

# Nicht-lineare Regression



**Johannes Lüken**, Diplom-Psychologe, ist Leiter des Bereichs Data Sciences bei IfaD.

jlueken@ifad.de



**Prof. Dr. Heiko Schimmelpfennig**, Diplom-Kaufmann, ist Projektleiter für Data Sciences bei IfaD.

hshimmelpfennig@ifad.de

Die lineare Regressionsanalyse geht davon aus, dass der Zusammenhang zwischen den unabhängigen und der abhängigen Variable linear ist. Diese Annahme ist zwar in vielen Fällen angebracht, mitunter ist aber der Einsatz der nicht-linearen Regression erforderlich, mit der auch andere Formen des Zusammenhangs untersucht werden können.

**Beispiel aus der Sensorik.** Es wird die Abhängigkeit der Kaufbereitschaft für eine Limonade von einer Eigenschaft wie der Süße betrachtet. Die Süße wird gemessen auf einer Skala von „nicht wahrnehmbar“ bis „äußerst intensiv“. Es ist zu vermuten, dass für die meisten Befragten eine Ausprägung zwischen den Extremen ideal ist, das heißt zu der höchsten Kaufbereitschaft führt. Geringe Abweichungen von der Idealausprägung beeinflussen die Kaufbereitschaft kaum, große Abweichungen sehr. Bei einem linearen Zusammenhang müsste die gleiche Veränderung der Süße immer zu derselben Veränderung der Kaufbereitschaft führen, unabhängig von der Intensität der Süße. Das ist hier nicht der Fall. Der Zusammenhang

zwischen Süße und Kaufbereitschaft kann vielmehr durch eine quadratische Funktion – eine nach unten geöffnete Parabel – gut beschrieben werden (siehe Abbildung 1).

## Nicht-lineare Regression

Bezeichnen  $x$  die Süße und  $y$  die Kaufbereitschaft, so lautet die Regressionsfunktion  $y = b_0 + b_1 \cdot x + b_2 \cdot x^2$

Diese Regressionsfunktion ist wegen  $x^2$  nicht-linear in den Variablen.

Ebenso wäre zum Beispiel eine Regressionsfunktion mit Regressoren  $x^{-1}$  oder  $e^x$  nicht-linear in den Variablen. Eine Regressionsfunktion ist nicht-linear in den Parametern, wenn sie Koeffizienten wie  $b^{-1}$  oder  $b^2$  enthält.

Die Koeffizienten einer nicht-linearen Regressionsfunktion können mit iterativen Verfahren geschätzt werden. Häufig ist eine Funktion jedoch linearisierbar. Dann lässt sich mit den gängigen Tools zur linearen Regressionsanalyse eine analytische Lösung finden. Die Linearisierung kann mittels Variablensubstitution bzw. -transformation möglich sein.

Durch  $x_1 = x$  und  $x_2 = x^2$  wird in dem Beispiel einer quadratischen Regressionsfunktion  $y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2$  zu einer Funktion, die linear in allen Parametern und Variablen ist. In der praktischen Umsetzung geht somit die Süße sowohl mit ihren Originalwerten für  $x_1$  als auch mit den quadrierten Werten für  $x_2$  in die Schät-

zung der Regressionskoeffizienten  $b_0$ ,  $b_1$  und  $b_2$  ein. Allgemein zeigt ein Vergleich des Bestimmtheitsmaßes einer nicht-linearen Regression mit dem der linearen Regression  $y = b_0 + b_1 \cdot x$ , ob von einem nicht-linearen statt linearen Zusammenhang ausgegangen werden sollte.

## Bestimmung von Treibertypen mittels nicht-linearer Regression

Im Hinblick auf den Zusammenhang zwischen dem Leistungsniveau einer Eigenschaft und der Zufriedenheit mit einem Produkt oder einer Dienstleistung werden im Allgemeinen drei Treibertypen unterschieden:

- **Basistreiber:** Es wird ein Mindestniveau erwartet. Wird dieses unterschritten, führt dies zu Unzufriedenheit. Wird es übertroffen, resultiert daraus aber keine Zufriedenheit.
- **Begeisterungstreiber:** An die Eigenschaft gibt es keine Erwartungen. Erreicht sie dennoch ein gutes Leistungsniveau, führt dies zu Zufriedenheit. Falls nicht, bleibt dies ohne Auswirkung.
- **Linearer Treiber:** Entsprechend dem Leistungsniveau können sowohl Zufriedenheit als auch Unzufriedenheit entstehen.

Die Kategorisierung der Treiber mit einer Kano-Analyse bedeutet hohen Erhebungsaufwand. Liegen die Zufriedenheit und die Beurteilung der Leistungsniveaus jedoch bereits vor, können die Treibertypen mittels Penalty- & Reward-Analyse, aber ebenso mithilfe der nicht-linearen Regression bestimmt werden.

Dazu wird eine quadratische Regression der Gesamtzufriedenheit auf das Leistungsniveau einer Eigenschaft durchgeführt. Der Koeffizient  $b_2$  bestimmt dann den Treibertyp (siehe Abbildung 2). Ist  $b_2$  positiv, ist die Parabel nach unten geöffnet. Der linke Ast spiegelt somit die Definition eines Basistreibers wider. Ein Begeisterungstreiber wird durch den rechten Ast einer bei negativem  $b_2$  nach oben geöffneten Parabel repräsentiert. Ist  $b_2$  nahe Null, so liegt ein linearer Treiber vor.

In Ausgabe 5/2019: Faktorenanalyse

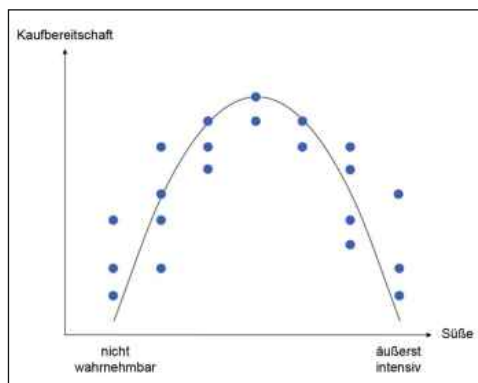


Abbildung 1: Quadratischer Zusammenhang zwischen Süße und Kaufbereitschaft



## Literatur

- von der Lippe, P.: Regressionsanalyse. In: Deskriptive Statistik, Stuttgart, Jena, 1993, S. 258-301.
- Schimmelpfennig, H.: Bekannte, aktuelle und neue Anforderungen an Treiberanalysen. In: Keller, B. et al. (Hrsg.): Marktforschung der Zukunft – Mensch oder Maschine?, Wiesbaden, 2016, S. 231-243.
- Strack, M.; Heise, N.; Karger, J.: Kanos Begeisterungsfaktoren mit quadratischem Term in der Treiberanalyse auffinden. Poster zur FG-Tagung A&O/Wirtschaftspsychologie, Wien, 2009

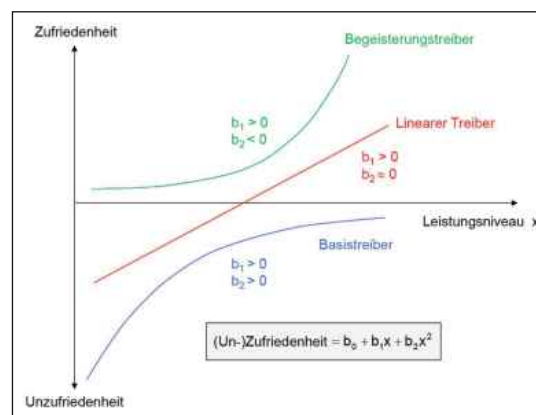


Abbildung 2: Darstellung der Treibertypen als lineare und quadratische Funktionen (Quelle: Schimmelpfennig (2016))