while backward stepwise selection produced a different model.

How to Test the Significance of a Regression Slope

How to Read and Interpret a Regression Table

A Guide to Multicollinearity in Regression