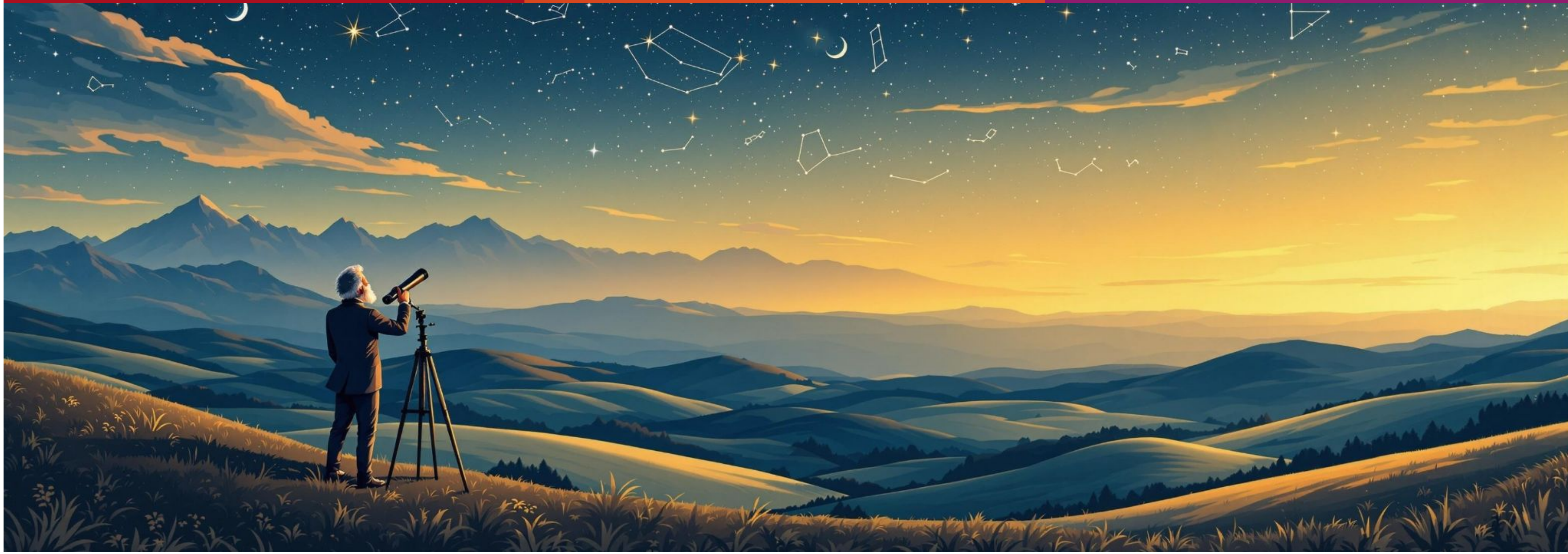


Abbildung: Erstellt mit GenSpark.ai (Mai 2025)

Least-Squares-Methode – Theorie und Anwendung

Von verrauschten Messdaten zum linearen Modell

Bachelor Studiengänge im 2. Semester

Überblick

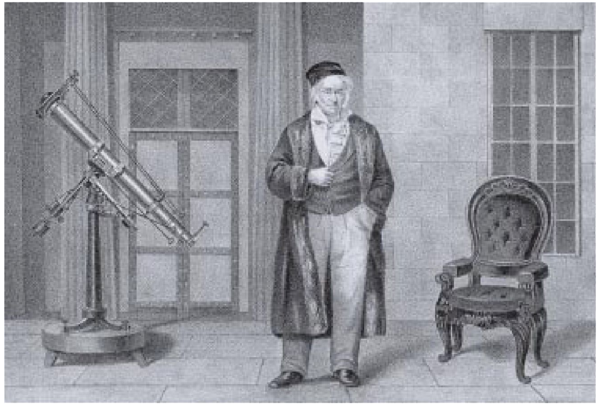
- Historischer Anwendungsfall
- Die Methode der kleinsten Quadrate am einfachen Beispiel
  - Modelleigenschaften
  - Fehlermaße
  - Berechnung der Ausgleichsgeraden
- Anwendungen von damals bis heute
- Mini-Quiz

Einsatz der Least-Squares-Methode damals & heute

- Im Jahr 1801 entdeckte der Astronom Giuseppe Piazzi den Himmelskörper Ceres
  - Ceres war nur wenige Tage sichtbar
  - Nur wenige Positionsdaten erfasst

- Gauss entwickelte die Methode der kleinsten Quadrate, mit der er die Bahn des Himmelskörpers bestimmen konnte

- Ende 1801 wurde Ceres tatsächlich genau an der von Gauss berechneten Stelle entdeckt.



Carl Friedrich Gauß auf der Terrasse „seiner" Sternwarte
Abbildung: Uni Göttingen, Uni|in|form, Ausgabe Nr.4/2004

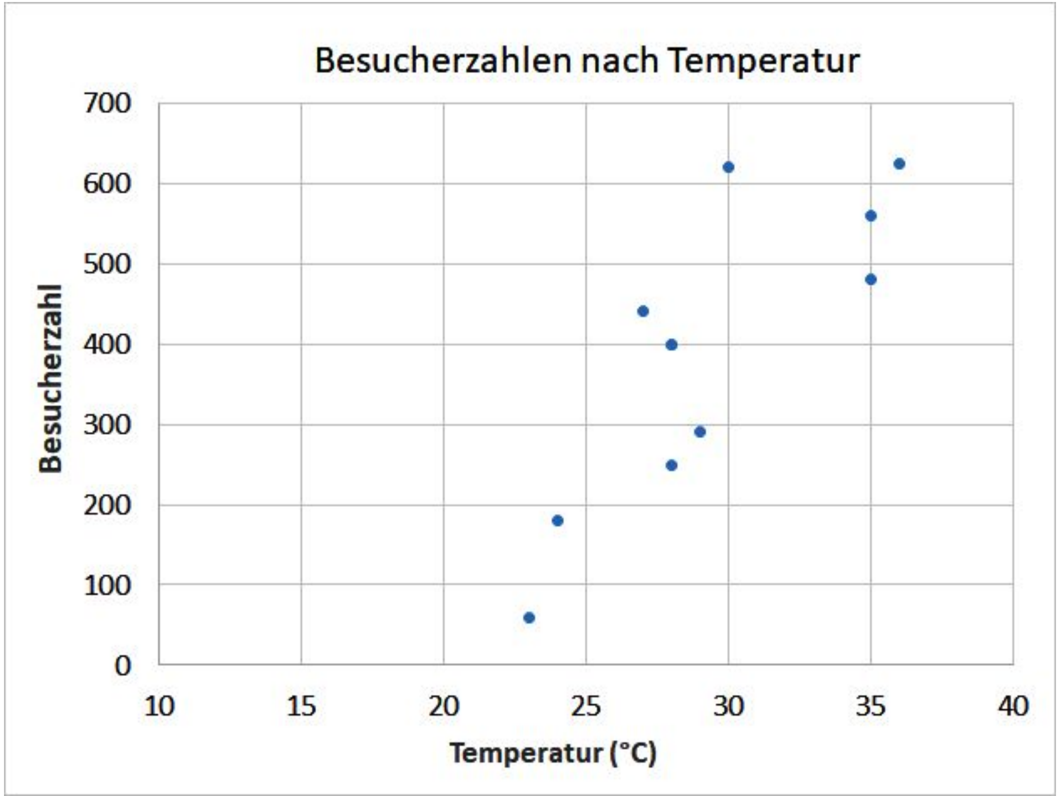
Die Methode der kleinsten Quadrate am einfachen Beispiel

- Besucherzahlen der letzten Tage bekannt
- Wetterdaten der letzten Tage bekannt
- Für morgen sind 33°C angesagt
- Wie viele Besucher sind morgen zu erwarten?

Temperatur in °C (x)	28	23	24	36	35	29	30	27	28	35	33
Gästezahl (y)	400	60	180	625	560	290	620	440	250	480	?

Das Verfahren am einfachen Beispiel

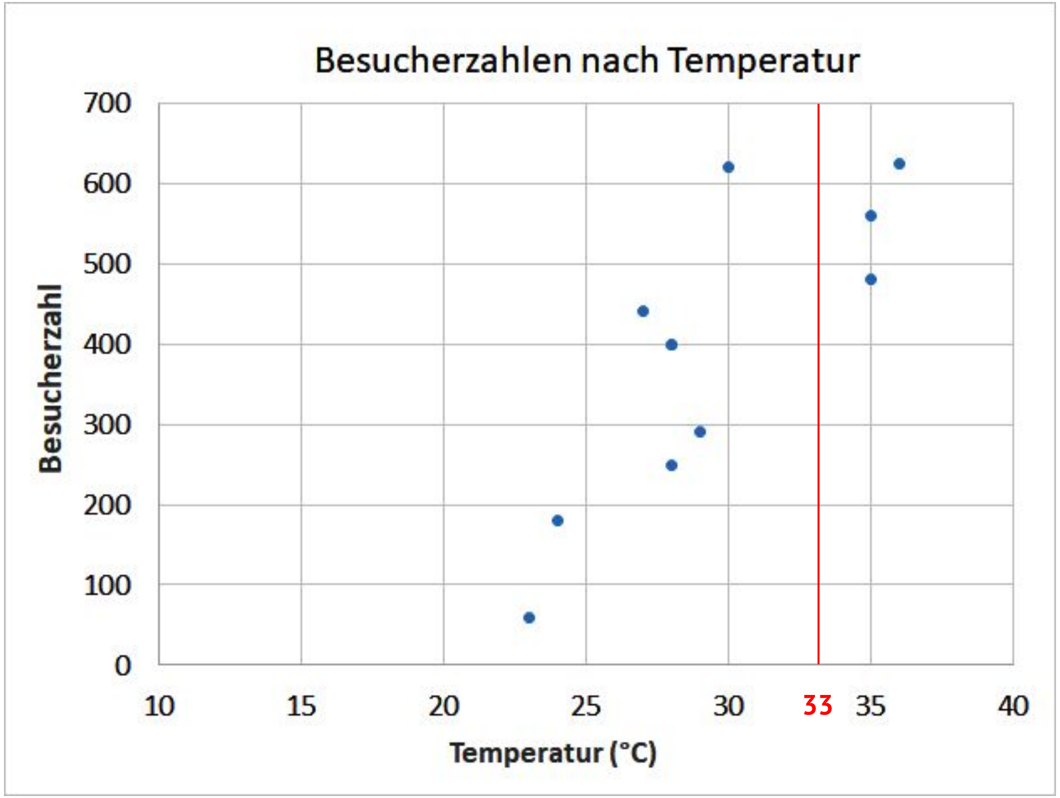
- Besucherzahlen der letzten Tage bekannt
- Wetterdaten der letzten Tage bekannt



Das Verfahren am einfachen Beispiel - Modelleigenschaften

- Besucherzahlen der letzten Tage bekannt
- Wetterdaten der letzten Tage bekannt

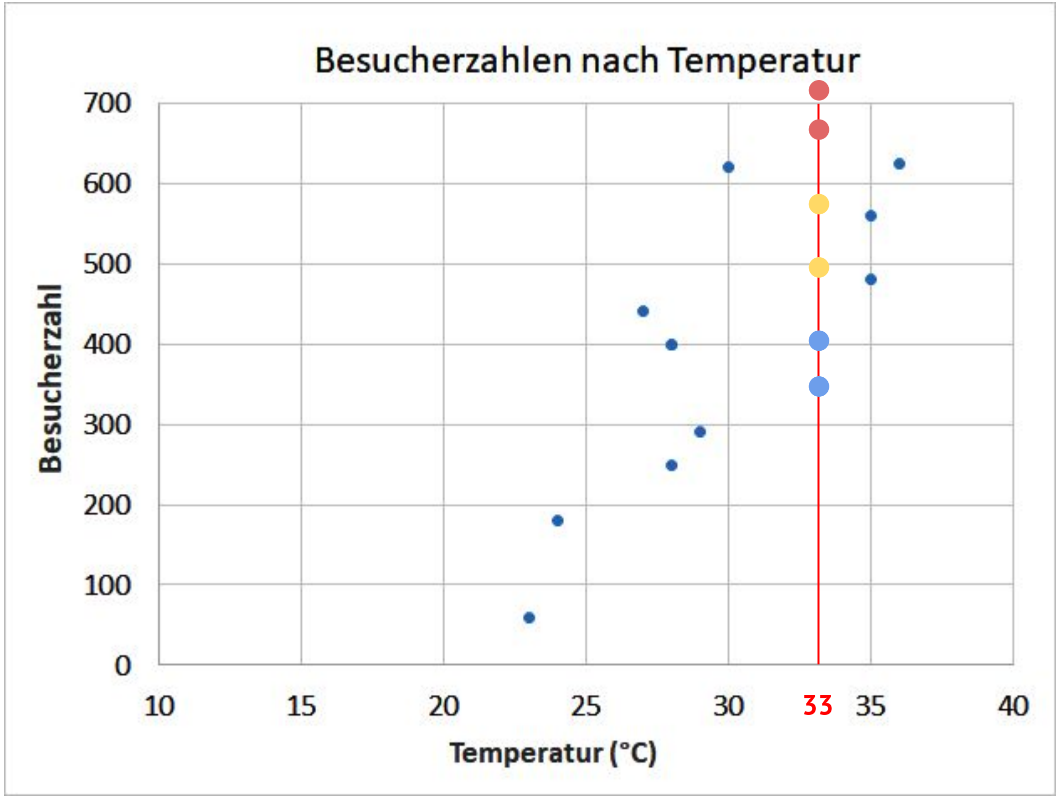
- Für morgen sind 33°C angesagt
- Wie viele Besucher sind morgen zu erwarten?



Das Verfahren am einfachen Beispiel - Modelleigenschaften

- Besucherzahlen der letzten Tage bekannt
- Wetterdaten der letzten Tage bekannt

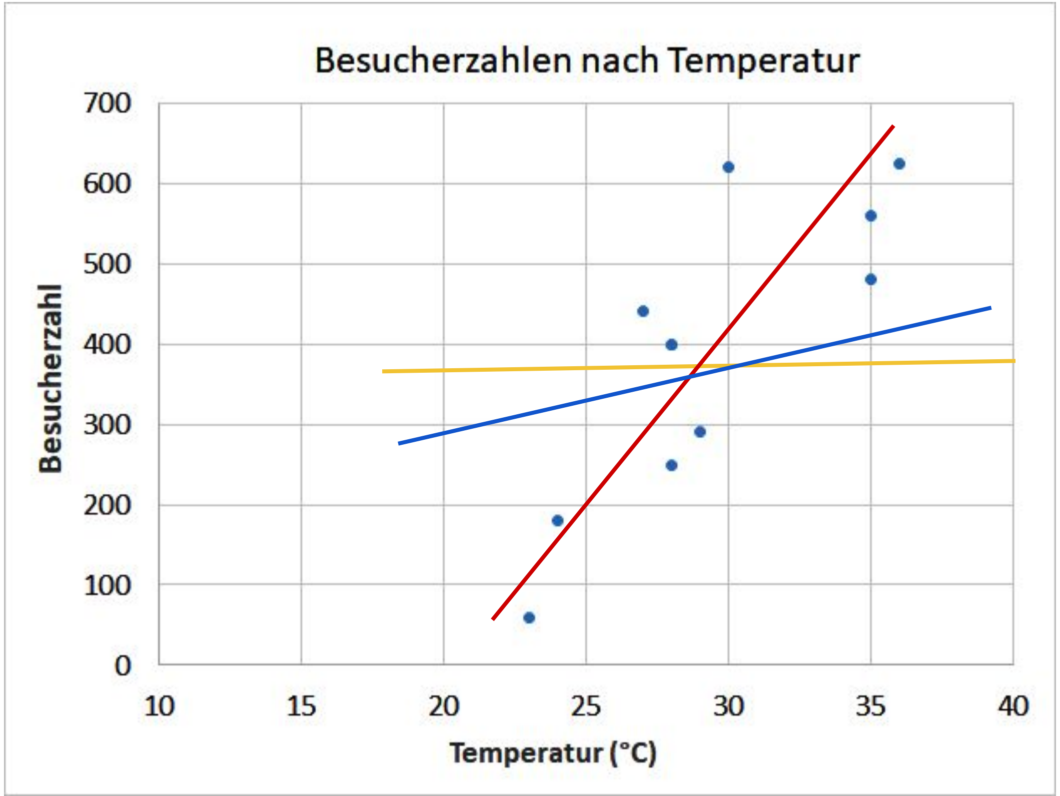
- Für morgen sind 33°C angesagt
- Wie viele Besucher sind morgen zu erwarten?



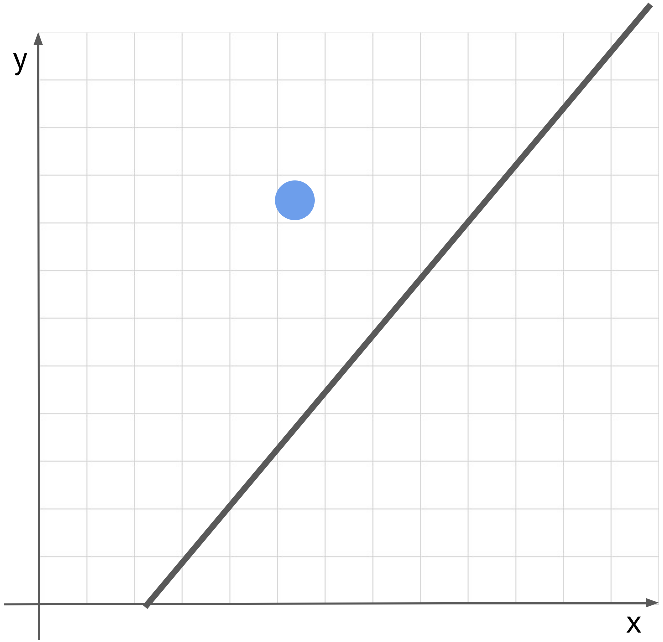
Das Verfahren am einfachen Beispiel - Modelleigenschaften

- Besucherzahlen der letzten Tage bekannt
- Wetterdaten der letzten Tage bekannt

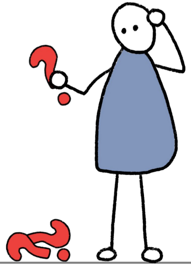
- Für morgen sind 33°C angesagt
- Wie viele Besucher sind morgen zu erwarten?



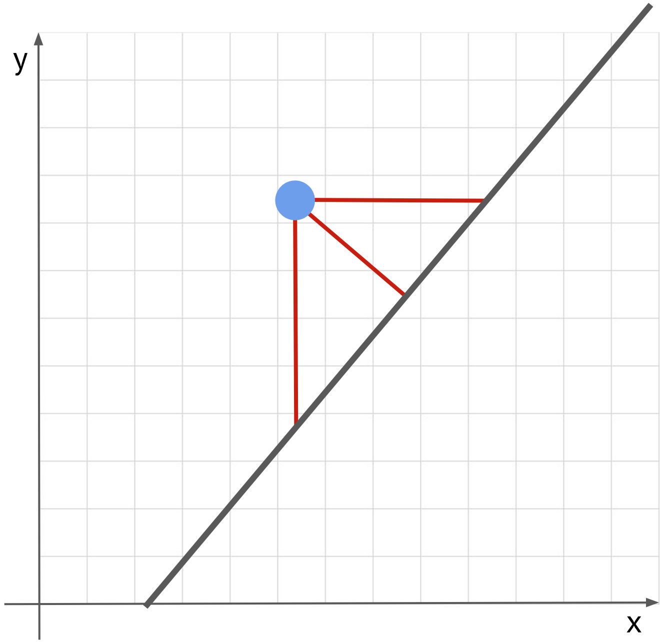
Das Verfahren am einfachen Beispiel — Fehlermaß



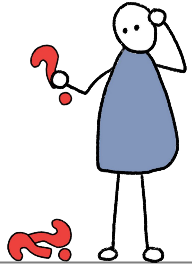
- Was ist der Abstand, der möglichst klein sein soll?



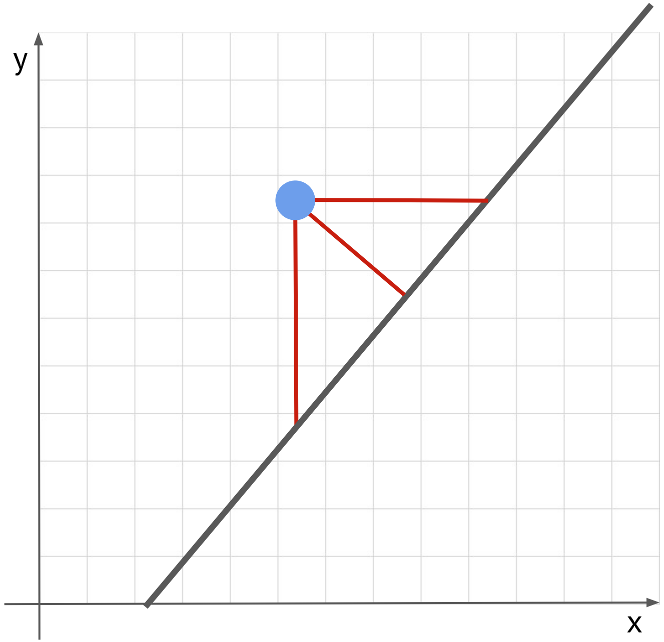
Das Verfahren am einfachen Beispiel — Fehlermaß



- Was ist der Abstand, der möglichst klein sein soll?

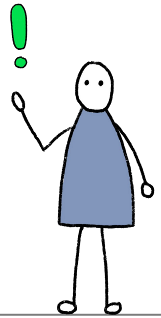


Das Verfahren am einfachen Beispiel — Fehlermaß

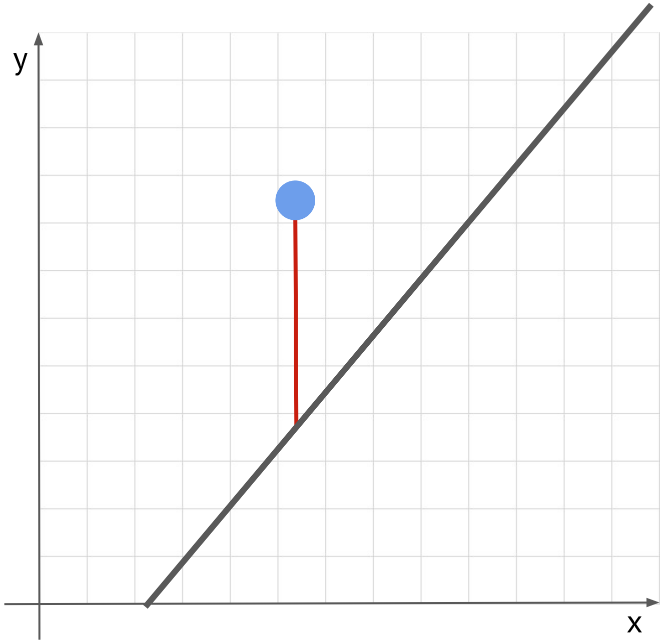


- Was ist der Abstand, der möglichst klein sein soll?

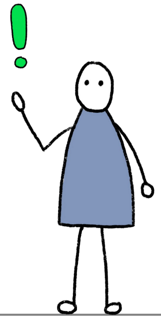
- Problemstellung hilft:
  - Modell: f(x) = y -- nicht in x und y gleichberechtigt
  - Temperatur ist bekannt (x, unabhängige Variable)
  - Besucherzahl ist gesucht (y, abhängige Variable)



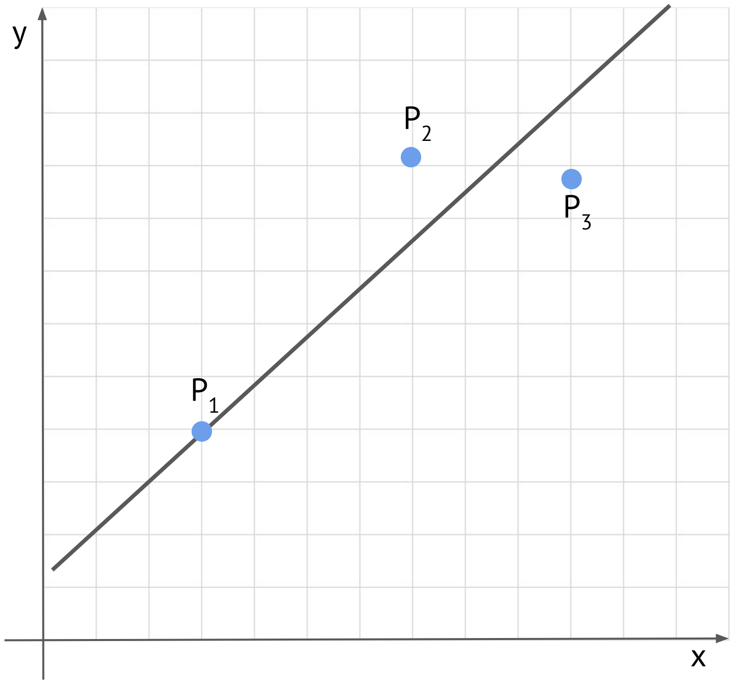
Das Verfahren am einfachen Beispiel — Fehlermaß



- Was ist der Abstand, der möglichst klein sein soll?
- Problemstellung hilft:
  - Modell: f(x) = y -- nicht in x und y gleichberechtigt
  - Temperatur ist bekannt (x, unabhängige Variable)
  - Besucherzahl ist gesucht (y, abhängige Variable)
- Fehler bewertet die Vorhersage des Modells
- **Residuum**: Differenz zwischen den tatsächlichen beobachteten Werten und den vorhergesagten Werten einer statistischen Analyse

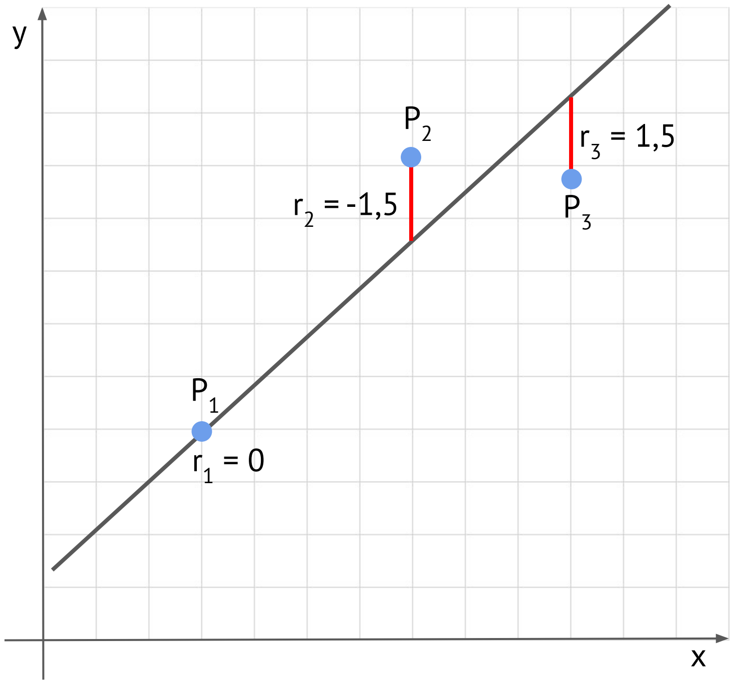


Das Verfahren am einfachen Beispiel — Fehlermaß

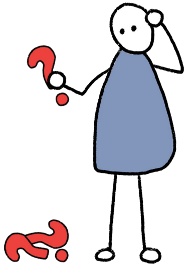


Gesamtfehler: $\sum_{i=1}^{n} r_i$, n = Anzahl der Punkte

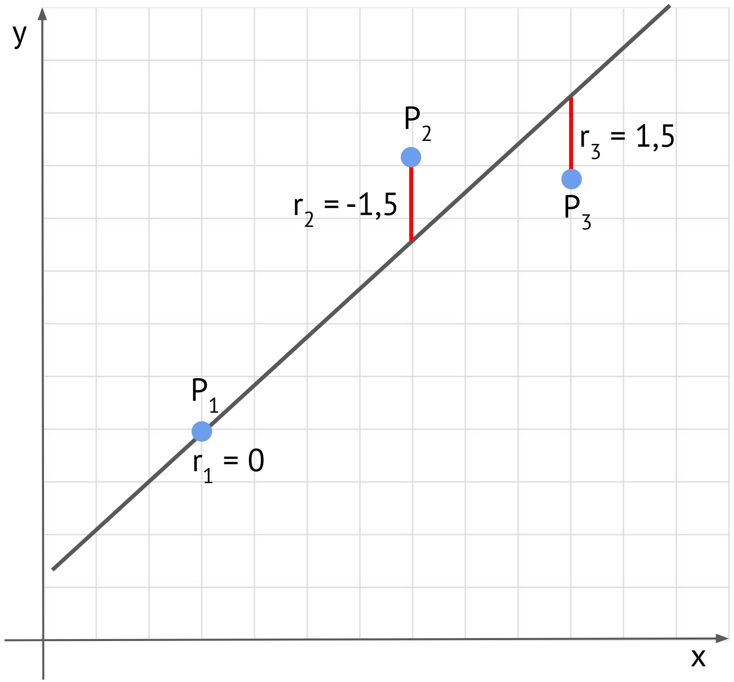
Das Verfahren am einfachen Beispiel — Fehlermaß



Gesamtfehler: $\sum_{i=1}^{n} r_i$, n = Anzahl der Punkte

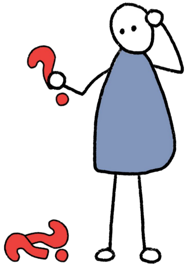


Das Verfahren am einfachen Beispiel — Fehlermaß

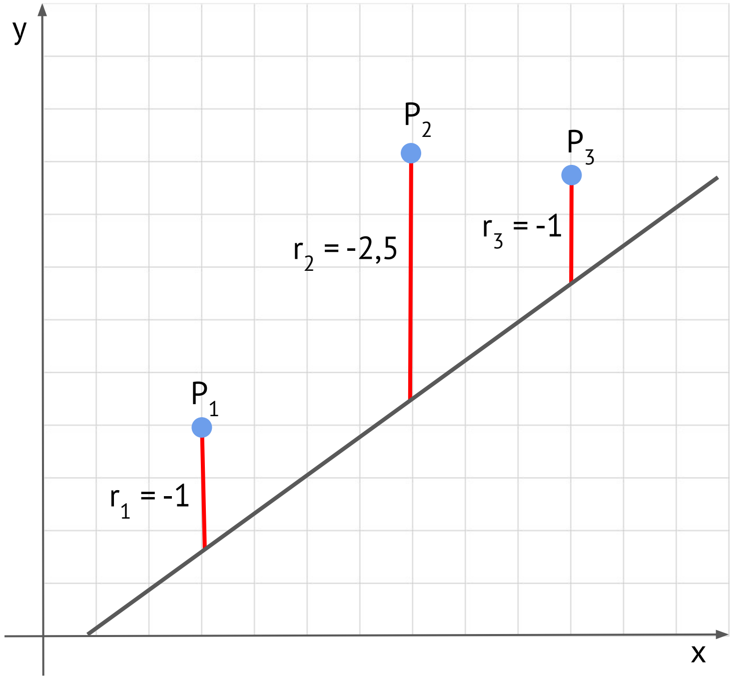


Gesamtfehler: $\sum_{i=1}^{n} r_i$, n = Anzahl der Punkte

SOE = $r_1 + r_2 + r_3 = 0 + (-1{,}5) + (1{,}5) = 0$

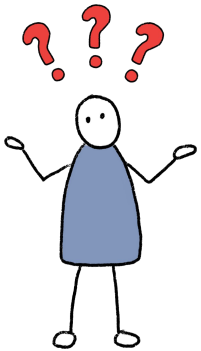


Das Verfahren am einfachen Beispiel — Fehlermaß



Gesamtfehler: $\sum_{i=1}^{n} r_i$, n = Anzahl der Punkte

Gesamtfehler = $r_1 + r_2 + r_3 = (-1) + (-2{,}5) + (-1) = -4{,}5$



Das Verfahren am einfachen Beispiel — Fehlermaß

Sum of Errors (SOE)

- Vorteile:
  - Intuitiv
- Nachteile:
  - Positive und negative Fehler heben sich auf

$$SOE = \sum_{i=1}^{n} r_i$$

Sum of Absolute Errors (SAE) / **Mean Absolute Error** (MAE)

- Vorteile:
  - Interpretierbar
- Nachteile:
  - Nicht differenzierbar bei Null

$$SAE = \sum_{i=1}^{n} |r_i|$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |r_i|$$

Das Verfahren am einfachen Beispiel — Fehlermaß

Sum of Squared Errors (SSE) / **Mean Squared Error** (MSE)

- Vorteile:
  - Differenzierbar
  - Bestraft große Fehler stark
- Nachteile:
  - Empfindlich gegenüber Ausreißern

$$SSE = \sum_{i=1}^{n} r_i^2$$

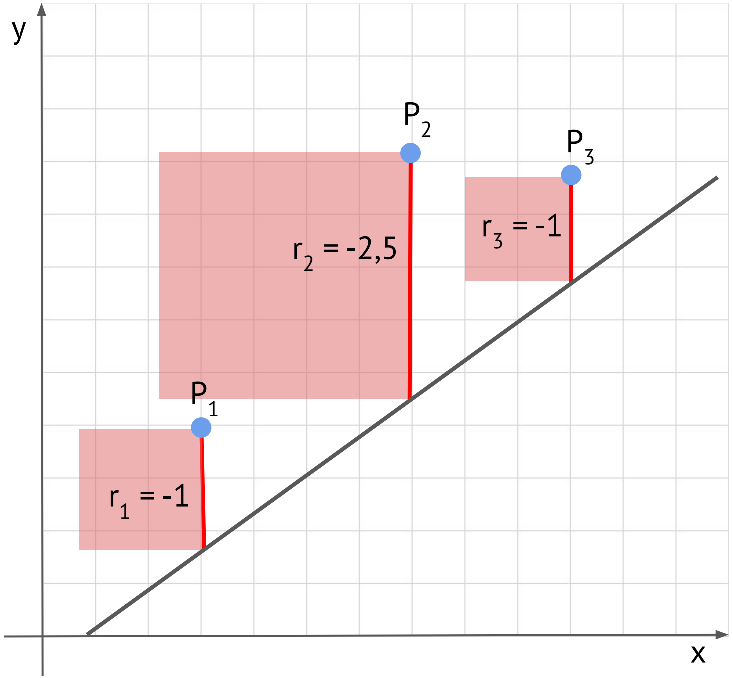
$$MSE = \frac{1}{n}\sum_{i=1}^{n} r_i^2$$

Root Mean Squared Error (RMSE)

- Vorteil:
  - Differenzierbar
  - Bestraft große Fehler stark
  - Fehlermaß interpretierbar
- Nachteil:
  - Empfindlich gegenüber Ausreißern

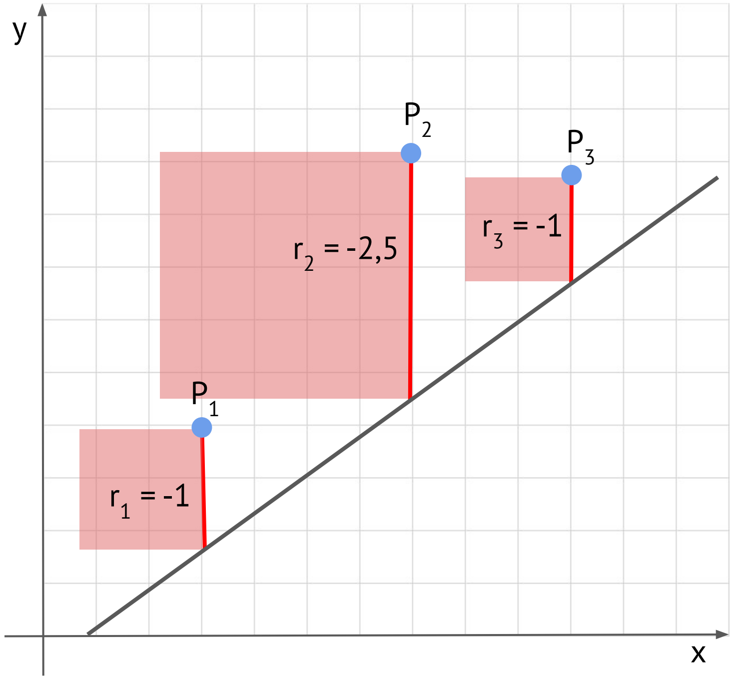
$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} r_i^2}$$

Das Verfahren am einfachen Beispiel — Fehlermaß



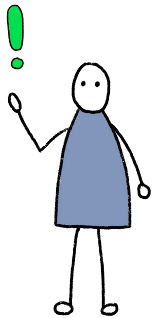
Gesamtfehler: $SSE = \sum_{i=1}^{n} r_i^2$, n = Anzahl der Punkte

Das Verfahren am einfachen Beispiel — Fehlermaß



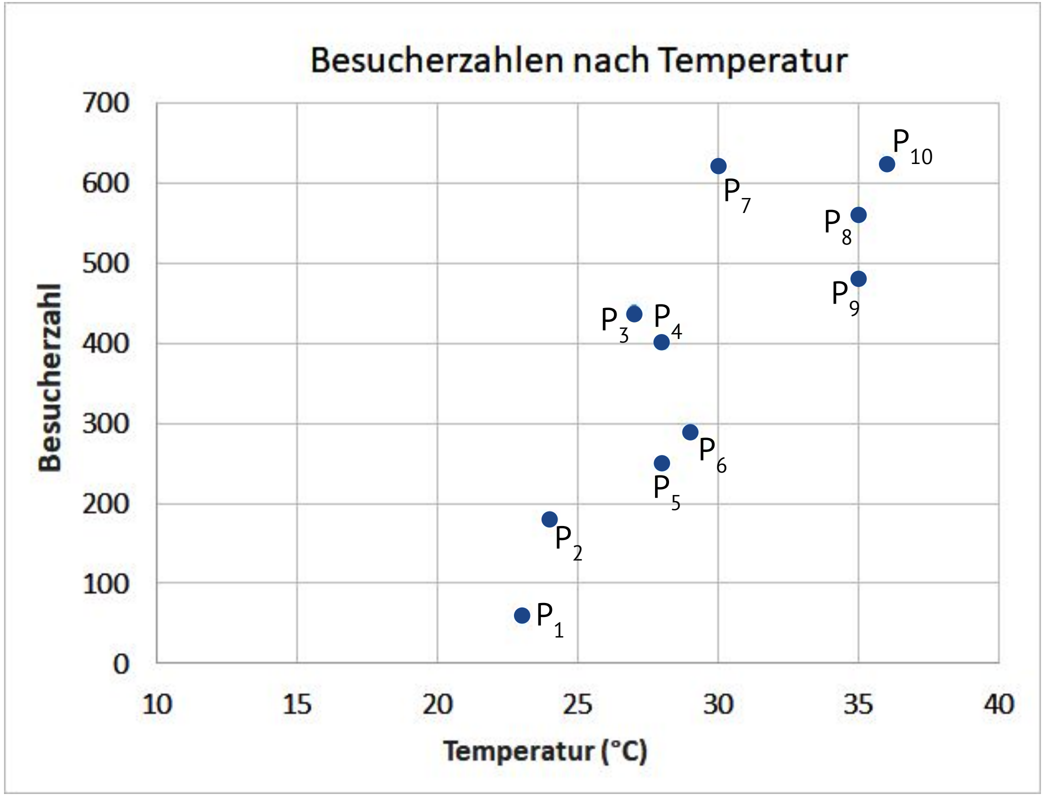
Gesamtfehler: $SSE = \sum_{i=1}^{n} r_i^2$, n = Anzahl der Punkte

SSE = $r_1^2 + r_2^2 + r_3^2 = (-1)^2 + (-2{,}5)^2 + (-1)^2 = 8{,}25$



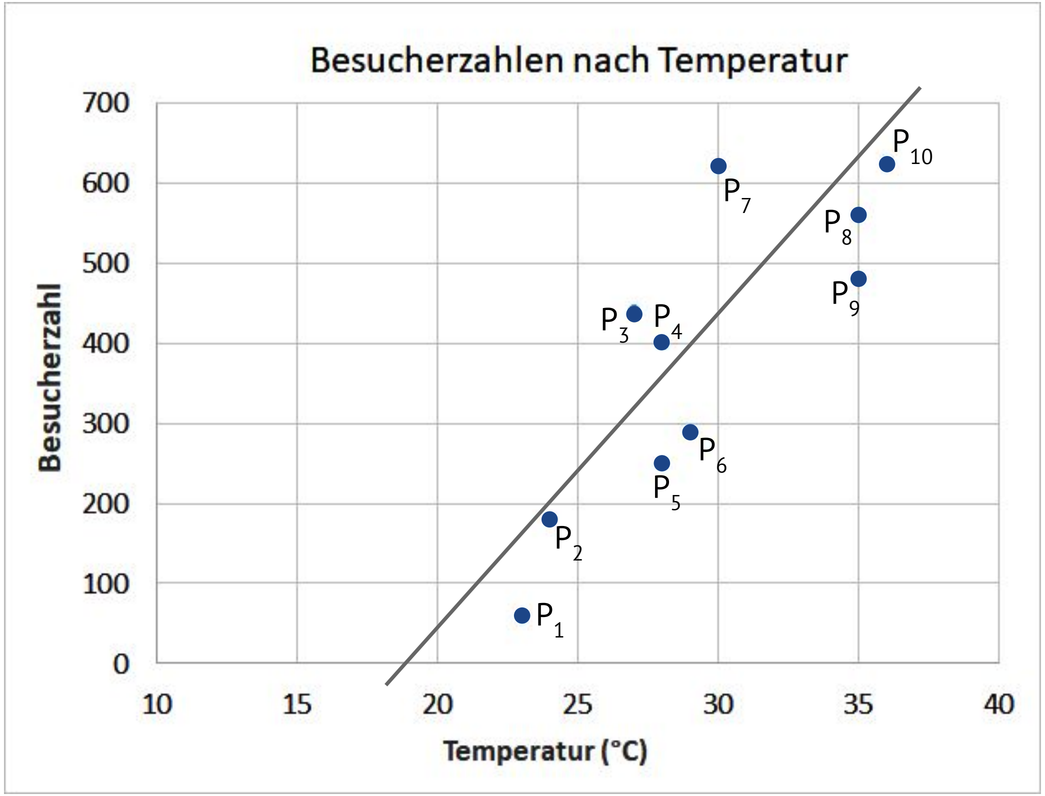
Das Verfahren am einfachen Beispiel

- 2-dimensionale Daten liegen vor
  - (x_i, y_i), *für* $i=1,\ldots,n$, *mit* $n=10$
- Annahme: Linearer Zusammenhang



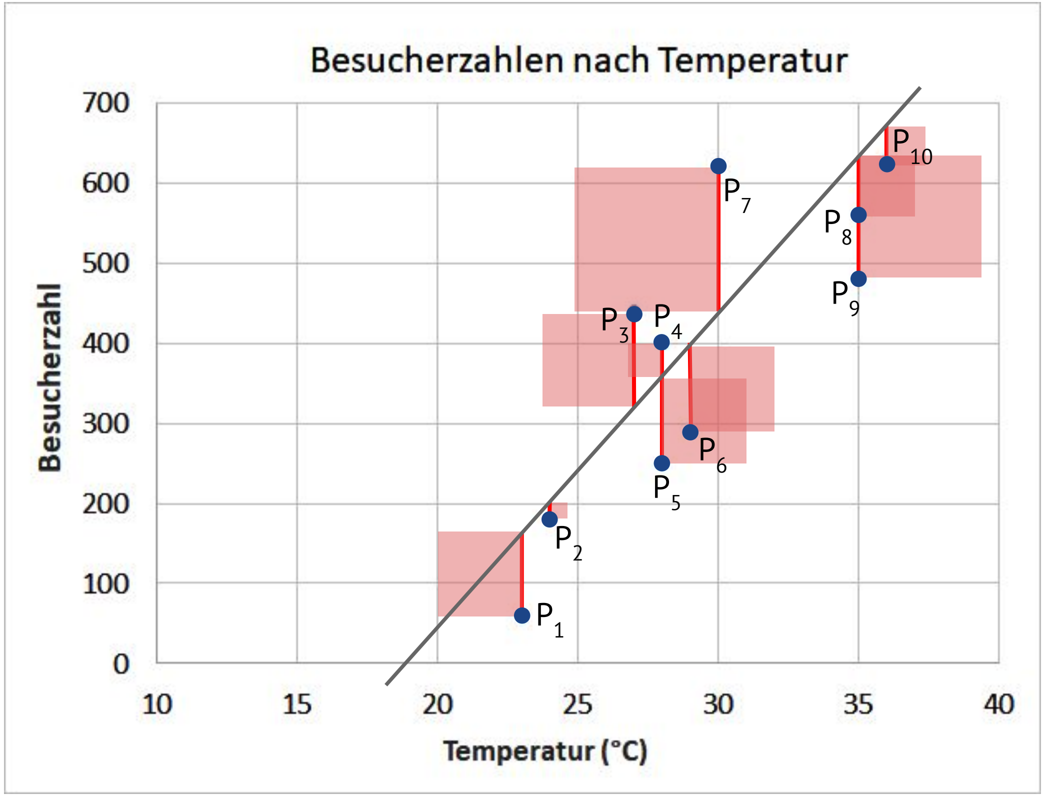
Das Verfahren am einfachen Beispiel

- 2-dimensionale Daten liegen vor
  - (x_i, y_i), *für* $i=1,\dots,n$, *mit* $n=10$
- Annahme: Linearer Zusammenhang
- Modell: $f(x) : \hat{y} = mx + b$



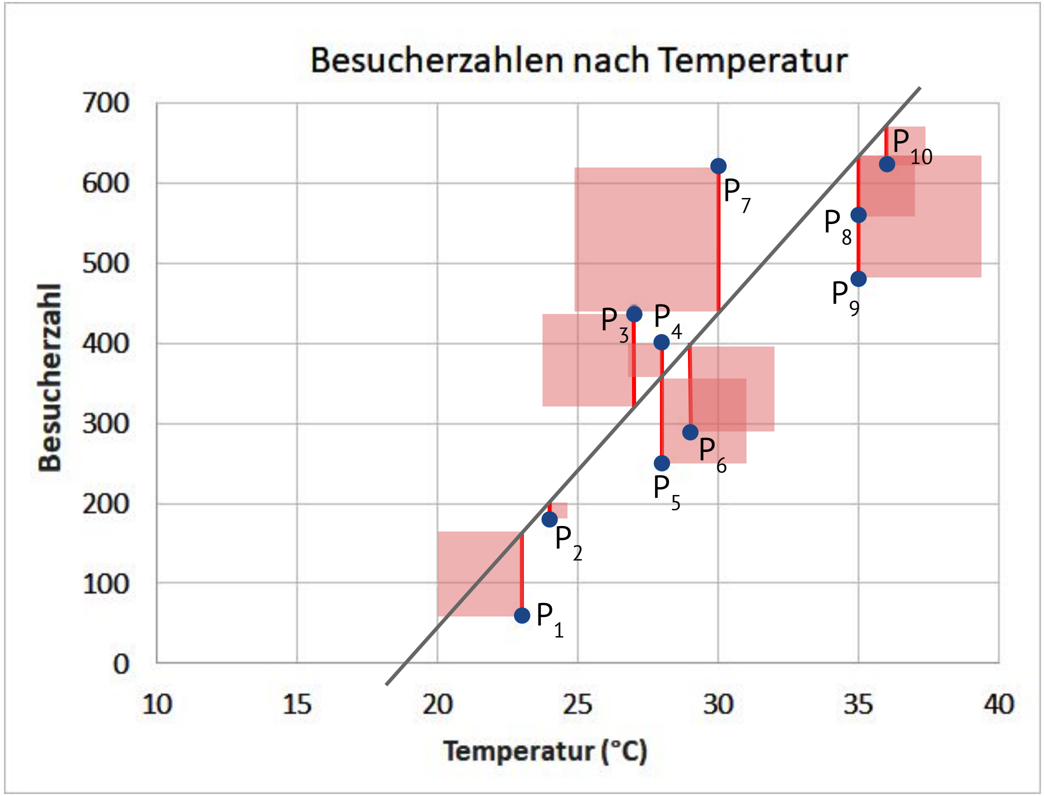
Das Verfahren am einfachen Beispiel

- 2-dimensionale Daten liegen vor
  - (x_i, y_i), *für* $i=1,\ldots,n$, *mit* $n=9$
- Annahme: Linearer Zusammenhang
- Modell: $f(x): \hat{y} = mx + b$
- "Beste" Gerade: $\min \sum_{i=1}^{n} (r_i)^2$



Das Verfahren am einfachen Beispiel

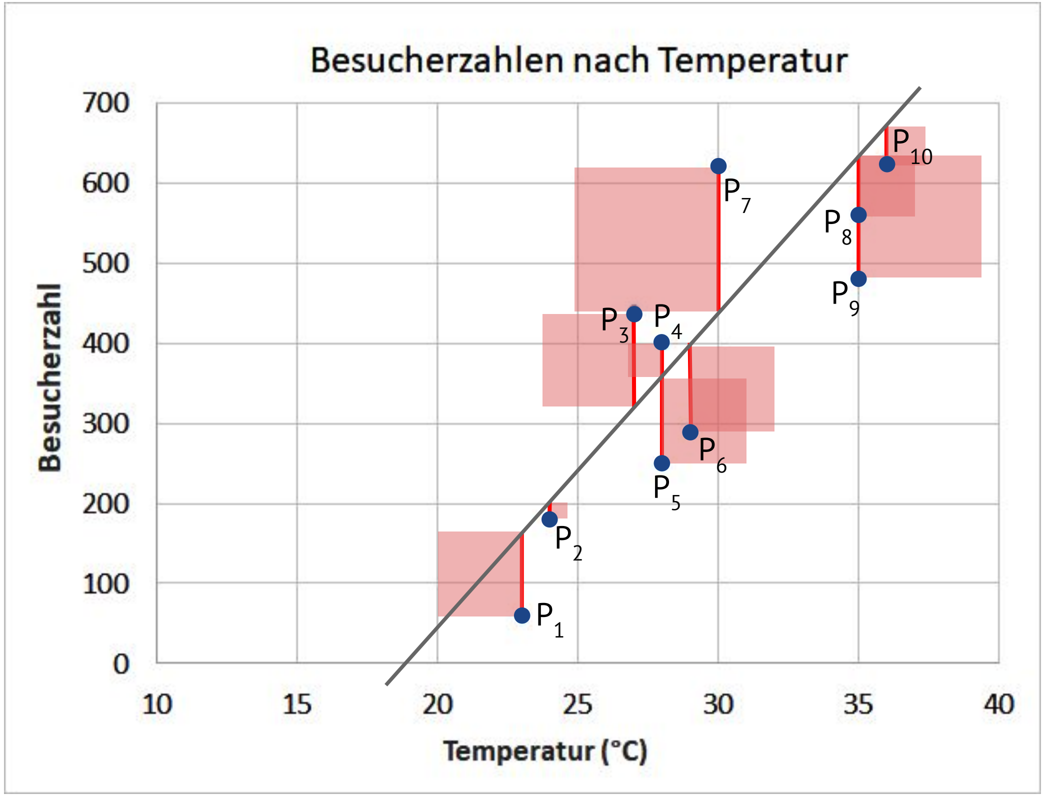
- 2-dimensionale Daten liegen vor
  - (x_i, y_i), für $i=1,\ldots,n$, mit $n=9$
- Annahme: Linearer Zusammenhang
- Modell: $f(x): \hat{y} = mx + b$
- "Beste" Gerade: $\min \sum_{i=1}^{n} (r_i)^2$
- Residuum: $r_i = \hat{y}_i - y_i$



Das Verfahren am einfachen Beispiel

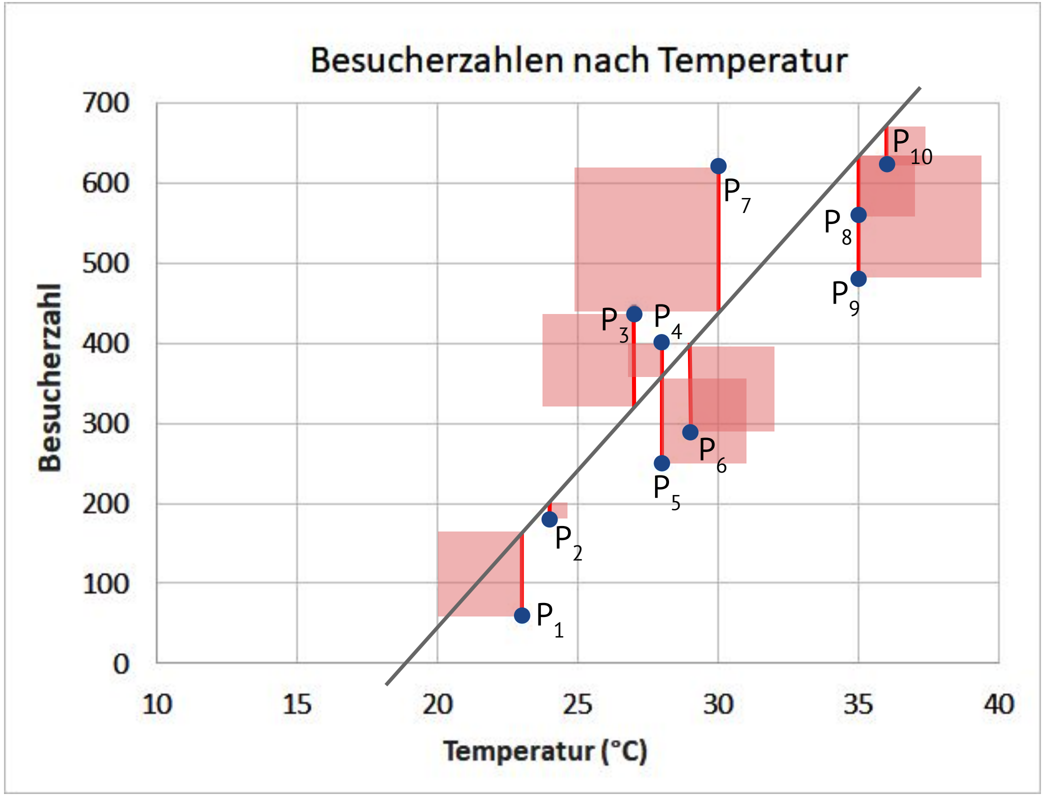
- 2-dimensionale Daten liegen vor
  - (x_i, y_i)*, für i=1,…,n, mit n=9*
- Annahme: Linearer Zusammenhang
- Modell: $f(x): \hat{y} = mx + b$
- "Beste" Gerade: $\min \sum_{i=1}^{n} (r_i)^2$
- Residuum: $r_i = \hat{y}_i - y_i$

  $\Leftrightarrow r_i = mx_i + b - y_i$



Das Verfahren am einfachen Beispiel

- 2-dimensionale Daten liegen vor
  - (x_i, y_i), *für i=1,…,n, mit n=9*
- Annahme: Linearer Zusammenhang
- Modell: $f(x) : \hat{y} = mx + b$
- "Beste" Gerade: $\min \sum_{i=1}^{n} (r_i)^2$
- Residuum: $r_i = \hat{y}_i - y_i$
  $\Leftrightarrow r_i = mx_i + b - y_i$
- $f(b, m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$



Mathematische Minimierung

- Gesucht: Minimum der Funktion $f(b, m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$

Mathematische Minimierung

- Gesucht: Minimum der Funktion $f(b, m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$

- Minimum: Partielle Ableitungen gleich null:

$$f_b'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i) = 0$$

$$f_m'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i) x_i = 0$$

Mathematische Minimierung

- Gesucht: Minimum der Funktion $f(b, m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$
- Minimum: Partielle Ableitungen gleich null:

$$f_b'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i) = 0$$

$$f_m'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i)\, x_i = 0$$

- Lineares Gleichungssystem (2 Lin. Unabh. Gleichungen & 2 Unbekannte: 1 Lösung)

Mathematische Minimierung

- Gesucht: Minimum der Funktion $f(b, m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$
- Minimum: Partielle Ableitungen gleich null:

$$f_b'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i) = 0$$

$$f_m'(b, m) = \sum_{i=1}^{n} 2(b + mx_i - y_i)\, x_i = 0$$

- Lineares Gleichungssystem (2 Lin. Unabh. Gleichungen & 2 Unbekannte: 1 Lösung)
- Lösung: $$m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad b = \bar{y} - m\bar{x}$$

Mit $\bar{x}$ und $\bar{y}$, den Mittelwerten von x_i und y_i

Anwendung der Least Squares Methode

$\{(x_i, y_i)\} = \{ (28, 400), (23, 60), (24, 180), (36, 625), (35, 560),$
$(29, 290), (30, 620), (27, 440), (28, 250), (35, 480) \}$

$$m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad b = \bar{y} - m\bar{x}$$

Anwendung der Least Squares Methode

$\{(x_i, y_i)\} = \{ (28, 400), (23, 60), (24, 180), (36, 625), (35, 560),$
$(29, 290), (30, 620), (27, 440), (28, 250), (35, 480) \}$

$$m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad b = \bar{y} - m\bar{x}$$

$\bar{x} = (28 + 23 + 24 + 36 + 35 + 29 + 30 + 27 + 28 + 35\,)\,/\,10 = 295\,/\,10 = 29{,}5$

$\bar{y} = (400 + 60 + 180 + 625 + 560 + 290 + 620 + 440 + 250 + 480\,)\,/\,10 = 3905\,/\,10 = 390{,}5$

Anwendung der Least Squares Methode

$\{(x_i, y_i)\} = \{ (28, 400), (23, 60), (24, 180), (36, 625), (35, 560),$
$(29, 290), (30, 620), (27, 440), (28, 250), (35, 480) \}$

$$m = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad b = \bar{y} - m\bar{x}$$

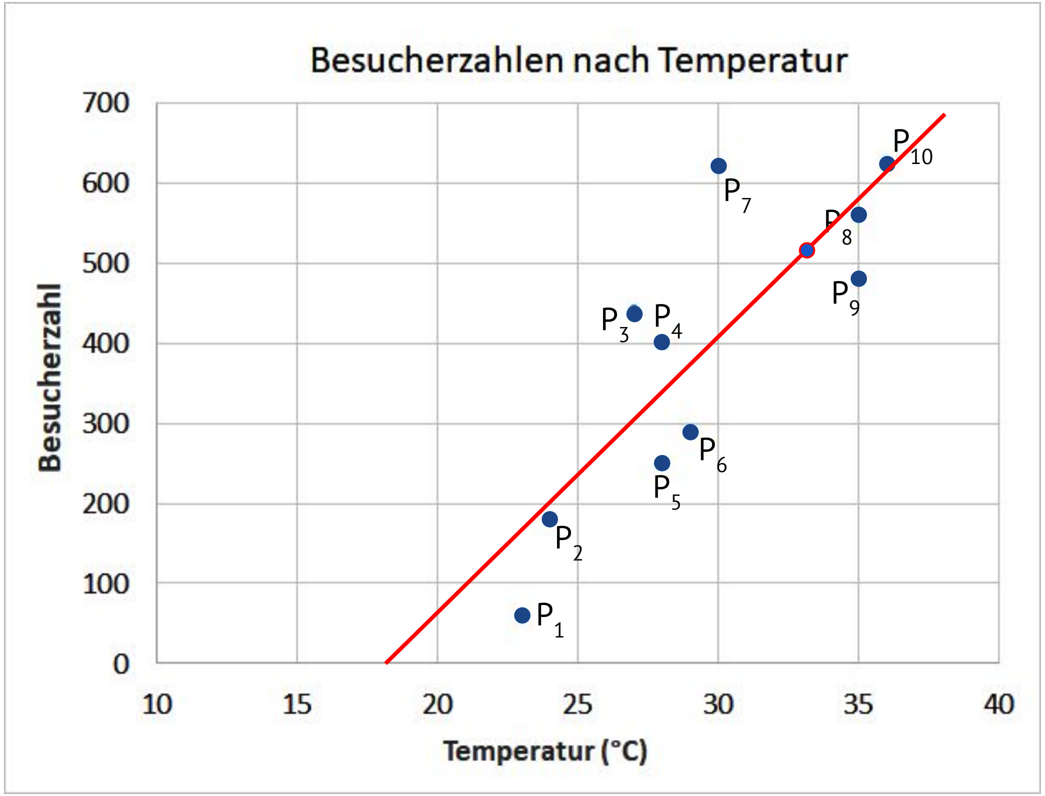
$\bar{x} = (28 + 23 + 24 + 36 + 35 + 29 + 30 + 27 + 28 + 35) / 10 = 295 / 10 = 29{,}5$

$\bar{y} = (400 + 60 + 180 + 625 + 560 + 290 + 620 + 440 + 250 + 480) / 10 = 3905 / 10 = 390{,}5$

$$m = \frac{6492{,}5}{186{,}5} \approx 34{,}812 \qquad b = 390{,}5 - (6492{,}5/186{,}5)\, 29{,}5 \approx -636{,}464$$

$f(x): \ \hat{y} = 34{,}812\, x - 636{,}464$

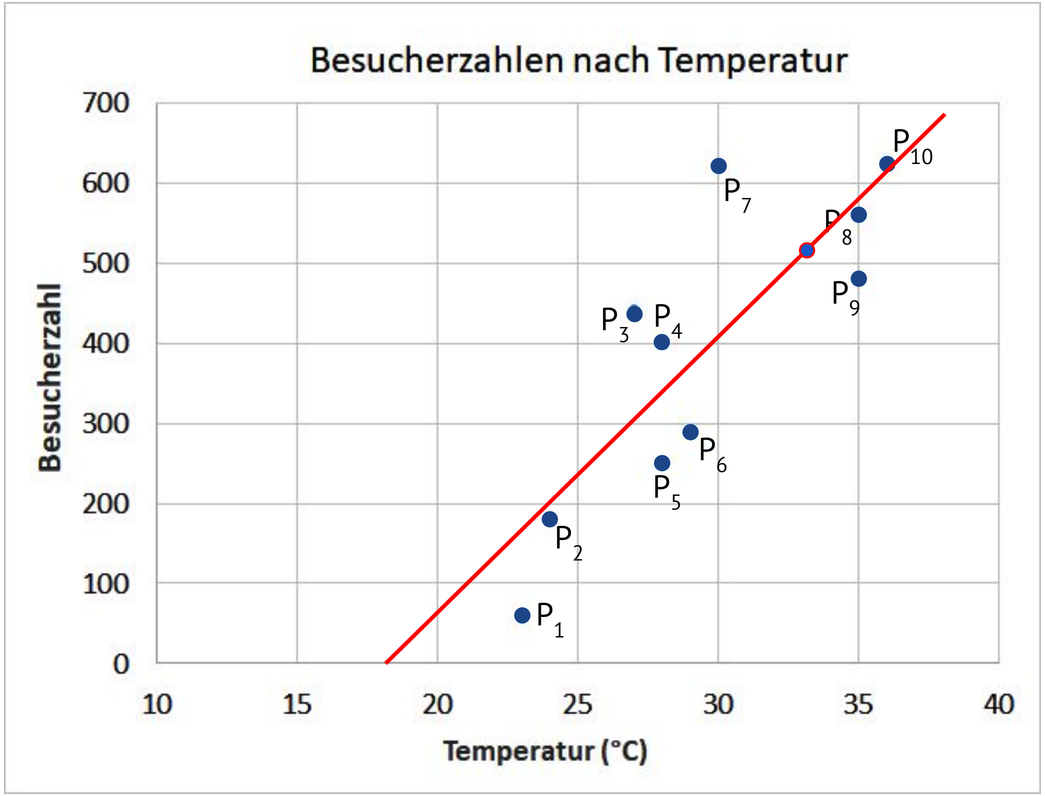
Anwendung der Least Squares Methode



$f(x): \ \hat{y} = 34.812\, x - 636.464$

$f(33) = 512.34$

Anwendung der Least Squares Methode



$f(x):\ \hat{y} = 34.812\, x - 636.464$

$f(33) = 512.34$

Bitte beachten Sie folgende Voraussetzungen:

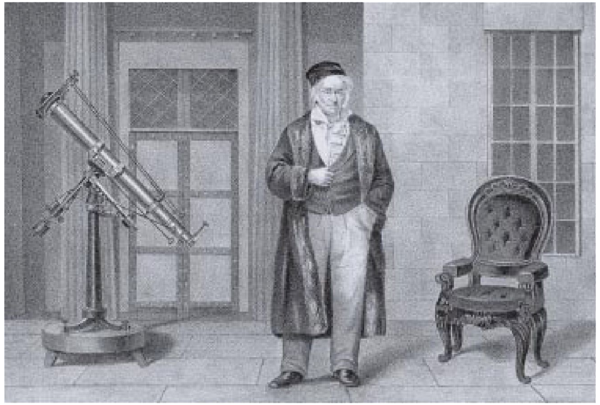
- Modell begrenzt gültig
- Modellwahl muss stimmen
- Empfindlich bei Ausreißern
- Annahme: Fehler normalverteilt

Least Squares Methode - Die erste Anwendung

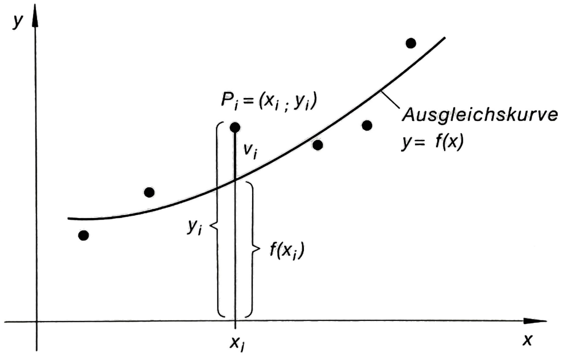
- Im Jahr 1801 entdeckte der Astronom Giuseppe Piazzi den Himmelskörper Ceres

- Lösungsansatz von Gauss:
  - Keplers 1. Gesetz:
    Die Bahn eines jeden Planeten ist eine Ellipse, wobei die Sonne in einem der beiden Brennpunkte steht.
  - Bestimme die Bahn, die die Summe der quadrierten Abweichungen minimiert.

- Ende 1801 wurde Ceres tatsächlich genau an der von Gauss berechneten Stelle entdeckt.



Carl Friedrich Gauß auf der Terrasse „seiner“ Sternwarte
Abbildung: Uni Göttingen, Uni|in|form, Ausgabe Nr.4/2004

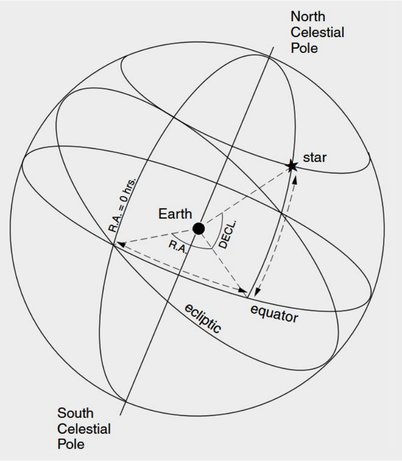


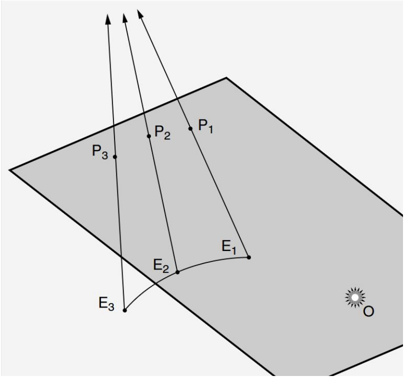
Messpunkte P_i mit “Ausgleichskurve”
L. Papula; *Mathematik für Ingenieure und Naturwissenschaftler,* Band 3, Vieweg+Teubner Verlag, 2011, https://doi.org/10.1007/978-3-8348-8133-5

Least Squares Methode - Die erste Anwendung

- Es gab nur drei unsichere Positionsmessungen
- Gauss arbeitete mit der Modellannahme einer Ellipse
- Er minimierte die Summe der Fehlerquadrate:

$$\text{Minimiere} \sum (y_i - f(x_i, \text{Parameter}))^2$$





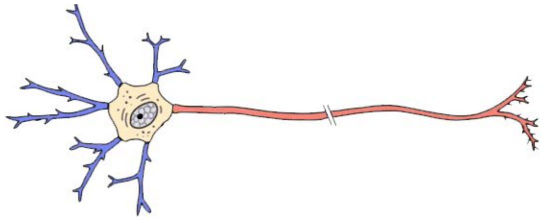
J. Tennenbaum, B. Director; *How Gauss Determined The Orbit of Ceres*
https://archive.schillerinstitute.com/fid_97-01/982_orbit_ceres.pdf

Einsatz der Least-Squares-Methode heute

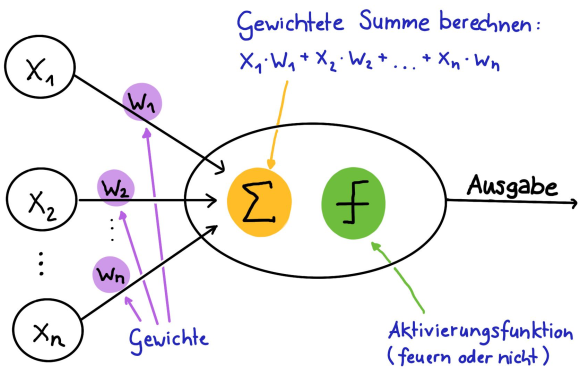
- Neuronale Netze kommen heute in vielen Gebieten zum Einsatz
  - Gesichtserkennung
  - Tumordiagnose
  - Generative KI

- Neuronale Netze sind KI-Modelle, die grob wie ein vereinfachtes Gehirn funktionieren
  - "künstlichen Neuronen" bilden echte Neuronen nach
  - Neuron berechnet gewichtete Summe
  - Aktivierungsfunktion entscheidet über Ausgabe

Neuron



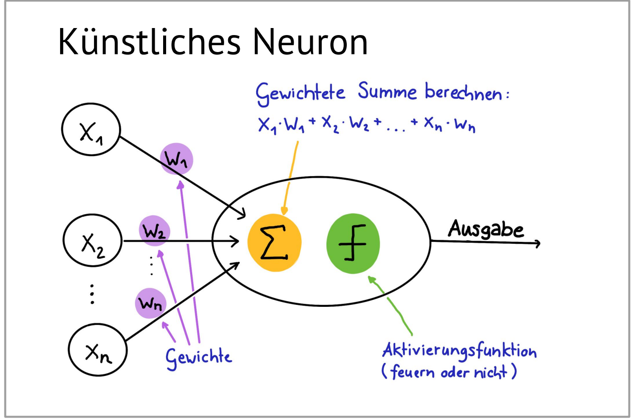
Künstliches Neuron

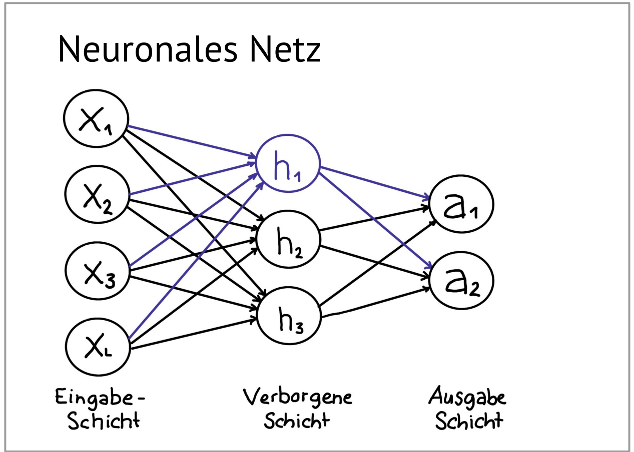


Einsatz der Least-Squares-Methode heute

- Neuronale Netze lernen durch Anpassung der Gewichte
  - Neuronen schichtweise angeordnet
  - Gewichte simulieren die Signalstärke
  - Lernen durch Anpassung der Gewichte

- Bewertung des Neuronalen Netzes anhand des Vorhersagefehlers
  - MSE: Der Mittlere Vorhersagefehler soll minimiert werden
  - Backpropagation: Gewichte werden angepasst um den Fehler zu minimieren





Mini-Quiz: Least Squares im Überblick

Was ist das Ziel der Methode der kleinsten Quadrate?

- Die Anzahl der Quadrate zu minimieren
- Die Summe der Abweichungen zwischen Vorhersage und Messwert zu minimieren
- Die Summe der quadrierten Abweichungen zwischen Vorhersage und Messwert zu minimieren

Mini-Quiz: Least Squares im Überblick

Was ist das Ziel der Methode der kleinsten Quadrate?

- Die Anzahl der Quadrate zu minimieren
- Die Summe der Abweichungen zwischen Vorhersage und Messwert zu minimieren
- Die Summe der quadrierten Abweichungen zwischen Vorhersage und Messwert zu minimieren

Mini-Quiz: Least Squares im Überblick

Welche Aussage über Residuen ist korrekt?

- Ein Residuum ist die Differenz zwischen beobachtetem und vorhergesagtem Wert
- Große Residuen deuten auf eine gute Modellanpassung hin
- Residuen sind immer positiv

Mini-Quiz: Least Squares im Überblick

Welche Aussage über Residuen ist korrekt?

- Ein Residuum ist die Differenz zwischen beobachtetem und vorhergesagtem Wert
- Große Residuen deuten auf eine gute Modellanpassung hin
- Residuen sind immer positiv

Mini-Quiz: Least Squares im Überblick

Welche Aussage zur Methode der kleinsten Quadrate ist korrekt?

- Die Methode kann auch bei nichtlinearen Zusammenhängen verwendet werden
- Die Methode funktioniert nur bei exakt linearen Daten
- Die Methode ist ungeeignet für verrauschte Daten

Mini-Quiz: Least Squares im Überblick

Welche Aussage zur Methode der kleinsten Quadrate ist korrekt?

- Die Methode kann auch bei nichtlinearen Zusammenhängen verwendet werden
- Die Methode funktioniert nur bei exakt linearen Daten
- Die Methode ist ungeeignet für verrauschte Daten

Vielen Dank!

Fragen & Diskussion

Vorlesungsfolien
Übung mit Lösung
Ergänzende Materialien

https://github.com/MartinaEchtenbruck/Least-Squares-Materials

Kontakt:
Dr. Martina Echtenbruck
martina.echtenbruck@th-koeln.de