KI für Alle 2: Verstehen, Bewerten, Reflektieren

Themenblock Prognosemodelle (Klassifikation und Regression):
02_03Verfahren_Regression_01

Die lineare Regression

Erarbeitet von
Dr. Katja Theune

Lernziele

- Du kannst die Geradengleichung erläutern
- Du kannst die Idee der Methode der kleinsten Quadrate erläutern
- Du kannst anhand eines einfachen Beispiels den prognostizierten Output für eine neue Beobachtung berechnen
- Du kannst Vor- und Nachteile der linearen Regression erläutern

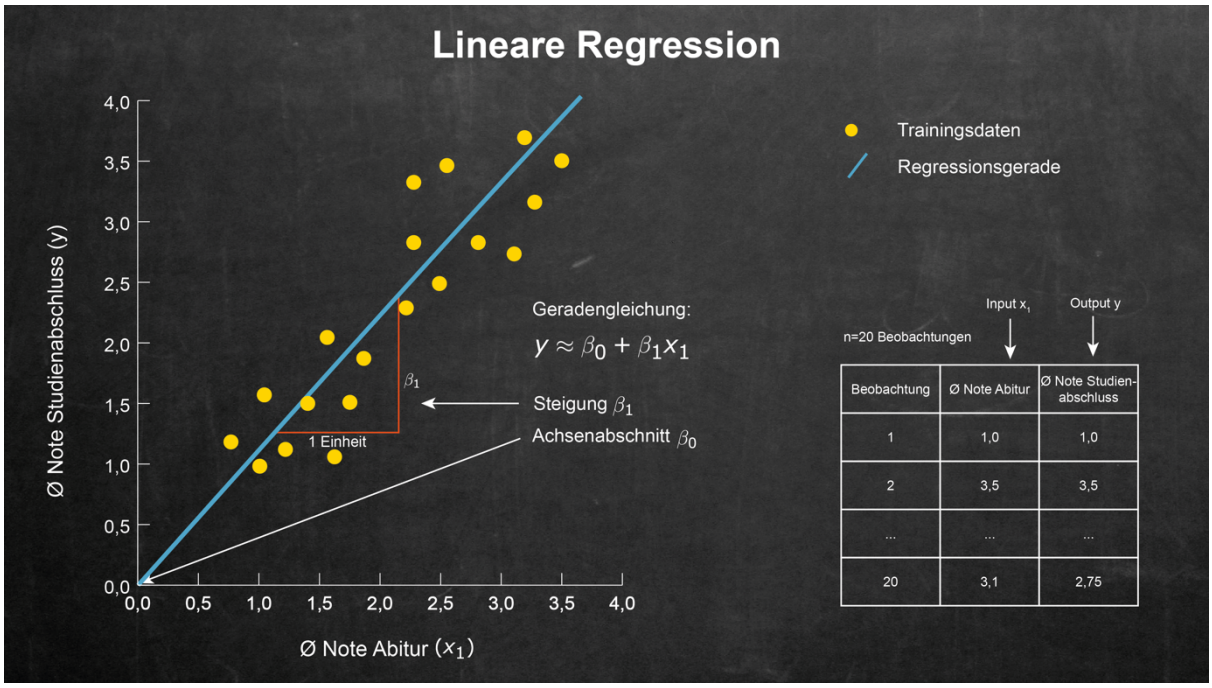
CC BY 4.0

Inhalt

Einstieg

Neben der Prognosegenauigkeit ist auch das Wissen über Zusammenhänge zwischen einem Output und verschiedenen Inputs in vielen Anwendungsszenarien von großer Bedeutung. Um diese Zusammenhänge herauszufinden, eignet sich z. B. eine lineare Regression.

Lineare Regression: Geradengleichung



Bei der linearen Regression nehmen wir einen linearen Zusammenhang zwischen einem metrischen Output und den Inputs an.

Zur Veranschaulichung verwenden wir einen fiktiven Datensatz im Kontext unserer Eingangsfragestellung. Sagen wir, er enthält 20 Beobachtungen, das sind hier unsere Studierenden. Als einzigen Input nehmen wir die durchschnittliche Abiturnote. Der Output muss hier metrisch sein, wir schauen uns daher die durchschnittliche Note zum Abschluss des Studiums an. Wir wollen jetzt wissen, welcher Zusammenhang zwischen der Abiturnote und der Studienabschlussnote besteht. Wir sehen, dass der Zusammenhang annähernd linear ist und sich durch eine Gerade abbilden lässt.

Diese Gerade wird durch folgende sogenannte Geradengleichung beschrieben (siehe Grafik).

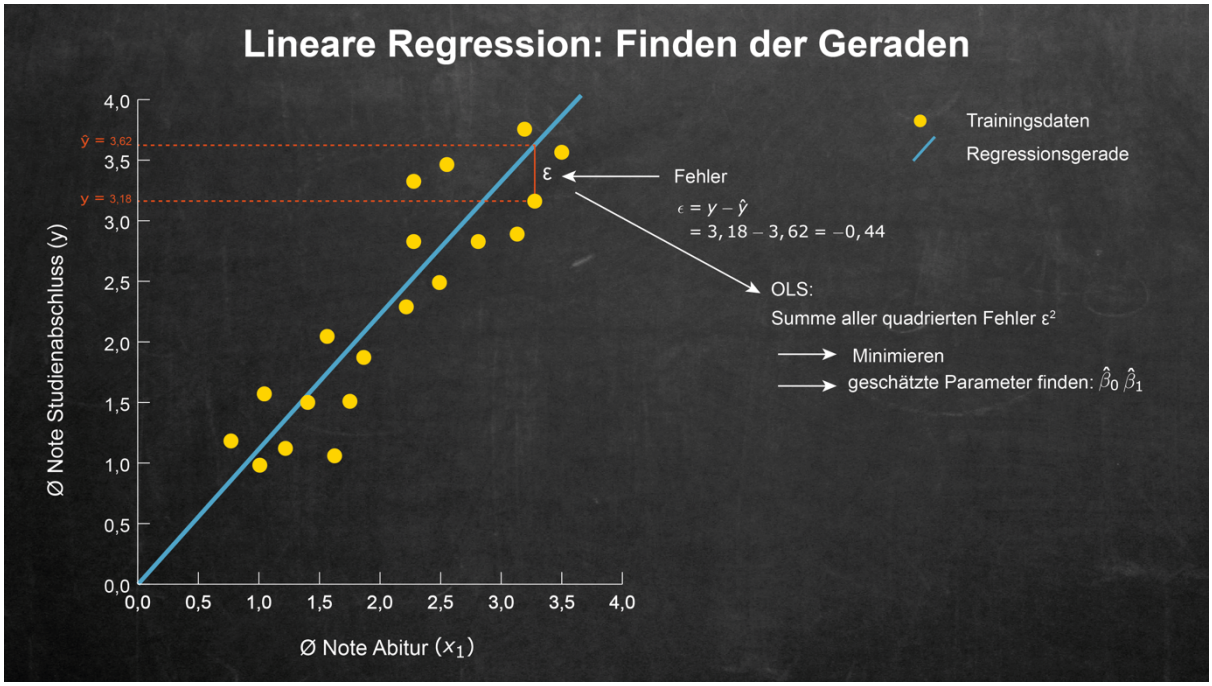
y sind die Output-Werte für unsere Beobachtungen, also die durchschnittlichen Noten zum Studienabschluss. x_1 sind die Werte unseres einzigen Inputs, also hier die Abiturnoten. Das gewellte Gleichheitszeichen bedeutet „ungefähr gleich".

β_0 ist der vertikale Achsenabschnitt. Er gibt den Wert von y an, wenn $x_1 = 0$ ist. Also hier den Wert der Studienabschlussnote, wenn die Abiturnote gleich 0 ist. Die Sinnhaftigkeit dessen lassen wir hier mal außen vor.
β_1 ist der Steigungsparameter. Er gibt an, um wie viel y sich verändert, wenn x_1 sich um eine Einheit verändert. Hier wäre das die Veränderung der Studienabschlussnote, wenn sich die Abiturnote um eine Einheit verändert. Dazu kommen wir aber noch.

Finden der Geraden: Methode der kleinsten Quadrate

Die beiden Betas (β) sind bisher noch unbekannte Parameter und müssen anhand der Daten geschätzt bzw. gelernt werden. Wir kennzeichnen sie üblicherweise mit einem Dach $\widehat{\ }$.

Wir suchen jetzt also die besten Werte für den Achsenabschnitt $\hat{\beta}_0$ und die Steigung $\hat{\beta}_1$. Wir wollen also die Gerade finden, die sich möglichst gut in die Punktwolke einpasst und gute Prognosen macht. Wir bezeichnen die prognostizierten Outputs ebenfalls mit $\hat{y}$.



Um die beste Gerade zu finden, müssen wir die vertikalen Abstände zwischen den Beobachtungen und der Geraden minimieren. Diese Abstände berechnen sich als Differenz zwischen den wahren Werten y aus den Trainingsdaten und den mit der Geraden prognostizierten Werten $\hat{y}$. Diese Abstände entsprechen also den Prognosefehlern. Wir bezeichnen sie mit Epsilon (ϵ) und man nennt sie auch Residuen. Diese Fehler sollen natürlich gering sein.

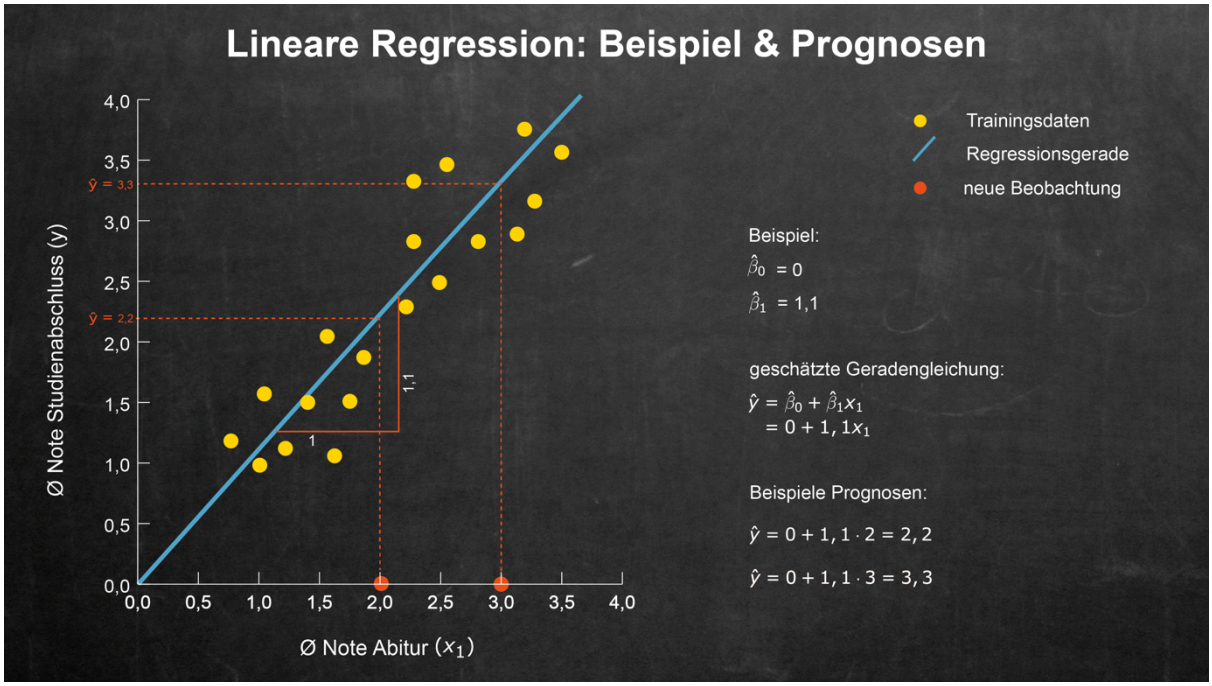
Für diese Beobachtung hier (siehe Grafik) ergibt sich der Fehler z. B. als Differenz zwischen dem wahren Wert 3,18 und dem prognostizierten Wert 3,62 und ist damit gleich - 0,44.

Für das Finden der besten Gerade wird häufig die Methode der kleinsten Quadrate verwendet, im englischen Ordinary Least Squares und meistens mit OLS Methode abgekürzt.

Sie nutzt dafür die Summe der quadrierten Fehler bzw. Abstände über alle Beobachtungen, hier also 20. Wir quadrieren diese Abstände, damit sich positive und negative Abstände nicht gegenseitig aufheben können.

Wir suchen jetzt die Parameter $\hat{\beta}_0$ und $\hat{\beta}_1$, welche diese Summe minimieren. Wie genau man dann zu den geschätzten Parametern kommt, ist hier aber erstmal nicht von Bedeutung.

Lineare Regression: Beispiel und Prognose



Sagen wir, für unser fiktives Beispiel haben wir jetzt für $\hat{\beta}_0$ einen Wert von 0 und für $\hat{\beta}_1$ einen Wert von 1,1 erhalten. Wir haben also einen Achsenabschnitt von 0. D. h., die Gerade berührt bei einer Abiturnote von 0 die vertikale Achse ebenfalls bei 0. Die Steigung beträgt 1,1 und besagt, dass mit der Steigung der Abiturnote um eine Einheit, die Studienabschlussnote um 1,1 Einheiten steigt. Wir sehen darin also die Stärke und auch die Richtung des Zusammenhangs zwischen Input und Output.

Wir erhalten damit die folgende geschätzte Geradengleichung (siehe Grafik). Dies ergibt dann genau die Gerade, die wir hier sehen. Damit können nun die Outputs für bestimmte Beobachtungen prognostiziert werden. Hier würden wir z. B. für jemanden mit einer Abiturnote von 2,0 eine Studienabschlussnote von 2,2 prognostizieren. Und für jemanden mit einer Abiturnote von 3,0 eine Studienabschlussnote von 3,3.

Multiple lineare Regression

Die multiple lineare Regression ist eine Erweiterung der einfachen linearen Regression. Hier haben wir wieder nur einen Output, aber mehrere Inputs. Wir schreiben dann:

$$y \approx \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots$$

Optisch hätten wir dann hier nicht eine Gerade, sondern im dreidimensionalen Raum z. B. eine Ebene, in höher dimensionalen Räumen eine Hyperebene. In unserem Beispiel könnten wir uns als weitere Inputs z. B. den Bildungshintergrund der Eltern oder die Zufriedenheit mit dem Studium vorstellen.

Diskussion, Vor- und Nachteile

Positiv hervorzuheben ist, dass sich mit der linearen Regression Zusammenhänge zwischen Output und Inputs darstellen lassen und wir einen Eindruck von Stärke und Richtung dieser Zusammenhänge erhalten. Das ist gerade für Anwendungsfälle sinnvoll, bei denen es nicht nur auf die Prognosegenauigkeit ankommt, sondern wir auch Informationen über wichtige Inputs und deren Einfluss benötigen.

Ein Nachteil ist, dass wir im Vorhinein einen linearen Zusammenhang zwischen Output und Inputs spezifizieren. Sollte diese Annahme falsch sein, kann das zu verzerrten Ergebnissen führen. Häufig erfordert die Anwendung der linearen Regression eine aufwendigere Datenaufbereitung, da sie z. B. nicht mit fehlenden Werten umgehen kann und auch nicht mit allen Arten von Inputs.

Abschluss

Wir haben jetzt das für das Verfahren der linearen Regression wichtige Konzept der Geradengleichung kennengelernt und mit Hilfe der Methode der kleinsten Quadrate können wir die beste Gerade finden. Mit diesem Verfahren erhalten wir neben Prognosen auch Informationen über mögliche Zusammenhänge von Inputs und Outputs.

Weiterführendes Material

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3. Auflage). Morgan Kaufmann.

James, G., Witten, D., Hastie, T., & Tibshirani, R., & Tylor, J. (2023). *An Introduction to Statistical Learning - with Applications in Python*. Springer.

Lantz, B. (2015). *Machine learning with R* (2. Auflage). Packt Publishing Ltd, Birmingham.

Disclaimer

Transkript zu dem Video „Prognosemodelle (Klassifikation und Regression): Die lineare Regression“, Dr. Katja Theune.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz CC-BY 4.0 veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.