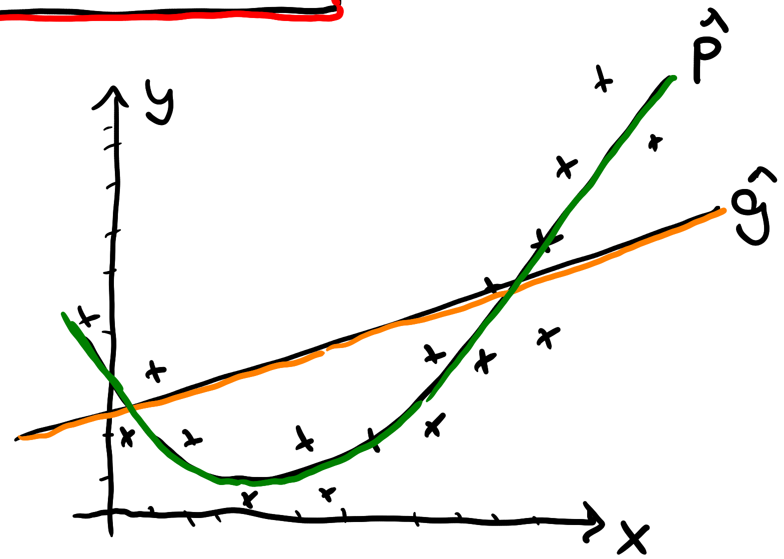
## 2.1.3) Linearisierbare Modelle

Bisher sind wir davon ausgegangen, daß die Modellfunktion $f(x_1,\dots,x_m)$ mit $\boxed{Y = f(X_1,\dots,X_m)+\varepsilon}$ linear ist, also

$$\boxed{f(X_1,\dots,X_m) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m}$$

Das ist aber leider nur selten der Fall ...

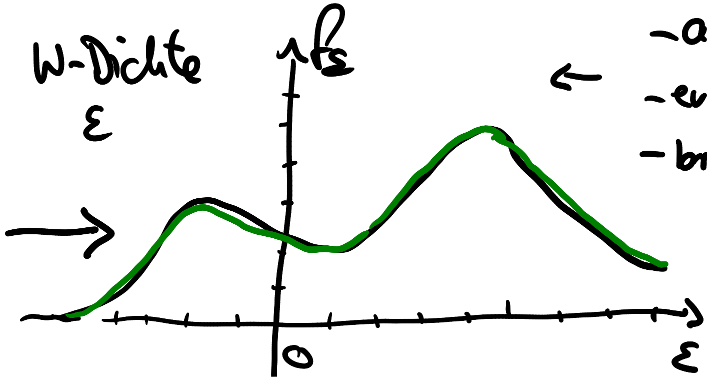



Bsp.:
- Jäger-Beute Beziehung
- Bakterienwachstum (Epidemie)
- periodische Abläufe

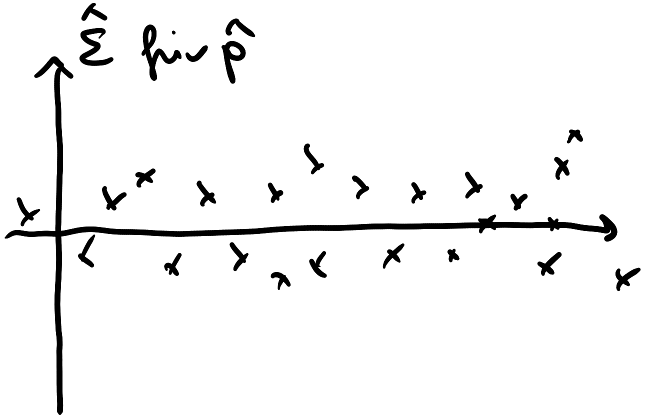
→ Schätzgerade $\hat{g}$ läßt sich (mathematisch gesehen) problemlos für die Daten berechnen

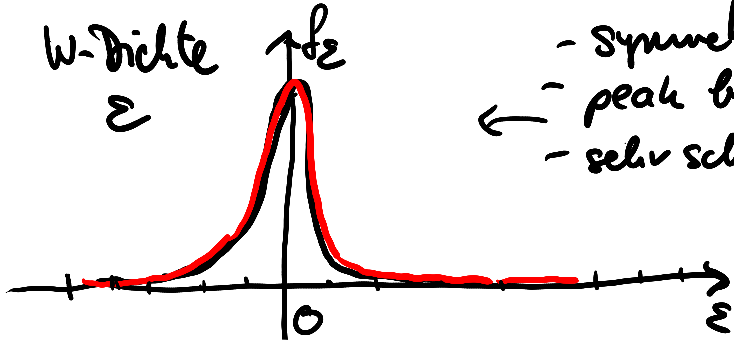
→ $\hat{p}$ paßt aber deutlich besser zur Zusammenhangsstruktur

Eine Analyse der Residuen könnte wie folgt aussehen:


- asymmetrisch
- evtl. mehrere peaks
- breit




- symmetrisch
- peak bei 0
- sehr schmal

⇨ Die Analyse der Residuen zeigt für $\hat{g}$ eine deutliche Systematik

$\hat{=}$ Systematik des Fehlers $\varepsilon$ $\hat{=}$ bias des linearen Modelles

**IDEE** Verwende für das Modell <u>keine</u> lineare Funktion sondern ein Polynom, Exponentialfunktion ...

Polynom: in $X_1, \dots, X_m$

$$\boxed{f(X_1, \dots, X_m) = b_0 + b_1 X_1 + b_2 X_1^2 + \dots + b_\ell X_m^\ell}$$

← ziemlich komplex!

auch vermischte Terme $b_j X_i X_k$ möglich!

**2 FRAGEN**

1. Wie findet man die unbekannten Parameter $b_j$?
2. Welche Modellstruktur ist geeignet, d.h. welche $X_j$ und bis zu welchem Grad bzw. in welchen Kombis $X_j X_k$?

Wir starten einfach:

I. Polynom (univariat) in $X$

Modell: $$\boxed{Y = b_0 + b_1 X + b_2 X^2 + \dots + b_k X^k + \varepsilon}$$

$k$: Grad des Polynoms

für Datenpunkte $i = 1, \dots, n$ $(x_i, y_i)$ gilt dann

$$\boxed{y_i = b_0 + b_1 x_i + b_2 x_i^2 + \dots + b_k x_i^k + \varepsilon_i}$$

**empirische Beziehung**

Lösung ?? $\hat{\vec{b}} = \begin{pmatrix} \hat{b}_0 \\ \vdots \\ \hat{b}_k \end{pmatrix} = ??$

IDEE ⇒ Die Struktur des Modells sieht aus wie

(∗) $$\boxed{Y = b_0 + b_1 X + b_2 X^2 + \dots + b_k X^k + \varepsilon}$$

$$\boxed{Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k + \varepsilon}$$

multiv. lineares Modell
Lösung bekannt !!

→ Variablentrafo:

$$\boxed{\begin{array}{l} X := X_1 \\ X^2 := X_2 \\ \vdots \qquad \vdots \\ X^k := X_k \end{array}}$$

(∗) hat (in den neuen Variablen) offensichtlich die gleiche Lösung

$$\boxed{\hat{\vec{b}} = (X^t \cdot X)^{-1} \cdot X^t \cdot \vec{y}}$$

⇝ Datenmatrix $X$ muß entsprechend befüllt werden ...

Es gilt:

$$X = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & & x_2^k \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{bmatrix}$$

$$\begin{array}{cccc} & \uparrow & \uparrow & \cdots & \uparrow \\ & X & X^2 & \cdots & X^k \end{array}$$

$$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

$$\boxed{\hat{\vec{b}} = (X \cdot X^t)^{-1} \cdot X^t \cdot \vec{y}}$$

Eine entsprechende Schätzung für $y_i$ (Prognose) ergibt sich als

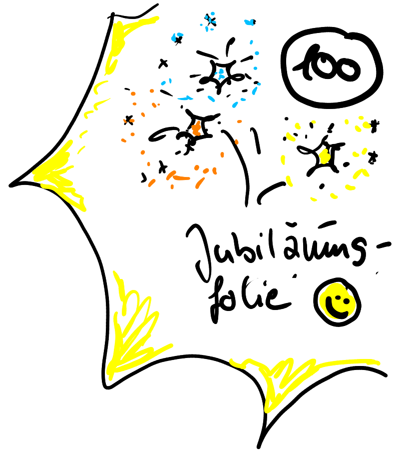
$$\boxed{\hat{y}_i = \hat{b}_0 + \hat{b}_1 x_i + \hat{b}_2 x_i^2 + \ldots + \hat{b}_k x_i^k} \qquad \boxed{\hat{\varepsilon}_i = y_i - \hat{y}_i} \leftarrow \text{Residuum}$$

mit Standardfehler aus

$$\boxed{\hat{\sigma}^2 = \frac{1}{n-(k+1)} \sum_{i=1}^{n} \hat{\varepsilon}_i^2} \leftarrow \text{abhängig vom Grad } k \text{ des Polynoms.}$$

Bemerkung: zu hohe Polynomgrade sind i. A. nicht sinnvoll (overfitting).

Faustregel: $k \leq 4$




## Bsp. 1) Verkaufszahlen

3 Produkte A, B, C
4 Kunden $(i = 1, \ldots, 4)$

| $i$ | $Y$ | $X_1$ | $X_2$ |
|---|---|---|---|
| 1 | 2 | 2 | 2 |
| 2 | 4 | 6 | 2 |
| 3 | 6 | 6 | 4 |
| 4 | 4 | 4 | 4 |
| | A | B | C |

$x_1, x_2, Y$ ZV

≙ "gekaufte Stückzahl A, B, C"

Aus der Anzahl der gekauften B und C soll eine Stückzahl für A ermittelt werden!

Vermutung $\boxed{Y = b_0 + b_1 X_1 + b_2 X_1^2 + \varepsilon}$

$$\vec{y} = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 4 \end{pmatrix} \qquad X = \begin{array}{c} \begin{bmatrix} 1 & 2 & 4 \\ 1 & 6 & 36 \\ 1 & 6 & 36 \\ 1 & 4 & 16 \end{bmatrix} \\ \quad\; x_1 \quad x_1^2 \end{array}$$

$$\Rightarrow \hat{\vec{b}} = (X^t \cdot X)^{-1} \cdot X^t \cdot \vec{y} \quad \xrightarrow{\text{Computer}} \quad \hat{\vec{b}} = \begin{pmatrix} -1 \\ 1{,}75 \\ -0{,}125 \end{pmatrix}$$

$$\Rightarrow \hat{y}_i = -1 + 1{,}75\, x_{i_1} - 0{,}125\, x_{i_1}^2 \quad \leftarrow \text{Prognose}$$

$\Rightarrow$

| | $y_i$ | $\hat{y}_i$ | $\hat{\varepsilon}_i$ |
|---|---|---|---|
| 1 | 2 | 2 | 0 |
| 2 | 4 | 5 | -1 |
| 3 | 6 | 5 | 1 |
| 4 | 4 | 4 | 0 |

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n-(k+1)} \sum_{i=1}^{4} \hat{\varepsilon}_i^2 = \frac{1}{4-3} \cdot 2 = 2$$

$$\Rightarrow \hat{\sigma} = \sqrt{2} = 1{,}41$$ ← Standardfehler

Um zu testen, ob $x_1^2$ im Modell Sinn macht

$$\boxed{\begin{aligned} H_0 &: b_2 = 0 \\ H_1 &: b_2 \neq 0 \end{aligned}}$$

mit $\alpha$-Fehler bekannt!

mit Teststatistik:

$$\boxed{\hat{T}_{\hat{b}_2} = \frac{\hat{b}_2 - 0}{\hat{\sigma}_{\hat{b}_2}}}$$

geschätzte Varianz von $b_2$

und krit. Bereich:

$$\boxed{K = \left(-\infty \,;\, -t(n-k)_{1-\frac{\alpha}{2}}\right) \cup \left(t(n-k)_{1-\frac{\alpha}{2}} \,;\, +\infty\right)}$$

$(1-\frac{\alpha}{2})$ Quantil der $t$-Verteilung mit $(n-k)$ FG (Tabelle oder Computer)

**Bemerkung:**

Die geschätzte Varianz von $b_2$ bekommt man aus der Kovarianzmatrix

$$\boxed{\hat{\Sigma} = \hat{\sigma}^2 (X^t \cdot X)^{-1}}$$

die Diagonalelemente von $\hat{\Sigma}$ entsprechen den geschätzten Varianzen von $b_0, b_1, \dots, b_k$

**Bsp. 1) Fortsetzung**

$$\hat{\Sigma} = \begin{bmatrix} 37 & * & * \\ * & 11{,}69 & * \\ * & * & 0{,}17 \end{bmatrix}$$

$\Rightarrow$ Konfidenzintervall für $b_1$:

$$\hat{I}_{b_1} = \left[\hat{b}_1 \pm t(n-k)_{1-\frac{\alpha}{2}} \, \hat{\sigma}_{\hat{b}_1}\right] =$$

$$= \left[1{,}75 \pm 12{,}706 \cdot \sqrt{11{,}69}\right] = [-41{,}69 ; 45{,}19]$$

Tabelle

ganz schön breit!!
$\Rightarrow$ Schätzung ungenau

**Hypothesentest:**

$$\boxed{\begin{array}{l} H_0: b_2 = 0 \\ H_1: b_2 \neq 0 \end{array}}$$

$$\hat{T} = \frac{-0{,}125 - 0}{\sqrt{0{,}17}} = -0{,}3$$

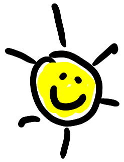
$$K = (-\infty ; -12{,}7) \cup (12{,}7 ; \infty)$$

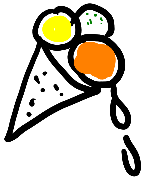
$\Rightarrow \hat{T} = -0{,}3 \notin K \Rightarrow H_0$ annehmen

$\Rightarrow H_1$ nicht bewiesen!! $X_1$ vermutlich nicht relevant

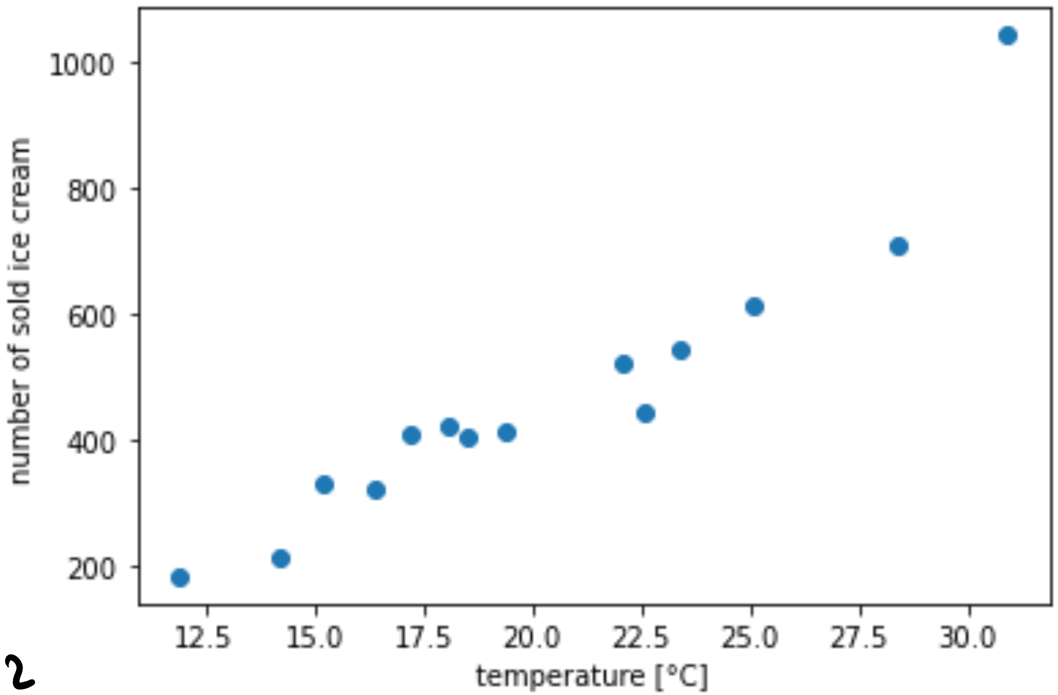
## Bsp.: Eisverkauf

Y (ZV)

je schöner das Wetter 

desto mehr  wird

verkauft




FRAGE Wie genau sieht der Zusammenhang aus?

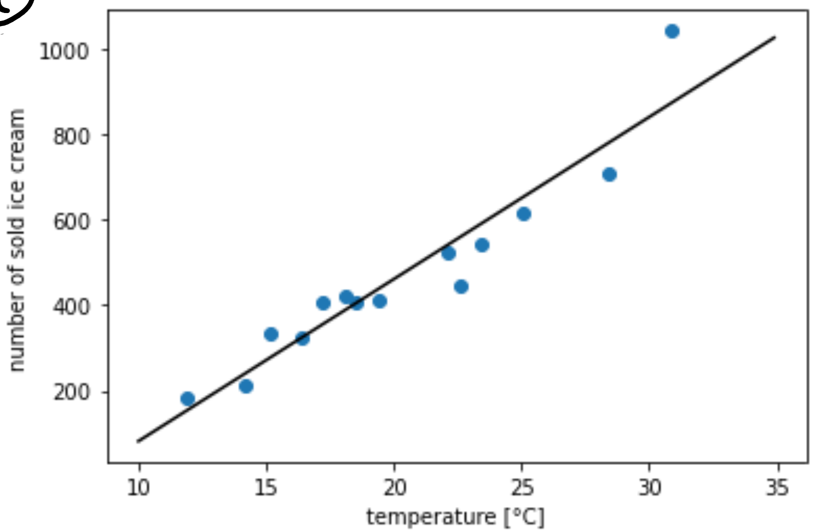
X (ZV)

Polynomialer Zusammenhang zwischen X und Y ⇒

$$\boxed{f(x) = b_0 + b_1 X + b_2 X^2 + \dots + b_k X^k}$$

... wir starten einfach (lin. Einfachregression) und schauen mal ...

1




Das Modell passt von seiner Struktur her gut zu den Daten. Einzelne Datenpunkte ändern das Ergebnis für $\hat{b}_0$, $\hat{b}_1$ kaum!

→ Polynom Grad 1

**geschätzte Parameter:**

$\hat{b}_0 = -340$

$\hat{b}_1 = 40{,}37$ (pos. lin. Zshg.)

**Bestimmtheitsmaß:**

$R^2 = 0{,}894$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.894
Model:                            OLS   Adj. R-squared:                  0.885
Method:                 Least Squares   F-statistic:                     101.4
Date:                Wed, 25 Oct 2023   Prob (F-statistic):           3.33e-07
Time:                        10:30:36   Log-Likelihood:                -80.096
No. Observations:                  14   AIC:                             164.2
Df Residuals:                      12   BIC:                             165.5
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       -340.9114     84.037     -4.057      0.002    -524.011    -157.811
x1            40.3691      4.010     10.067      0.000      31.632      49.106
==============================================================================
Omnibus:                        5.634   Durbin-Watson:                   1.301
Prob(Omnibus):                  0.060   Jarque-Bera (JB):                2.540
Skew:                           0.875   Prob(JB):                        0.281
Kurtosis:                       4.138   Cond. No.                         82.8
==============================================================================
```

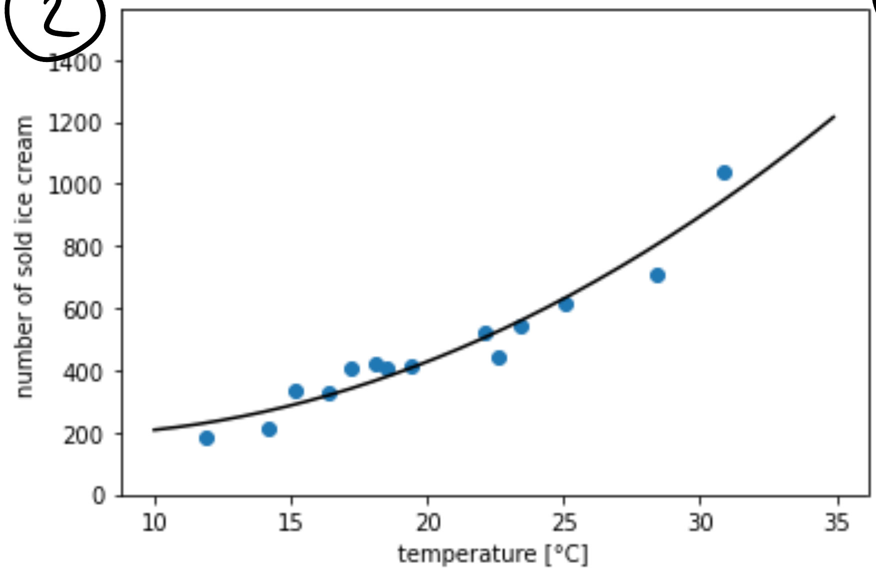
2

05

Polynom vom Grad 2 →

$$\boxed{f(x) = b_0 + b_1 X + b_2 X^2}$$

$$X = \begin{array}{c} \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \text{const} & X & X^2 \end{array} \end{array}$$




```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.942
Model:                            OLS   Adj. R-squared:                  0.931
Method:                 Least Squares   F-statistic:                     89.15
Date:                Wed, 25 Oct 2023   Prob (F-statistic):           1.60e-07
Time:                        10:34:47   Log-Likelihood:                -75.897
No. Observations:                  14   AIC:                             157.8
Df Residuals:                      11   BIC:                             159.7
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        349.1834    238.543      1.464      0.171    -175.847     874.214
x1           -28.0818     22.976     -1.222      0.247     -78.651      22.487
x2             1.5881      0.528      3.007      0.012       0.426       2.751
==============================================================================
Omnibus:                        1.343   Durbin-Watson:                   1.561
Prob(Omnibus):                  0.511   Jarque-Bera (JB):                0.806
Skew:                          -0.130   Prob(JB):                        0.668
Kurtosis:                       1.854   Cond. No.                     7.22e+03
==============================================================================
```

click to scroll output; double click to hide

← $R^2 = 0{,}942$ (gestiegen)

← **ABER**: p-Werte gestiegen ...

⇒ kleine Änderungen im Datensatz zeigen schon einen Einfluss ...

Polynom vom Grad 3 $\rightarrow$

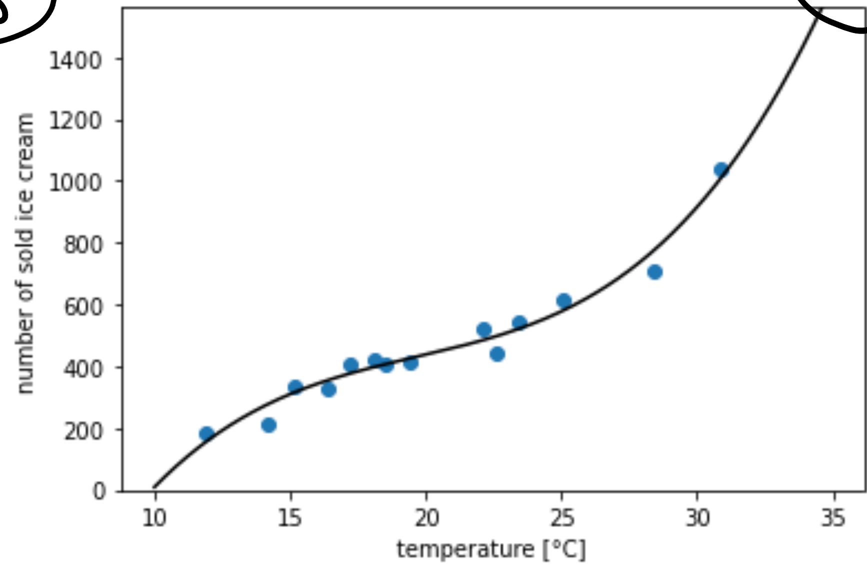
3

$$X = \begin{array}{cccc} \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 & x_m^3 \end{bmatrix} \end{array}$$

$$\begin{array}{cccc} \text{const} & \uparrow & \uparrow & \uparrow \\ & X & X^2 & X^3 \end{array}$$

$$\boxed{f(x) = b_0 + b_1 X + b_2 X^2 + b_3 X^3}$$

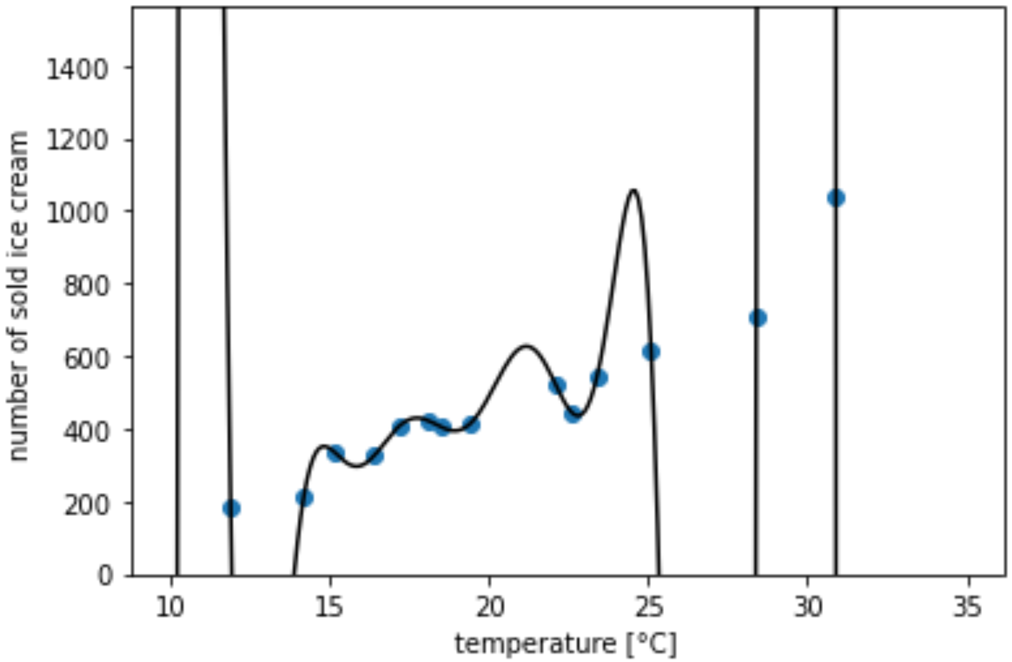
106




```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.978
Model:                            OLS   Adj. R-squared:                  0.972
Method:                 Least Squares   F-statistic:                     150.0
Date:                Wed, 25 Oct 2023   Prob (F-statistic):           1.30e-08
Time:                        10:36:06   Log-Likelihood:                -69.014
No. Observations:                  14   AIC:                             146.0
Df Residuals:                      10   BIC:                             148.6
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const      -1871.2232    563.948     -3.318      0.008   -3127.777    -614.670
x1           311.6791     84.354      3.695      0.004     123.726     499.632
x2           -14.9056      4.046     -3.684      0.004     -23.921      -5.890
x3             0.2546      0.062      4.091      0.002       0.116       0.393
==============================================================================
Omnibus:                        2.607   Durbin-Watson:                   3.015
Prob(Omnibus):                  0.272   Jarque-Bera (JB):                1.795
Skew:                          -0.710   Prob(JB):                        0.408
Kurtosis:                       1.971   Cond. No.                     6.96e+05
==============================================================================
```

$\leftarrow R^2 = 0{,}978$

$\leftarrow$ p-Werte klein




↑ deutliches overfitting !!

→ Polynom Grad 12

⚠ 14 Datenpunkte! 13 Parameter!

⇒ die Schätzung ist bereits numerisch instabil!

läßt man 1 Datenpunkt weg, dann ändert sich die Form des Polynoms komplett!

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.986
Model:                            OLS   Adj. R-squared:                  0.974
Method:                 Least Squares   click to scroll output; double click to hide   81.15
Date:                Wed, 25 Oct 2023   Prob (F-statistic):           4.10e-06
Time:                        10:37:00   Log-Likelihood:                -66.021
No. Observations:                  14   AIC:                             146.0
Df Residuals:                       7   BIC:                             150.5
Df Model:                           6
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       2.731e-08   3.91e-08      0.699      0.507     -6.5e-08     1.2e-07
x1         -1.666e-06   2.38e-06     -0.699      0.507     -7.3e-06    3.97e-06
x2          2.159e-06   3.08e-06      0.700      0.507    -5.14e-06    9.45e-06
x3          1.646e-05   2.35e-05      0.700      0.506    -3.91e-05    7.21e-05
x4             0.0001      0.000      0.701      0.506       -0.000       0.000
x5             0.0005      0.001      0.702      0.505       -0.001       0.002
x6             0.0014      0.002      0.706      0.503       -0.003       0.006
x7            -0.0004      0.001     -0.608      0.562       -0.002       0.001
x8          3.953e-05   7.04e-05      0.561      0.592       -0.000       0.000
x9         -2.327e-06    4.3e-06     -0.541      0.605    -1.25e-05    7.84e-06
x10         7.767e-08   1.45e-07      0.535      0.609    -2.65e-07    4.21e-07
x11        -1.383e-09   2.57e-09     -0.538      0.607    -7.46e-09    4.69e-09
x12          1.02e-11   1.87e-11      0.546      0.602     -3.4e-11    5.44e-11
==============================================================================
Omnibus:                        0.921   Durbin-Watson:                   3.362
Prob(Omnibus):                  0.631   Jarque-Bera (JB):                0.806
Skew:                          -0.366   Prob(JB):                        0.668
Kurtosis:                       2.081   Cond. No.                     8.42e+22
==============================================================================
```

⇒ p-Werte zeigen, daß das Modell deutlich zu komplex ist!

(obwohl $R^2$ top ist auf der Trainingsmenge!) ← Testmenge !!

⇝ Vermutung: Allgemein wäre es gut auch $X_2, \ldots, X_p$ berücksichtigen zu können – mit beliebigem polynomialen Grad …

**II. Polynom multivariat $X_1, \ldots, X_m$**

**Modell** $\boxed{Y = b_0 + b_1 X_1 + b_2 X_1^2 + \ldots + b_j X_i X_j + \ldots + \ldots + \varepsilon}$

⇒ schon bei 2 erklärenden Größen wird das i.A. sehr komplex..

$$\Rightarrow \quad (*) \quad \boxed{Y = b_0 + b_1 X_1 + b_2 X_1^2 + b_3 X_1 X_2 + b_4 X_2 + b_5 X_2^2 + \varepsilon}$$

Die Lösungsidee (VARIABLENTRAFO) bleibt aber die gleiche

setze

$$\boxed{\begin{array}{c} X_1 = \tilde{X}_1 \\ X_1^2 = X_2^2 \\ X_1 X_2 = \tilde{X}_3 \\ X_2 = \tilde{X}_4 \\ X_2^2 = \tilde{X}_5 \\ \vdots \end{array}} \quad \xRightarrow{\text{in } (*)} \quad \boxed{Y = b_0 + b_1 \tilde{X}_1 + b_2 \tilde{X}_2 + \ldots + b_5 \tilde{X}_5 + \varepsilon}$$

⚠ Vorsicht beim Befüllen der Datenmatrix!

Reihenfolge ist entscheidend!

$$X = \begin{array}{c} \begin{array}{cccc} \phantom{1} & x_1 & x_1^2 & \ldots \\ & \downarrow & \downarrow & \end{array} \\ \left[\begin{array}{cccc} 1 & x_{11} & \cdots & \cdots \\ 1 & x_{21} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cdots & \cdots & \cdots \end{array}\right] \\ \begin{array}{cccc} & \tilde{X}_1 & \tilde{X}_2 \ldots & \tilde{X}_m \end{array} \end{array}$$

Bsp. 1) Forts.

| i | Y | $X_1$ | $X_2$ |
|---|---|---|---|
| 1 | 2 | 2 | 2 |
| 2 | 4 | 6 | 2 |
| 3 | 6 | 6 | 4 |
| 4 | 4 | 4 | 4 |
| | A | B | C |

z.B.

$$\Rightarrow \boxed{Y = b_0 + b_1 X_1^2 + b_2 X_2^2 + \varepsilon}$$

$$X = \begin{array}{c} \begin{bmatrix} 1 & 4 & 4 \\ 1 & 36 & 4 \\ 1 & 36 & 16 \\ 1 & 16 & 16 \end{bmatrix} \\ \begin{array}{ccc} \overset{1}{\text{const.}} & X_1^2 & X_2^2 \end{array} \end{array} \qquad \vec{y} = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 4 \end{pmatrix}$$

$$\Rightarrow \vec{b} = (X^t \cdot X)^{-1} \cdot X^t \cdot \vec{y} = \begin{pmatrix} 1{,}019 \\ 0{,}073 \\ 0{,}130 \end{pmatrix}$$

$$\hat{y}_i = 1{,}019 + 0{,}073\, x_{i_1}^2 + 0{,}130\, x_{i_2}^2$$

$$\hat{\varepsilon}_i = y_i - \hat{y}_i \qquad (i = 1,2,3,4)$$

| i | $y_i$ | $\hat{y}_i$ | $\hat{\varepsilon}_i$ |
|---|---|---|---|
| 1 | 2 | 1,83 | 0,17 |
| 2 | 4 | 4,17 | -0,17 |
| 3 | 6 | 5,73 | 0,27 |
| 4 | 4 | 4,27 | -0,27 |

größer als vorher

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n-(k+1)} \sum_i \hat{\varepsilon}_i^2 = \frac{1}{4-3} \sum_i \hat{\varepsilon}_i^2 = 0{,}20 \Rightarrow \hat{\sigma} = 0{,}45$$

**Bemerkung**: mit $\boxed{\begin{array}{l} H_0 : b_2 = 0 \\ H_1 : b_2 \neq 0 \end{array}}$ ← Test ob $X_2^2$ im Modell Sinn macht ...

<u>Bemerkung</u>: die Lösungsidee läßt sich auf beliebige Modelle erweitern, die eine lineare "Grundstruktur" haben ...

III. Fourier-Synthese

Modell

Bsp.:

$$Y = b_0 + b_1 \cos(X) + b_2 \cos(2X) + b_3 \sin(X) + b_4 \sin(2X) + \varepsilon \quad (*)$$

... und beliebige weitere Terme ... ← (linear)

TRAFO

$\cos(X) = \tilde{X}_1$

$\cos(2X) = \tilde{X}_2$

$\sin(X) = \tilde{X}_3$

$\sin(2X) = \tilde{X}_4$

$\xRightarrow{\text{in } (*)}$

$$Y = b_0 + b_1 \tilde{X}_1 + b_2 \tilde{X}_2 + b_3 \tilde{X}_3 + b_4 \tilde{X}_4 + \varepsilon$$

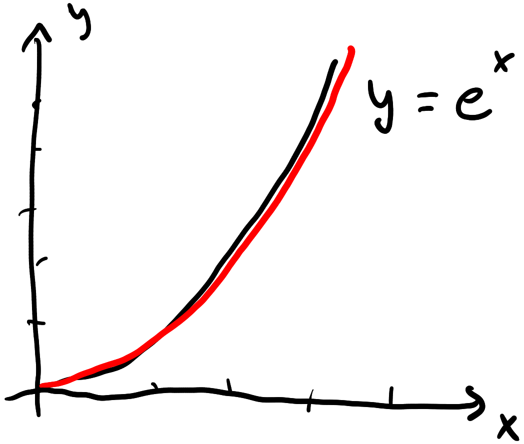
<u>dann</u>:

$$X = \begin{array}{c} \left[\begin{matrix} 1 & \cos(x_1) & \cos(2x_1) & \cdots & \sin(2x_1) \\ 1 & \cos(x_2) & \cos(2x_2) & & \sin(2x_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos(x_n) & \cos(2x_n) & & \sin(2x_n) \end{matrix}\right] \\ \begin{matrix} \uparrow & \uparrow & & \uparrow \\ \cos(X) & \cos(2X) & \cdots & \sin(2X) \end{matrix} \end{array}$$

⇒ ist besonders dann geeignet, wenn Y periodisch von X abhängt!

→ Signalverarbeitung, Sound, etc.

Mit den bisherigen Überlegungen läßt sich schon viel machen. Was aber, wenn




← ganz anderer funktionaler Zusammenhang z.B. exponentiell...

IV. Exponentielle Regression

MODELL

$$\boxed{Y = b_0 \cdot e^{b_1 X} + \varepsilon}$$

exponentielles Modell

ZV $X, Y$ mit Realisierungen $(x_i, y_i)$ $(i = 1, \ldots, n)$

⇒ empirische Beziehung

$$\boxed{y_i = b_0 \cdot e^{b_1 x_i} + \varepsilon_i}$$

→ jetzt ist eine gute Idee gefragt ☺

**IDEE** $\Rightarrow$ betrachte die empirische Beziehung:

$$\boxed{y_i = b_0 \cdot e^{b_1 x_i}}$$

$\leftarrow$ e-Funktion "stört"

wende die Umkehrfunktion $\ln(\dots)$ auf beiden Seiten der Gleichung an

$$\boxed{\ln(e^x) = x = e^{\ln(x)}}$$

$$\Rightarrow \quad y_i = b_0 \cdot e^{b_1 x_i} \qquad \circ \Big| \ln(\dots)$$

$$(*) \quad \boxed{\ln(y_i) = \ln(b_0) + b_1 x_i}$$

Setze:

$$\boxed{\begin{aligned} \ln(y_i) &= \tilde{y}_i \\ \ln(b_0) &= \tilde{b}_0 \end{aligned}}$$

TRAFO

in $(*)$ $\Rightarrow$

$$\boxed{\tilde{y}_i = \tilde{b}_0 + b_1 x_i}$$

$\leftarrow$ lineare Einfachregression mit bekannter Lösung

... um auf das ursprüngliche Modell zu kommen muß man am Ende zurück transf. !

Bsp. 1) Forts.

| i | Y | X |
|---|---|---|
| 1 | 2 | 2 |
| 2 | 4 | 6 |
| 3 | 6 | 6 |
| 4 | 4 | 4 |
| | A | B |

← Vermutung: $\boxed{Y = b_0 \cdot e^{b_1 X} + \varepsilon}$

| $x_i$ | $y_i$ | $\ln(y_i) = \tilde{y}_i$ |
|---|---|---|
| 2 | 2 | 0,69 |
| 6 | 4 | 1,39 |
| 6 | 6 | 1,79 |
| 4 | 4 | 1,39 |

$\bar{x} = 4{,}5 \quad s_X^2 = 2{,}75$

$\bar{\tilde{y}} = 1{,}315$

$\mathrm{cov}(x, \tilde{y}) = 0{,}5875$

Lösung:

$$\left.\begin{aligned} \hat{\tilde{b}}_0 &= \bar{\tilde{y}} - \hat{b}_1 \bar{x} \\ \hat{b}_1 &= \frac{\mathrm{cov}(x, \tilde{y})}{s_x^2} \end{aligned}\right\} \Rightarrow \begin{aligned} \hat{b}_1 &= \frac{0{,}5875}{2{,}75} = 0{,}21 \\ \hat{\tilde{b}}_0 &= 1{,}315 - 0{,}21 \cdot 4{,}5 = 0{,}37 \end{aligned}$$

! $\hat{\tilde{b}}_0$ muß rücktransf. werden: $\hat{b}_0 = e^{\hat{\tilde{b}}_0} = 1{,}45 \Rightarrow \boxed{\hat{y}_i = 1{,}45 \cdot e^{0{,}21 x_i}}$

Prognose

**Auswertung:**

| $y_i$ | $\hat{y}_i$ | $\hat{\varepsilon}_i$ |
|---|---|---|
| 2 | 2,20 | -0,2 |
| 4 | 5,11 | -1,11 |
| 6 | 5,11 | 0,89 |
| 4 | 3,36 | 0,64 |

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{4} \hat{\varepsilon}_i^2 = 1{,}24$$

$$\Rightarrow \hat{\sigma} = 1{,}11$$

(Polynommodell $\hat{\sigma} = 1{,}41$ $\Rightarrow$ leicht besser)

**Bemerkung:**

Das Modell läßt sich auch multivariat formulieren.

Dann
$$\boxed{Y = b_0 \cdot e^{b_1 x_1 + b_2 x_2 + \ldots + b_m x_m} + \varepsilon}$$

logarithmiere $\Rightarrow$ $$\boxed{\ln(Y) = \ln(b_0) + b_1 x_1 + \ldots + b_m x_m + \varepsilon}$$

$\Rightarrow$ Lösung wie beim multivariaten linearen Modell!

# ZUSAMMENFASSUNG:



| Modell | $f(X)$ | TRAFO | LÖSUNG | METRIK |
|---|---|---|---|---|
| lin. Einfach-regression | $b_0+b_1X$ | keine | $\hat{b}_0=\bar{y}-\hat{b}_1\bar{x}$ <br> $\hat{b}_1=\frac{cov(x,y)}{s_x^2}$ (*) | $\hat{\sigma}^2=\frac{1}{n-2}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$ <br> $R^2=\frac{SQE}{SQT}$ |
| multiple Regression | $b_0+b_1X_1+\ldots+b_kX_k$ | keine | $\vec{b}=(X^t\cdot X)^{-1}\cdot X^t\cdot\vec{y}$ | $\hat{\sigma}^2=\frac{1}{n-(k+1)}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$ <br> korrigierten $R^2$ |
| Polynom in $X$ | $b_0+b_1X+b_2X^2+\ldots b_kX^k$ | $X=\tilde{X}_1$ <br> $X^2=\tilde{X}_2$ <br> $\vdots$ <br> $X^k=\tilde{X}_k$ | $\vec{b}=(\tilde{X}^t\tilde{X})^{-1}\cdot\tilde{X}^t\cdot\vec{y}$ | — " — |
| Polynom in $X_1,\ldots$ (multip.) (max. deg≈4 sinnvoll) | Bsp. <br> $b_0+b_1X_1+b_2X_2+b_3X_1X_2+b_4X_1^2+b_5X_2^2$ | $X_1=\tilde{X}_1$ <br> $X_2=\tilde{X}_2$ <br> $X_1X_2=\tilde{X}_3$ <br> $\vdots$ | — " — | — " — |
| Exponentialmodell | $b_0\cdot e^{b_1X}$ | $\ln(Y)=\tilde{Y}$ <br> $\ln(b_0)=\tilde{b}_0$ | $\hat{\tilde{b}}_0=\bar{\tilde{Y}}-\hat{b}_1\bar{x}$ sh. → (*) <br> $\hat{b}_1=\frac{cov(x,\tilde{Y})}{s_x^2}$ | $\hat{\sigma}^2=\frac{1}{n-2}\sum_i\hat{\varepsilon}_i^2$ <br> $R^2=\frac{SQE}{SQT}$ |