Fit the Bass Model using R

Jens Hüsers

1 11 2018

Die Erklärung von Innovationsmechanismen und die Entwicklung innovativer Produkte von großem Interessen. Dies wird zudem verstärkt von immer schneller verlaufenden Innovationszyklen.

Jedoch ist die Problematik, Innovationszyklen zu erklären und Produktentwicklungen vorherzusagen schon 1960 wissenschaftlich angestoßen worden. Insbesondere Rogers und Bass befassten sich der Erklärung und Modellierung von Innovationen. Bass gelang es als erster ein mathematisches Modell zur Erklärung von Innovationen und der Produktentwicklung zu erstellen. Das heute unter dem Namen Bass-Modell bekannte Modell wurde nach der Publikation der Theoretischen Grundlagen durch empirische Daten untermauert.

Bei diesem mathematischen Modell werden im Wesentlichen zwei Parameter geschätzt, um den Innovationsprozess zu erklären. Zum einen der Innovationskoeffizient p und zum anderen der Immitationskoeffizient q. (Ich merke mir diese Zuordnung indem ich mit das in unserer Lesrichtung von links nach rechts schauende p als progressiv und damit innovativ. Das zurückblickende q ist der Immitator.) Diese beiden Koeffizienten erklären, in wie fern Innovatoren und Immitatoren wirken. Dazu kann auf Grundlage des Modells zusätzlich geschätzt werden, wie sich die Produktentwicklung am Marktniederschlägt.

Es ist jedoch zu beachten, dass es sich bei diesen Modellen im Wesentlichen um eine Vereinfachung der komplexen Marktrealität handelt. Zudem werden die Prognosen auf Basis der Marktverhältnisse der Vergangenheit getroffen. Verändern sich diese Verhältnisse ist das Modell nicht mehr statthaft, da es diese Veränderungen nicht modellieren kann. Dies ist jedoch ein Problem vieler Vorhersagemodelle und gilt nicht exklusiv für das Bass-Modell.

Das Bass Modell hat sich in der Vergangenheit als nützlich herausgestellt und wird auch heute noch zur Erforschung der Digitalisierung in verschiedenen wissenschaftlichen Publikationen genutzt. Zudem wird das Modell auch für die Absatzplanung in Unternehmen eingesetzt.

Es gibt verschiedene Softwarelösungen, welche eine Schätzung des Modells vornehmen. R bietet sich als Open Source Lösung an, denn das Programm hält die Funktionen zur Modellschätzung bereit. Zudem exisiteren Programmerweiterungen, die den Umgang mit dem Bass-Modell erleichtern.

In diesem Beispiel möchte ich ein kurzes Beispiel demonstrieren, wie mit den Basisfunktionen in R ein Bass-Modell erstellt werden kann, das auf den beobachteten kumulierten Adoptionsraten eines Produktes basieren.

Für die Modellschätzung werden die Marktdaten benötigt. Dabei ist zu beachten, dass diese Daten in jeweils gleichen Zeitabständen vorliegen. Zum beispiel jährlich, quartalsweise oder monatlich. Fehlende Daten können unter Umständen ersetzt werden.

We try to model the cumulative proportions of adopters. Therfore we model the cumulative proporitions of adopters by using the differential function of the of the Bass Model Curve to estimate the coefficient of innovation p and coefficient of immitation q.

$F(t) = 1 - \frac{e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}}$

```r
# bass function
bass_F <- function(Time, p, q) {
  (1 - exp(-(p + q)*Time))/(1 + (q/p)*exp(-(p+q)*Time))
}

x <- 1:10
y <- c(.1, .2, .3, .35, .5, .6, .66, .78, .8, .84)
Time = x
```

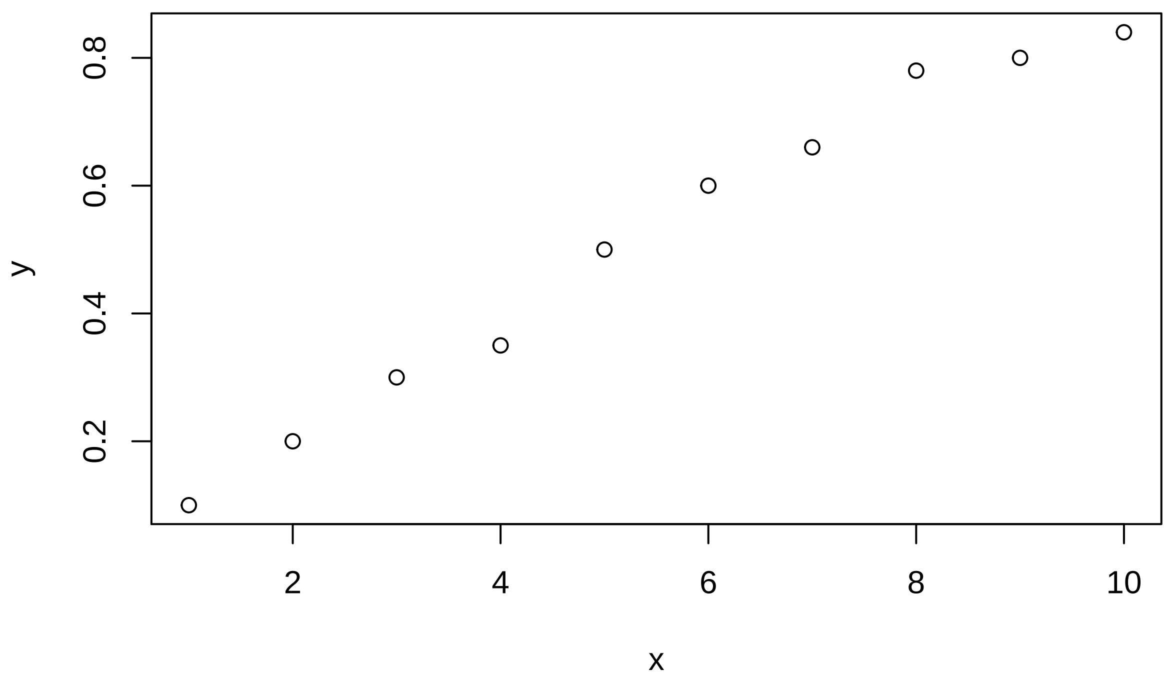
```r
fit <- stats::nls(y ~ bass_F(Time, p, q), start=list(p = .1, q = .5 ))

plot(x, y)
fit <- nls(formula = y ~ bass_F(Time = x, p, q), start=list(p=.1, q=.5))
summary(fit)
```

```
##
## Formula: y ~ bass_F(Time = x, p, q)
##
## Parameters:
##   Estimate Std. Error t value Pr(>|t|)
## p 0.080839   0.007181  11.258 3.48e-06 ***
## q 0.234345   0.028850   8.123 3.91e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02264 on 8 degrees of freedom
##
## Number of iterations to convergence: 6
## Achieved convergence tolerance: 8.376e-07
```

```r
stats::AIC(fit)
```

```
## [1] -43.61239
```

```r
f <- broom::augment(fit)$.fitted
```




```r
mean(sum(f^2)) / mean(x)
```

```
## [1] 0.5923237
```

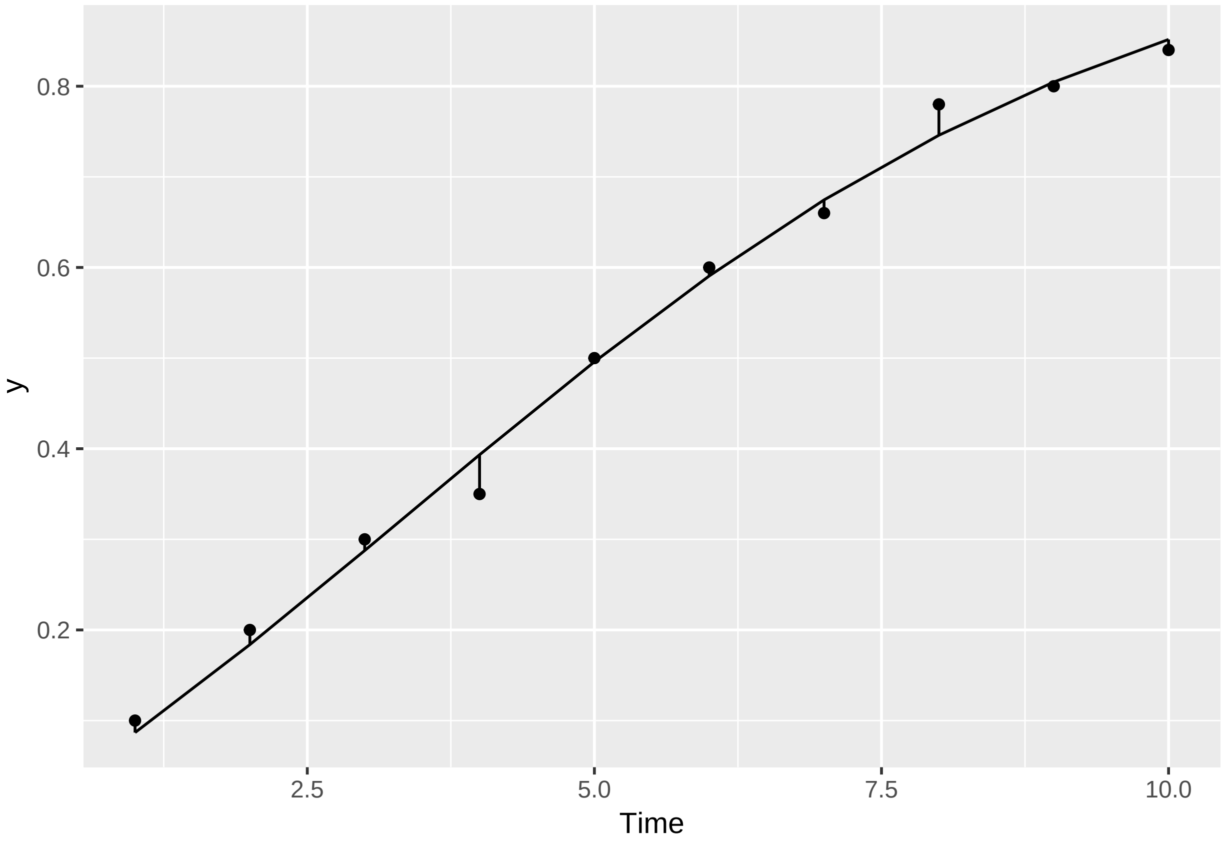
```r
sd(f) / sqrt(length(Time))
```

```
## [1] 0.08442616
```

```r
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------

## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  1.4.2     v dplyr   0.7.6
## v tidyr   0.8.1     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0

## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
broom::augment(fit) %>%
  mutate(y = y) %>%
  ggplot(aes(Time, y)) +
  geom_point() +
  geom_line(aes(x = Time, y = .fitted)) +
  geom_linerange(aes(ymin = .fitted, ymax = y))
```



We used the statistical computing language R to etimate the model's parameters p and q based on the observed cumulated proportions of adopters. In more detail, we used the `nls` Function whicht fits a non linear model using ordinary least squares as the cost function with the Newton-Gauss algorithm. To compare both models we compared the Stanard Errors of the coefficients and calculated the Aikaikes Information Criterion for both models.