Technische
Hochschule
Wildau [FH]
Technical University
of Applied Sciences

Maschinelles Lernen

Univariate Lineare Regression und einfacher Gradientenabstieg

Prof. Dr. Rainer Stollhoff

Supervised Learning

Technische Hochschule Wildau
Technical University of Applied Sciences

Supervised Learning

1. Aufgabe A

   Vorhersage $\hat{Y} = A(X)$

2. Qualität Q

   Verlustfunktion $\mathrm{L}(\hat{Y}, \mathrm{Y})$

3. Erfahrung E

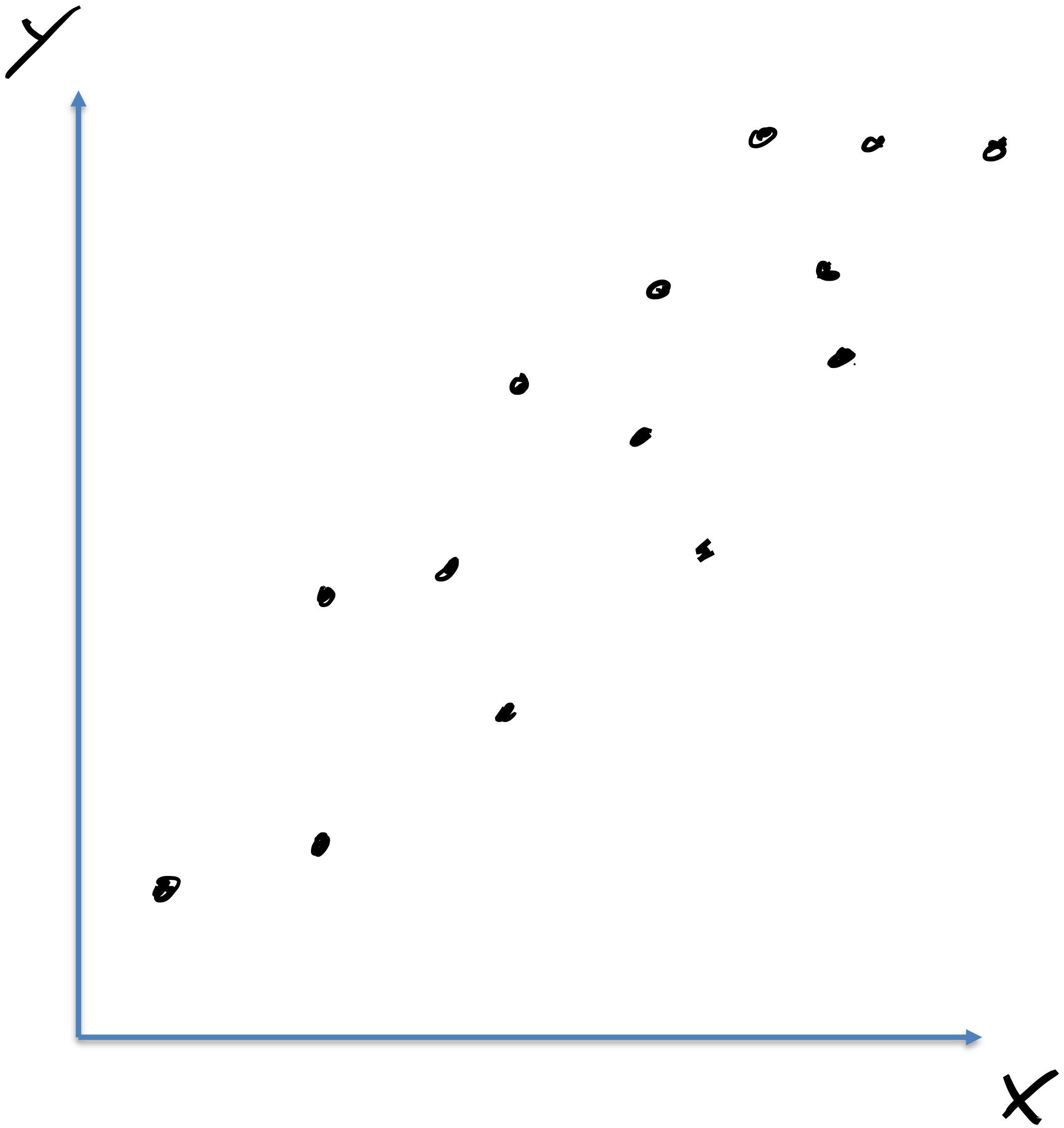
Datensatz
(x_i, y_i) für $i = 1, \cdots, n$

4. Maschine

Eine Maschine ***lernt*** aus Erfahrung E eine Aufgabe A mit der Qualität Q, wenn die Qualität Q beim erfüllen der Aufgabe A mit Erfahrung E steigt (T. Mitchell, MIT, 1988)

Einfache univariate Regression

Technische Hochschule Wildau
Technical University of Applied Sciences
WILDAU



- **Aufgabe**

  sage anhand von x den Zielwert y vorher

  $$\hat{y} = f(x)$$

- **Erfahrung**

  Beobachtungen: (x_i, y_i)

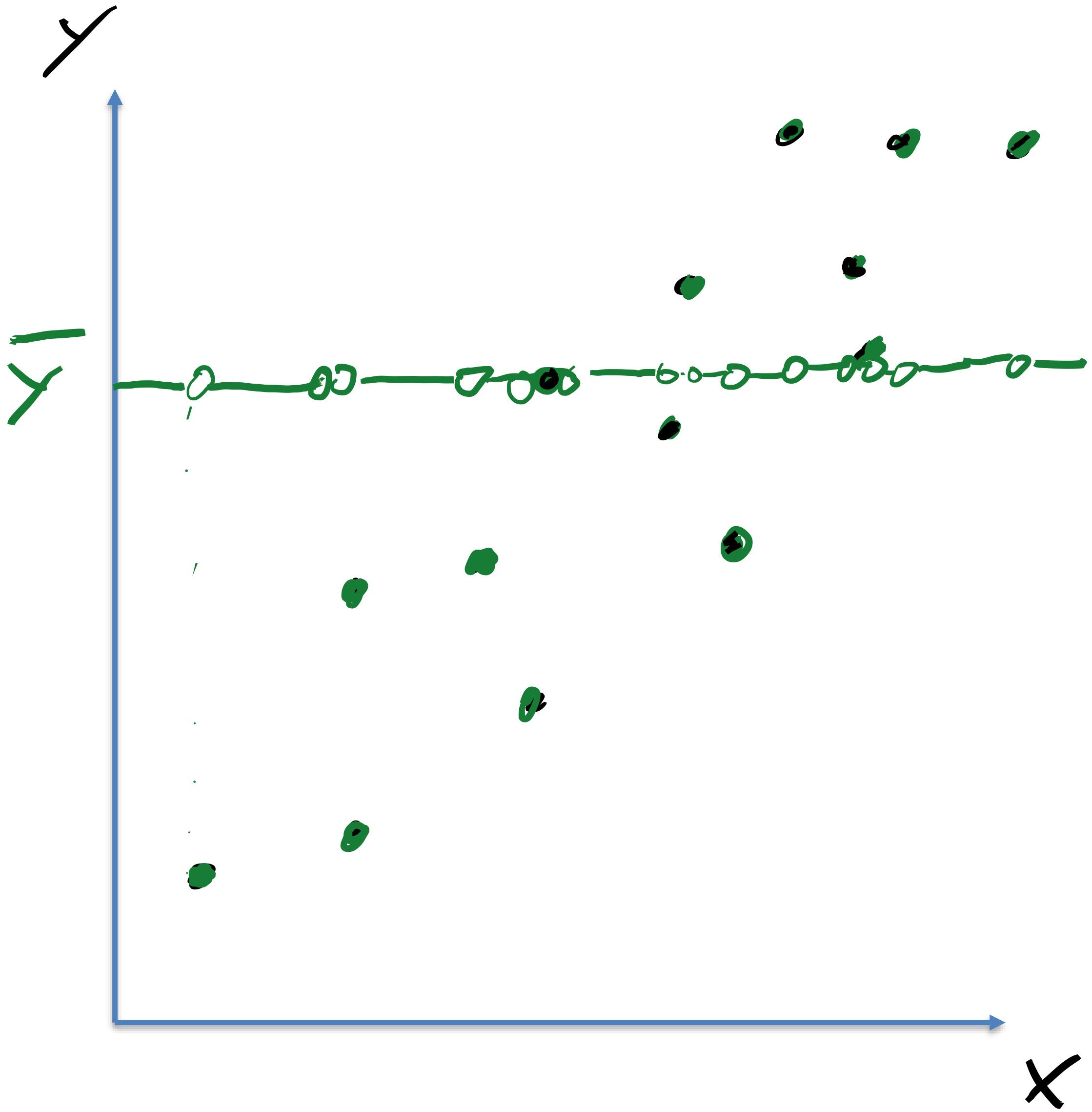
- **Qualität**

  sage den Zielwert möglichst gut vorher

  Gleichheit: $\hat{y} == y$

  Abstand: $(\hat{y} - y)^2$

Einfache univariate Regression



- **Aufgabe**

  sage anhand von x den Zielwert y vorher

  $$\hat{y} = \hat{y}(x) = f(x)$$

- **Erfahrung**

  Beobachtungen: (x_i, y_i)

- **Qualität**

  sage den Zielwert <u>möglichst</u> <u>gut</u> vorher

  Gleichheit: $\hat{y} == y$

  Abstand: $(\hat{y} - y)^2$

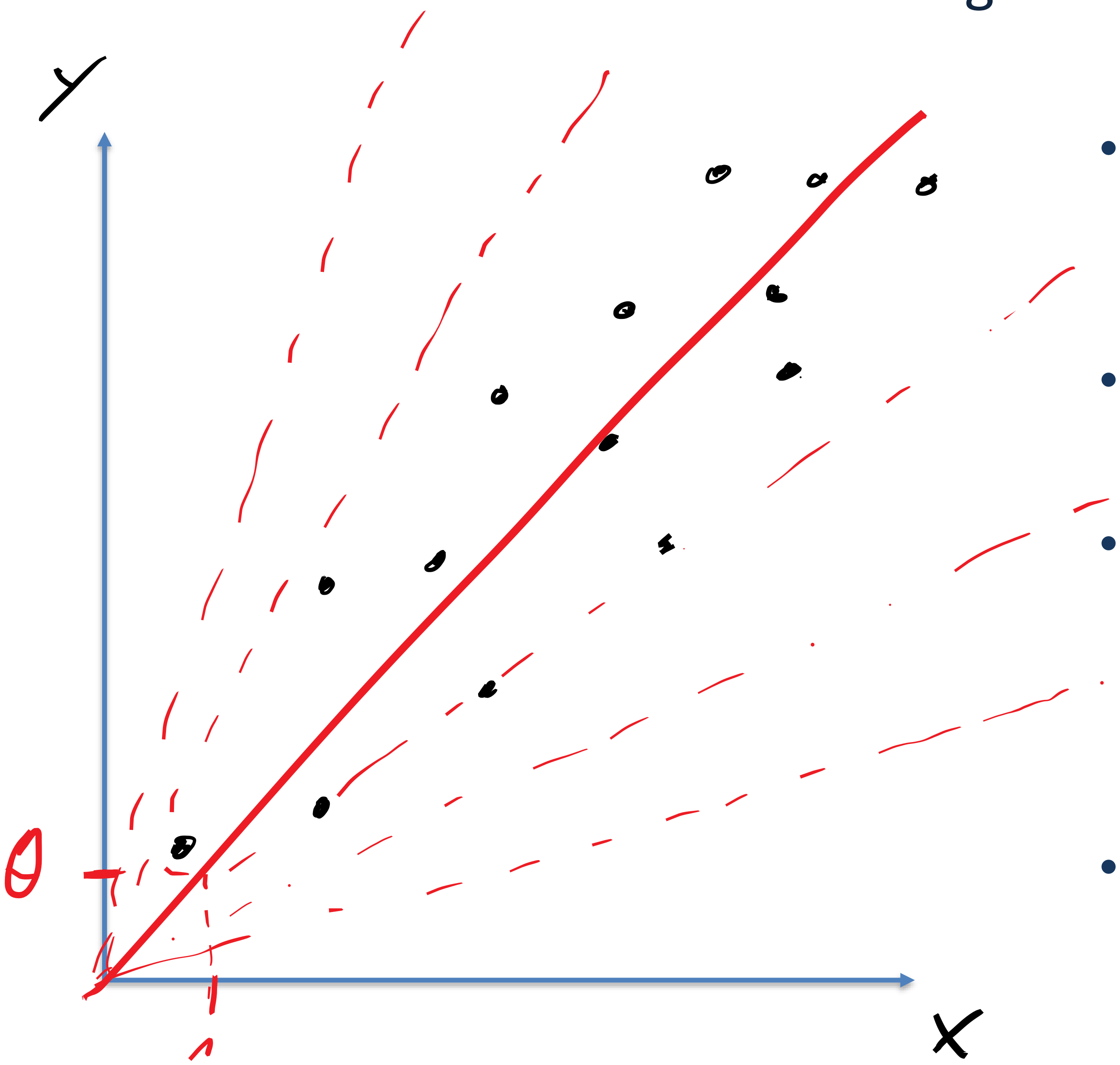
- Maschine bzw. Modell

  Look-up-table: $\hat{y}(x) = y_i$ für $\mathrm{x} = x_i, \bar{y}$ sonst

  Lineares Modell: $\hat{y} = \theta \cdot x$

Einfache univariate Regression

Technische Hochschule Wildau
Technical University of Applied Sciences
WILDAU



- **Aufgabe**

  sage anhand von x den Zielwert y vorher

  $$\hat{y} = \hat{y}(x) = f(x)$$

- **Erfahrung**

  Beobachtungen: (x_i, y_i)

- **Qualität**

  sage den Zielwert <u>möglichst</u> <u>gut</u> vorher

  Gleichheit: $\hat{y} == y$

  Abstand: $(\hat{y} - y)^2$

- Maschine bzw. Modell

  Look-up-table: $\hat{y}(x) = y_i$ für $x = x_i$, $\bar{y}$ sonst

  Lineares Modell: $\hat{y} = \theta \cdot x$

Einfache univariate Regression – analytische Lösung

Technische Hochschule Wildau
Technical University of Applied Sciences

Aufgabe: Regression, d.h. Vorhersage $\hat{y} = \hat{y}(x) = f(x)$

Erfahrung: Datensatz $(x_i, y_i)_{i=1}^{n}$

Qualität: Quadratische Verlustfunktion

$$L(y, \hat{y}) = \sum_{i=1}^{n} (y_i - \hat{y_i})^2 = \sum_{i=1}^{n} (y_i - f(x_i; \theta))^2 = L(\theta)$$

Maschine: vereinfachte lineare Regression mit

$$\mathrm{f}(x; \theta) = \theta \cdot x$$

Lernen: Finde einen Wert für θ , der die quadratische Verlustfunktion minimiert

Analytisch über die Nullstelle der ersten Ableitung

$$\frac{d}{d\theta} L(\theta) = \frac{d}{d\theta} \sum_{i=1}^{n} (y_i - f(x_i; \theta))^2 = \frac{d}{d\theta} \sum_{i=1}^{n} (y_i - x_i \cdot \theta)^2 = \sum_{i=1}^{n} 2 \cdot (y_i - x_i \cdot \theta) \cdot (-x_i) = 0$$

$$\theta_{opt} = {\sum_{i=1}^{n} y_i x_i} \Big/ {\sum_{i=1}^{n} x_i^2} = {Cov(x, y)} \Big/ {Var(x)}$$

Einfache univariate Regression – Gradientenabstieg

Aufgabe: Regression, d.h. Vorhersage $\hat{y} = \hat{y}(x) = f(x)$

Erfahrung: Datensatz $(x_i, y_i)_{i=1}^{n}$

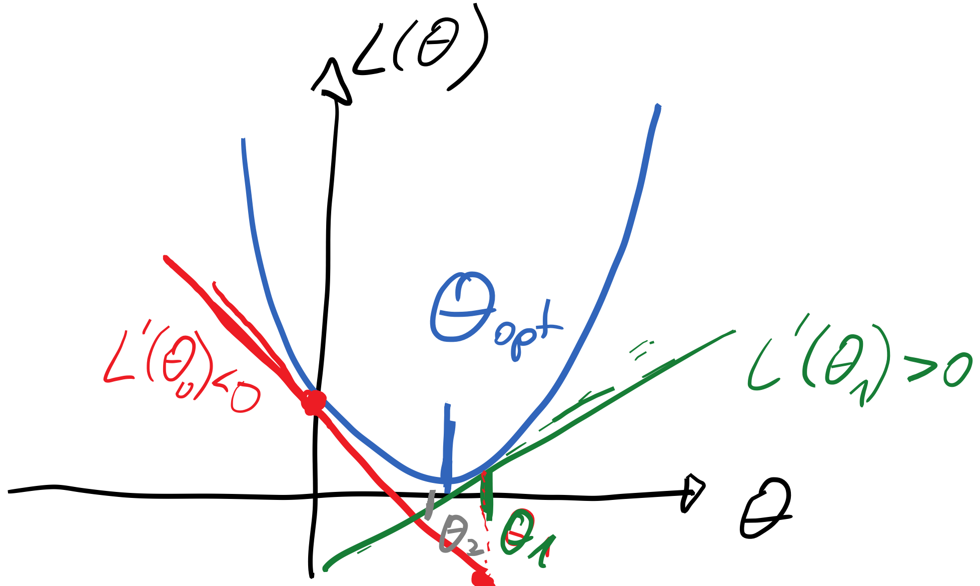
Qualität: Quadratische Verlustfunktion

$$L(y, \hat{y}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - f(x_i;\theta))^2 = L(\theta)$$

Maschine: vereinfachte lineare Regression mit

$$\mathrm{f}(x;\theta) = \theta \cdot x$$

Lernen: Finde einen Wert für θ , der die quadratische Verlustfunktion minimiert



Startwert $\theta_0 = 0$ $\quad L(\theta_0) \quad L'(\theta_0)$

$\theta_1 = \theta_0 + \alpha \cdot L'(\theta_0)$ mit $\alpha = -1$

$\theta_2 = \theta_1 + \alpha \cdot L'(\theta_1)$

......

$\theta_{opt} = \theta_t + \alpha \cdot 0$

Einfache univariate Regression – analytische Lösung

Technische Hochschule Wildau
Technical University of Applied Sciences
WILDAU

Aufgabe: Regression, d.h. Vorhersage $\hat{y} = \hat{y}(x) = f(x)$

Erfahrung: Datensatz $(x_i, y_i)_{i=1}^{n}$

Qualität: Quadratische Verlustfunktion

$$L(y, \hat{y}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - f(x_i; \theta))^2 = L(\theta)$$

Maschine: vereinfachte lineare Regression mit

$$\mathrm{f}(x; \theta) = \theta \cdot x$$

Lernen: Finde einen Wert für θ , der die quadratische Verlustfunktion minimiert

Durch geeignete Wahl von θ in einem iterativen Prozess (Gradientenabstiegsverfahren):

1. Wähle Startwert z.B. $\theta^0 = 0$
2. Berechne Ableitung $\frac{d}{d\theta} L(\theta) = \frac{d}{d\theta} \sum_{i=1}^{n} (y_i - f(x_i; \theta))^2 = \frac{d}{d\theta} \sum_{i=1}^{n} (y_i - x_i \cdot \theta)^2$
   $$= \sum_{i=1}^{n} 2 \cdot (y_i - x_i \cdot \theta) \cdot (-x_i)$$
3. Update $\theta^{t+1} = \theta^t + \alpha \cdot \frac{d}{d\theta} L(\theta^t) = \theta^t - \alpha \cdot 2 \sum_{i=1}^{n} (y_i - x_i \cdot \theta^t) \cdot x_i = \theta^t + 2 \sum_{i=1}^{n} (y_i - x_i \cdot \theta^t) \cdot x_i$

mit $\alpha = -1$ (steilster Abstieg)