Polynomiale Regression

Künstliche neuronale Netze

Marina Klein

09.11.2018

Inhaltsverzeichnis

Polynomiale Regression

- Lineare Regression

$$y = \beta_0 + \beta_1 x + \epsilon$$

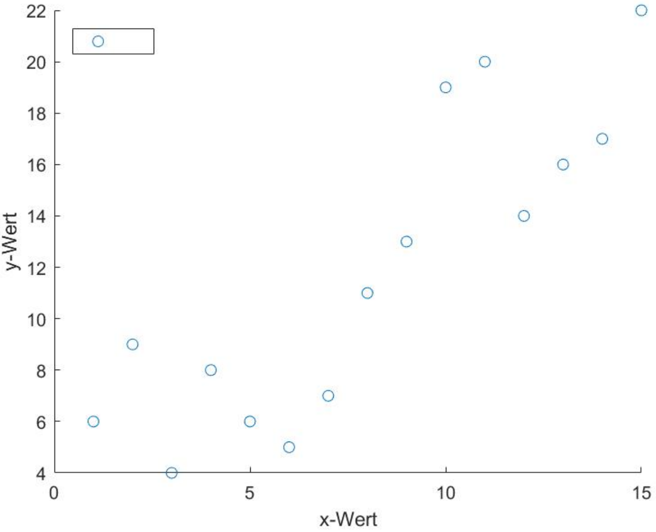
- Lineare Regression

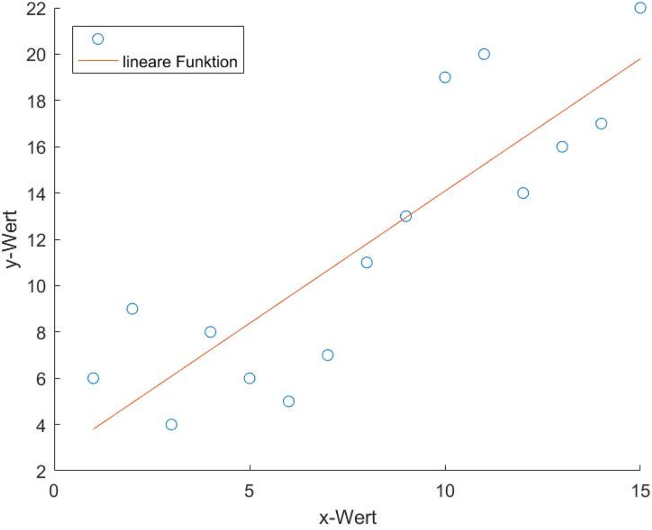
$$y = \beta_0 + \beta_1 x + \epsilon$$

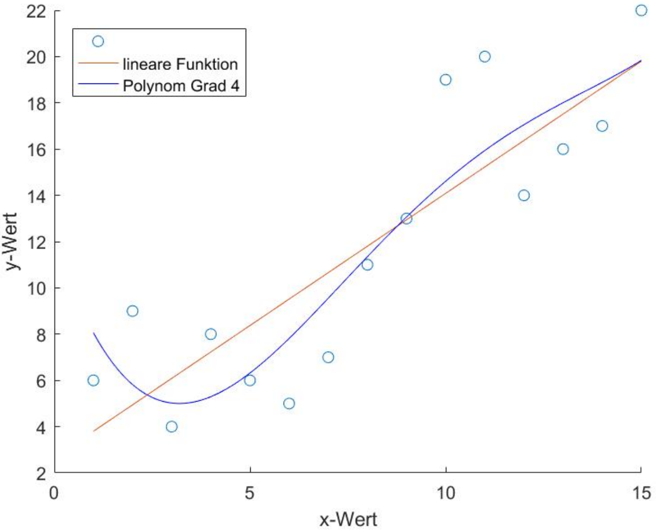
- Polynomiale Regression:

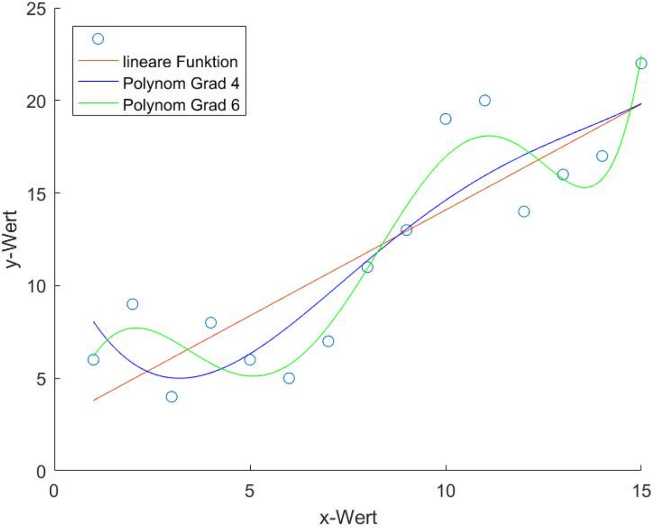
$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + ... + \beta_d x^d + \epsilon$$

,ϵ ist der Fehler


22
20
18
16
14
12
10
8
6
4
y-Wert
0
5
10
15
x-Wert


lineare Funktion
y-Wert
x-Wert
22
20
18
16
14
12
10
8
6
4
2
0
5
10
15


lineare Funktion
Polynom Grad 4
y-Wert
x-Wert
22
20
18
16
14
12
10
8
6
4
2
0
5
10
15



lineare Funktion
Polynom Grad 4
Polynom Grad 6
y-Wert
x-Wert
0
5
10
15
20
25

Die Stufenfunktion

- konstruiere $K+1$ neue Variablen mit den Punkten $c_1, c_2, ..., c_K$ aus X:

- konstruiere $K+1$ neue Variablen mit den Punkten $c_1, c_2, ..., c_K$ aus X:

$$
\begin{aligned}
C_0(X) &= I(X < c_1), \\
C_1(X) &= I(c_1 \leq X < c_2), \\
C_2(X) &= I(c_2 \leq X < c_3), \\
&... \\
C_{K-1}(X) &= I(c_{K-1} \leq X < c_K), \\
C_K(X) &= I(c_K \leq X)
\end{aligned}
$$

, $I(\cdot)$ Indikatorfunktion

- konstruiere $K+1$ neue Variablen mit den Punkten $c_1, c_2, ..., c_K$ aus X:

$$
\begin{aligned}
C_0(X) &= I(X < c_1), \\
C_1(X) &= I(c_1 \leq X < c_2), \\
C_2(X) &= I(c_2 \leq X < c_3), \\
&\dots \\
C_{K-1}(X) &= I(c_{K-1} \leq X < c_K), \\
C_K(X) &= I(c_K \leq X)
\end{aligned}
$$

, $I(\cdot)$ Indikatorfunktion

- $C_0(X) = I(X < c_1) = \begin{cases} 1 & \text{, falls } X < c_1, \\ 0 & \text{, sonst.} \end{cases}$

- konstruiere $K+1$ neue Variablen mit den Punkten $c_1, c_2, ..., c_K$ aus X:

$$\begin{aligned} C_0(X) &= I(X < c_1), \\ C_1(X) &= I(c_1 \leq X < c_2), \\ C_2(X) &= I(c_2 \leq X < c_3), \\ &... \\ C_{K-1}(X) &= I(c_{K-1} \leq X < c_K), \\ C_K(X) &= I(c_K \leq X) \end{aligned}$$

, $I(\cdot)$ Indikatorfunktion

- $C_0(X) = I(X < c_1) = \begin{cases} 1 & \text{, falls } X < c_1, \\ 0 & \text{, sonst.} \end{cases}$
- $C_0(X) + C_1(X) + ... + C_K(X) = 1$

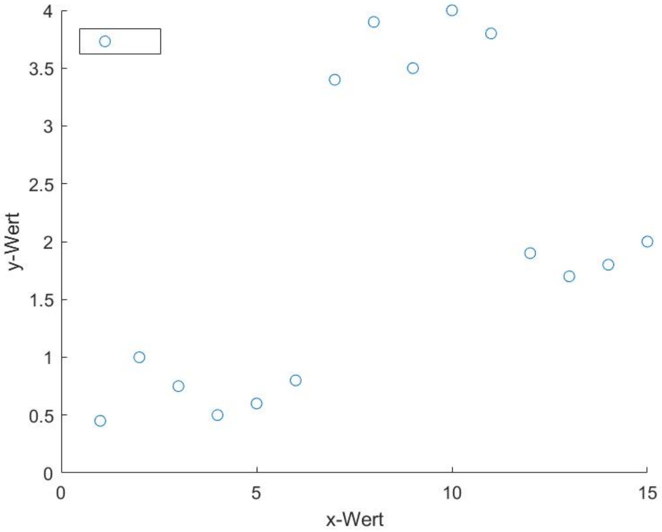
$$y = \beta_0 + \beta_1 C_1(x) + \beta_2 C_2(x) + ... + \beta_K C_K(x) + \epsilon$$

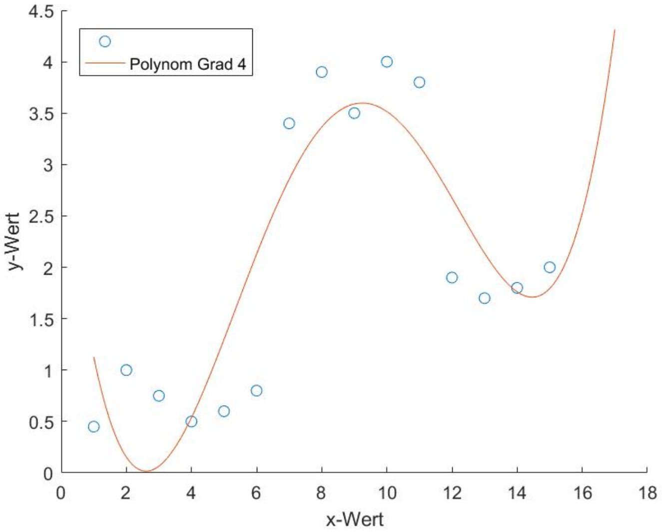
$$y = \beta_0 + \beta_1 C_1(x) + \beta_2 C_2(x) + ... + \beta_K C_K(x) + \epsilon$$

- $X < c_1 \Rightarrow Y = \beta_0$

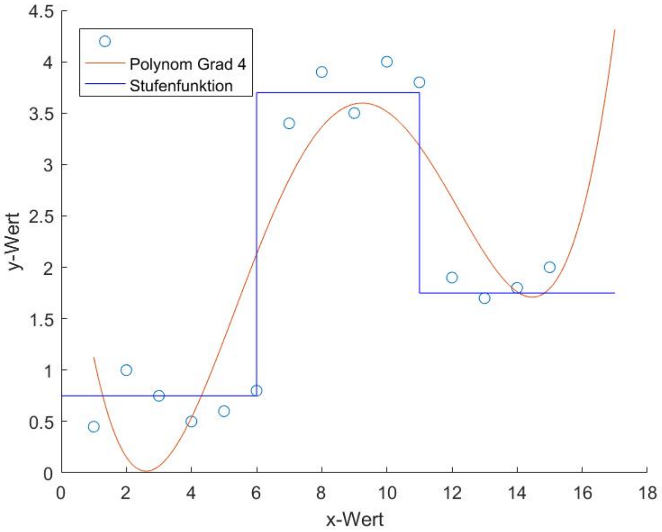
$$y = \beta_0 + \beta_1 C_1(x) + \beta_2 C_2(x) + ... + \beta_K C_K(x) + \epsilon$$

- $X < c_1 \Rightarrow Y = \beta_0$
- $Y = \beta_0 + \beta_j \ , \ c_j \leq X < c_{j+1}$


4
3.5
3
2.5
2
1.5
1
0.5
0
y-Wert
0
5
10
15
x-Wert



Polynom Grad 4
y-Wert
x-Wert
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
2
4
6
8
10
12
14
16
18



4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
2
4
6
8
10
12
14
16
18
y-Wert
x-Wert
Polynom Grad 4
Stufenfunktion

Die Basisfunktion

- Finde eine Familie von Funktionen oder Transformationen, die auf eine Variable X angewendet werden kann:

$$b_1(X), b_2(X), ..., b_K(x)$$

$\Rightarrow\; y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + \beta_3 b_3(x) + ... + \beta_K b_K(x) + \epsilon$
, Basisfunktionen $b_1(\cdot), b_2(\cdot), ..., b_K(\cdot)$ sind bekannt und fest

- Finde eine Familie von Funktionen oder Transformationen, die auf eine Variable X angewendet werden kann:

$$b_1(X), b_2(X), ..., b_K(x)$$

$\Rightarrow \; y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + \beta_3 b_3(x) + ... + \beta_K b_K(x) + \epsilon$
, Basisfunktionen $b_1(\cdot), b_2(\cdot), ..., b_K(\cdot)$ sind bekannt und fest

- Basisfunktionen für
  - Polynomiale Regression: $b_j(x) = x^j$

- Finde eine Familie von Funktionen oder Transformationen, die auf eine Variable X angewendet werden kann:

$$b_1(X), b_2(X), ..., b_K(x)$$

$\Rightarrow \; y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + \beta_3 b_3(x) + ... + \beta_K b_K(x) + \epsilon$
, Basisfunktionen $b_1(\cdot), b_2(\cdot), ..., b_K(\cdot)$ sind bekannt und fest

- Basisfunktionen für
  - Polynomiale Regression: $b_j(x) = x^j$
  - stückweise konstante Funktionen: $b_j(x) = I(c_j \leq x < c_{j+1})$

Spline-Regression

stückweise Polynome

- Finde Polynome mit einem niedrigen Grad für die verschiedenen Bereiche von X

stückweise Polynome

- Finde Polynome mit einem niedrigen Grad für die verschiedenen Bereiche von X
- stückweise kubisches Polynom:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon$$

, $\beta_0,\ \beta_1,\ \beta_2$ *und* β_3 für verschiedene Bereiche von X

stückweise Polynome

- Finde Polynome mit einem niedrigen Grad für die verschiedenen Bereiche von X
- stückweise kubisches Polynom:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon$$

  , $\beta_0,\ \beta_1,\ \beta_2\ und\ \beta_3$ für verschiedene Bereiche von X
- Knoten: Punkte an denen sich die Koeffizienten ändern

Beispiele für $d = 3$

Beispiele für $d = 3$

- Keine Knoten:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon$$

Beispiele für $d = 3$

- Keine Knoten:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon$$

- Ein Knoten am Punkt c:

$$y = \begin{cases} \beta_{01} + \beta_{11} x + \beta_{21} x^2 + \beta_{31} x^3 + \epsilon & \text{, falls } x < c, \\ \beta_{02} + \beta_{12} x + \beta_{22} x^2 + \beta_{32} x^3 + \epsilon & \text{, falls } x \geq c. \end{cases}$$

Beispiele für $d = 3$

- Keine Knoten:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \epsilon$$

- Ein Knoten am Punkt c:

$$y = \begin{cases} \beta_{01} + \beta_{11} x + \beta_{21} x^2 + \beta_{31} x^3 + \epsilon & \text{, falls } x < c, \\ \beta_{02} + \beta_{12} x + \beta_{22} x^2 + \beta_{32} x^3 + \epsilon & \text{, falls } x \geq c. \end{cases}$$

- Zwei Knoten an den Stellen c,d ($c < d$):

$$y_i = \begin{cases} \beta_{01} + \beta_{11} x + \beta_{21} x^2 + \beta_{31} x^3 + \epsilon & \text{, falls } x < c, \\ \beta_{02} + \beta_{12} x + \beta_{22} x^2 + \beta_{32} x^3 + \epsilon & \text{, falls } c \leq x < d, \\ \beta_{03} + \beta_{13} x + \beta_{23} x^2 + \beta_{33} x^3 + \epsilon & \text{, falls } x \geq d. \end{cases}$$

Beispiele für lineare Funktionen

Beispiele für lineare Funktionen

- Kein Knoten:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Beispiele für lineare Funktionen

- Kein Knoten:

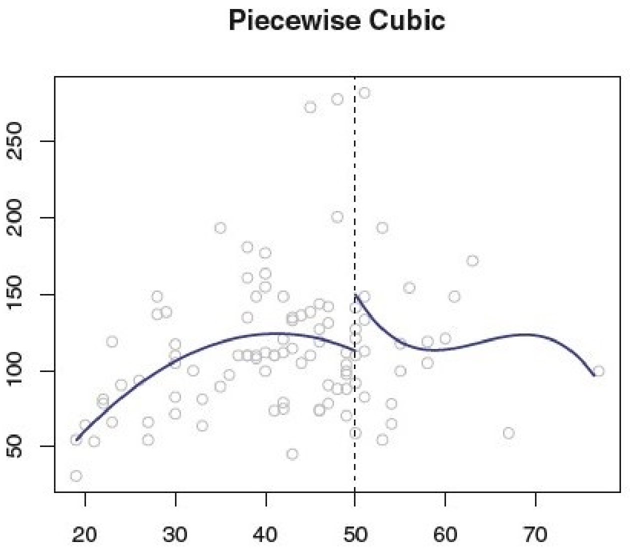
$$y = \beta_0 + \beta_1 x + \epsilon$$

- Ein Knoten an der Stelle c:

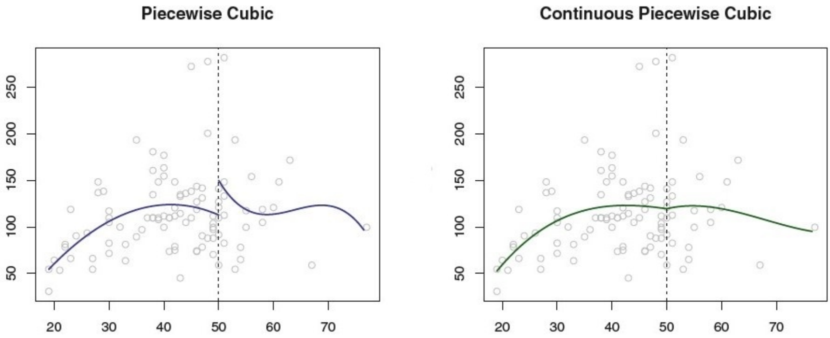
$$y = \begin{cases} \beta_{01} + \beta_{11} x + \epsilon & \text{, falls } x < c, \\ \beta_{02} + \beta_{12} x + \epsilon & \text{, falls } x \geq c. \end{cases}$$

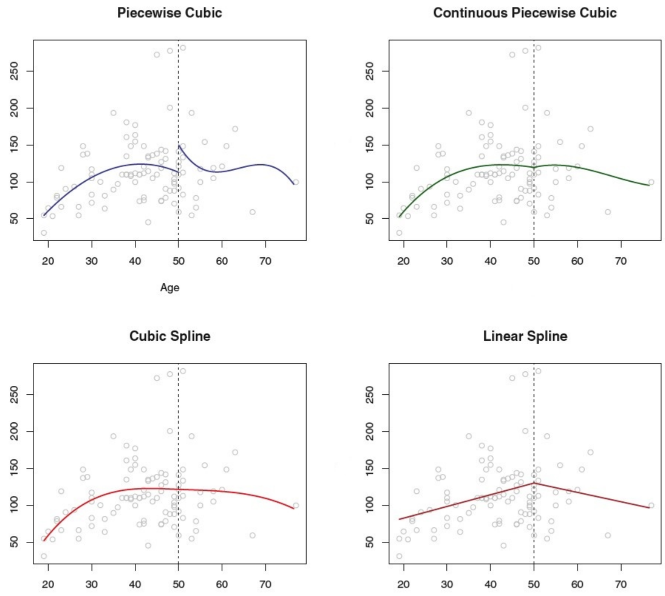
- mehr Knoten $\Rightarrow$ flexibleres stückweises Polynom

- mehr Knoten $\Rightarrow$ flexibleres stückweises Polynom
- bei K unterschiedlichen Knoten in X erhalten wir $K + 1$ verschiedene Polynome


Piecewise Cubic
50
100
150
200
250
20
30
40
50
60
70

Bedingungen und Splines




Piecewise Cubic
Continuous Piecewise Cubic
Cubic Spline
Linear Spline
Age
50
100
150
200
250
20
30
40
50
60
70

Definition eines Splines vom Grad d

Ein stückweises Polynom vom Grad d mit Stetigkeit der Ableitungen bis Grad $d-1$ in jedem Knoten ist ein Spline vom Grad d.

Die Spline Basis Repräsentation

Die Spline Basis Repräsentation

- Kubischer Spline mit K Knoten:

$$y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + \ldots + \beta_{K+3} b_{K+3}(x) + \epsilon$$

Die Spline Basis Repräsentation

- Kubischer Spline mit K Knoten:

$$y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + ... + \beta_{K+3} b_{K+3}(x) + \epsilon$$

- Beginnen mit einer Basis für Kubische Polynome: x, x^2, x^3

Die Spline Basis Repräsentation

- Kubischer Spline mit K Knoten:

$$y = \beta_0 + \beta_1 b_1(x) + \beta_2 b_2(x) + ... + \beta_{K+3} b_{K+3}(x) + \epsilon$$

- Beginnen mit einer Basis für Kubische Polynome: x, x^2, x^3
- Füge einen abgeschnittene Potenzfunktion pro Knoten hinzu:

$$h(x,\xi) = (x-\xi)_+^3 = \begin{cases} (x-\xi)^3 & \text{, falls } x > \xi, \\ 0 & \text{, sonst.} \end{cases}$$

,ξ ist der Knoten

Kubischer Spline mit K Knoten

Kubischer Spline mit K Knoten

- nehmen für einen Abschnitt die Regression mit der kleinsten Quadratur und $3 + K$ Variablen

Kubischer Spline mit K Knoten

- nehmen für einen Abschnitt die Regression mit der kleinsten Quadratur und $3 + K$ Variablen
- Form:$X,\ X^2,\ X^3,\ h(X, \xi_1),\ h(X, \xi_2), ..., h(X, \xi_K)$, $\xi_1, ..., \xi_K$ sind die Knoten

Kubischer Spline mit K Knoten

- nehmen für einen Abschnitt die Regression mit der kleinsten Quadratur und $3 + K$ Variablen
- Form:$X,\ X^2,\ X^3,\ h(X, \xi_1),\ h(X, \xi_2), ..., h(X, \xi_K)$, $\xi_1, ..., \xi_K$ sind die Knoten
- $K + 4$ Koeffizienten
- Freiheitsgrad $K + 4$

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)
  - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)
  - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten
  - Mache eine Vorhersage für den ausgelassenen Teil

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)
  - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten
  - Mache eine Vorhersage für den ausgelassenen Teil
  - Wiederhole dies, bis jede Beobachtung einmal ausgelassen wurde

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)
  - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten
  - Mache eine Vorhersage für den ausgelassenen Teil
  - Wiederhole dies, bis jede Beobachtung einmal ausgelassen wurde
  - Berechne RSS

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
  - Entferne einen Teil der Daten (10%)
  - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten
  - Mache eine Vorhersage für den ausgelassenen Teil
  - Wiederhole dies, bis jede Beobachtung einmal ausgelassen wurde
  - Berechne RSS
  - Kann für verschiedene Anzahlen von Knoten wiederholt werden

Wahl der Anzahl und Position der Knoten

- Wie wähle ich die Anzahl der Knoten?
- Verschiedene Anzahlen von Knoten ausprobieren und dann die am besten passendste Kurve wählen
- Kreuzvalidierung:
    - Entferne einen Teil der Daten (10%)
    - Finde einen passenden Spline mit einer gewählten Anzahl von Knoten zu den Daten
    - Mache eine Vorhersage für den ausgelassenen Teil
    - Wiederhole dies, bis jede Beobachtung einmal ausgelassen wurde
    - Berechne RSS
    - Kann für verschiedene Anzahlen von Knoten wiederholt werden
    - Der Wert für K mit dem kleinsten RSS Wert wird ausgewählt

Vergleich zur Polynomialen Regression

Vergleich zur Polynomialen Regression

- Die Regressiom mit Splines gibt meist ein besseres Ergebnis als die Polynomiale Regression

Vergleich zur Polynomialen Regression

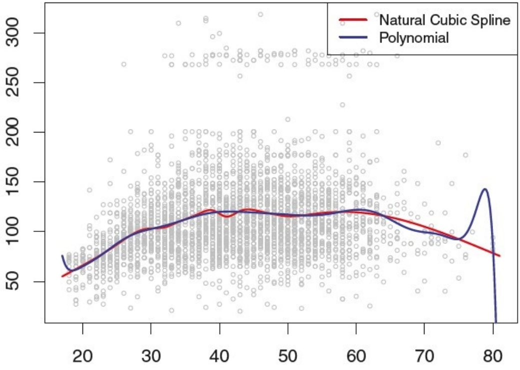
- Die Regressiom mit Splines gibt meist ein besseres Ergebnis als die Polynomiale Regression
- Polynome brauchen einen hohen Grad für eine flexible Form

Vergleich zur Polynomialen Regression

- Die Regressiom mit Splines gibt meist ein besseres Ergebnis als die Polynomiale Regression
- Polynome brauchen einen hohen Grad für eine flexible Form
- Splines erhöhen ihre Flexibilität durch eine höhere Anzahl an Knoten, der Grad bleibt jedoch gleich

Vergleich zur Polynomialen Regression

- Die Regressiom mit Splines gibt meist ein besseres Ergebnis als die Polynomiale Regression
- Polynome brauchen einen hohen Grad für eine flexible Form
- Splines erhöhen ihre Flexibilität durch eine höhere Anzahl an Knoten, der Grad bleibt jedoch gleich
- Es können mehr Knoten in Regionen mit einer starken Änderung verwendet werden und weniger in Regione, in welchen die Kurve stabiler ist


Natural Cubic Spline
Polynomial
50
100
150
200
250
300
20
30
40
50
60
70
80

Literatur

Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani:
An Introduction to Statistical Learning
with Applications in R

Vielen Dank für die Aufmerksamkeit!