Quadratmittelprobleme

Thea Fenzlein

Universität Rostock

19.11.2024

Gliederung

KLEINSTE QUADRATE

Kleinste Quadrate

- Häufig verwendeter Ansatz zur näherungsweisen Lösung von überbestimmten oder ungenau definierten Gleichungssystemen.
- Statt eine exakte Lösung zu finden, wird versucht, die Summe der Quadrate der Residuen zu minimieren:

$$\|Ax - b\|_2^2 = \sum_{i=1}^{m} (A_i x - b_i)^2$$

- Bedeutungen der Terme:
  - $A_i x$: Die Vorhersage des Modells für die i-te Gleichung (bzw. Messwert).
  - b_i: Der tatsächliche gemessene Wert.
  - m: Die Anzahl der Gleichungen (Datenpunkte).
- Gesucht wird der Vektor x, der die Summe der quadrierten Abweichungen minimiert.

MODELLE UND KURVENANPASSUNG

Modelle und Kurvenanpassung

- Kurvenanpassung ist ein häufiges Mittel der Wahl.
- Sei t eine unabhängige Variable und $y(t)$ eine unbekannte Funktion von t, die wir approximieren möchten.
- Angenommen, es gibt m Beobachtungen, d.h. Werte von y, die an bestimmten t-Werten gemessen wurden:

$$y_i = y(t_i), \quad i = 1, \ldots, m$$

- Die Idee besteht darin, $y(t)$ durch eine lineare Kombination von n Basisfunktionen zu modellieren:

$$y(t) = \beta_1 \phi_1(t) + \cdots + \beta_n \phi_n(t)$$

Designmatrix und Modellbeschreibung

- Die Designmatrix X ist eine rechteckige Matrix der Dimension $m \times n$ mit Elementen:

$$x_{i,j} = \phi_j(t_i)$$

- Die Designmatrix hat in der Regel mehr Zeilen als Spalten. Das Modell wird geschrieben als:

$$y \approx X\beta$$

- Basisfunktionen $\phi_j(t)$:
  - Können nichtlineare Funktionen von t sein.
  - Können nichtlineare Parameter $\alpha_1, \ldots, \alpha_p$ enthalten.
- Unbekannte Parameter β_j treten linear im Modell auf:

$$y(t) \approx \beta_1\phi_1(t,\alpha) + \cdots + \beta_n\phi_n(t,\alpha)$$

Lineare Parameter β

Die unbekannte Funktion $y(t)$ wird modelliert als:

$$y(t) = \beta_1\phi_1(t) + \beta_2\phi_2(t) + \cdots + \beta_n\phi_n(t)$$

- Die Parameter β sind die linearen Gewichtungsfaktoren der Basisfunktionen und werden durch lineare Regression geschätzt.
- Designmatrix X mit Elementen $x_{i,j} = \phi_j(t_i)$.
- Das Modell kann als Vektorform geschrieben werden:

$$y \approx X\beta$$

Nichtlineare Parameter α

Die Basisfunktionen können nichtlineare Parameter enthalten:

$$y(t) \approx \beta_1\phi_1(t,\alpha) + \beta_2\phi_2(t,\alpha) + \cdots + \beta_n\phi_n(t,\alpha)$$

- $\alpha_1, \ldots, \alpha_p$: nichtlineare Parameter in den Basisfunktionen.
- Beispiel:

$$\phi_1(t,\alpha_1) = e^{\alpha_1 t}, \quad \phi_2(t,\alpha_2) = \sin(\alpha_2 t)$$

- Die Parameter α steuern die Form der Basisfunktionen und müssen durch nichtlineare Optimierung geschätzt werden.

Zusammenspiel von α, β und der Designmatrix X

- Die Designmatrix X hat die Dimension $m \times n$ und enthält d
  Basisfunktionen:

$$X = \begin{pmatrix} \phi_1(t_1, \alpha) & \phi_2(t_1, \alpha) & \dots & \phi_n(t_1, \alpha) \\ \phi_1(t_2, \alpha) & \phi_2(t_2, \alpha) & \dots & \phi_n(t_2, \alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(t_m, \alpha) & \phi_2(t_m, \alpha) & \dots & \phi_n(t_m, \alpha) \end{pmatrix}$$

- Die Elemente $x_{i,j} = \phi_j(t_i, \alpha)$ hängen von den Beobachtungswerten t_i und den nichtlinearen Parametern α ab.
- Das Modell wird wie folgt formuliert:

$$y \approx X\beta$$

Schätzung der Parameter

- Das Modell $y(t) \approx X\beta$ kann gelöst werden, wenn die Parameter α bekannt sind.
- Die Designmatrix X hängt von t_i und den nichtlinearen Parametern α ab.
- Um β_j und α_p zu schätzen, werden häufig iterative Verfahren verwendet:
  - Zuerst β mit festen α schätzen.
  - Dann α mit fixierten β optimieren.

Beispiele für Modelle

Gerade Linie:

$$y(t) \approx \beta_1 t + \beta_2$$

Polynome:

$$\phi_j(t) = t^{n-j}, \quad y(t) \approx \beta_1 t^{n-1} + \cdots + \beta_{n-1} t + \beta_n$$

Rationale Funktionen:

$$\phi_j(t) = \frac{t^{n-j}}{\alpha_1 t^{n-1} + \cdots + \alpha_{n-1} t + \alpha_n}, \quad y(t) \approx \frac{\beta_1 t^{n-1} + \cdots + \beta_n}{\alpha_1 t^{n-1} + \cdots + \alpha_n}$$

Weitere Modellbeispiele

Exponentielle Funktionen:

$$\phi_j(t) = e^{-\lambda_j t}, \quad y(t) \approx \beta_1 e^{-\lambda_1 t} + \cdots + \beta_n e^{-\lambda_n t}$$

Log-linear:

$$y(t) \approx K e^{\lambda t}, \quad \log(y) \approx \lambda t + \beta_2 \quad \text{mit} \quad \beta_1 = \lambda,\ \beta_2 = \log(K)$$

Gauss-Funktionen:

$$\phi_j(t) = e^{-\frac{(t-\mu_j)^2}{\sigma_j^2}}, \quad y(t) \approx \beta_1 e^{-\frac{(t-\mu_1)^2}{\sigma_1^2}} + \cdots + \beta_n e^{-\frac{(t-\mu_n)^2}{\sigma_n^2}}$$

Normalgleichungen

Wenn das Gleichungssystem überbestimmt ist, gibt es keine exakte Lösung. Stattdessen wird im Sinne der kleinsten Quadrate gerechnet:

$$\min_{\beta} \|X\beta - y\|$$

Durch das Multiplizieren beider Seiten mit X^T wird das System auf eine $n \times n$-Matrix reduziert:

$$X^T X\beta = X^T y$$

Wenn es viele Beobachtungen, aber nur wenige Parameter gibt, ist die Designmatrix X sehr groß, während $X^T X$ relativ klein ist. Die Lösung ist die Projektion von y auf den Raum, der von den Spalten von X aufgespannt wird:

$$\beta = (X^T X)^{-1} X^T y$$

Beispiel: Lineare Regression

Gegeben seien die Messpunkte:

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 4.5 \\ 6.1 \\ 7.6 \\ 10.9 \end{pmatrix}$$

Das Modell lautet:

$$y(x) = \beta_1 x + \beta_2, \quad \beta = \begin{pmatrix} m \\ b \end{pmatrix}$$

Die Designmatrix X:

$$X = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix}$$

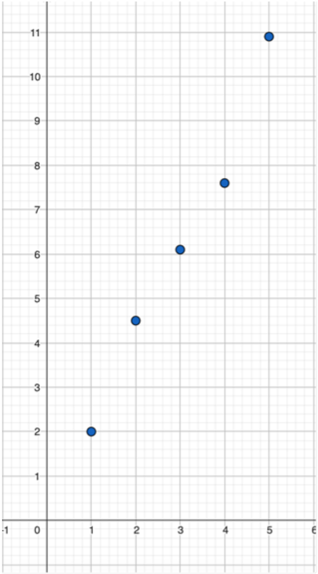

Abbildung 1

Schritt 1: Berechnung von $X^T X$

Anwendung der Normalengleichung:

$$\beta = \left(X^T X\right)^{-1} X^T y$$

$$X^T X = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 55 & 15 \\ 15 & 5 \end{pmatrix}$$

Schritte 2 und 3: Berechnung von $X^T y$ und $(X^T X)^{-1}$

Berechnung von $X^T y$:

$$X^T y = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 4.5 \\ 6.1 \\ 7.6 \\ 10.9 \end{pmatrix} = \begin{pmatrix} 114.2 \\ 31.1 \end{pmatrix}$$

Berechnung von $(X^T X)^{-1}$:

$$\begin{aligned} \left(X^T X\right)^{-1} &= \frac{1}{(55 \cdot 5) - (15 \cdot 15)} \begin{pmatrix} 5 & -15 \\ -15 & 55 \end{pmatrix} \\ &= \frac{1}{50} \begin{pmatrix} 5 & -15 \\ -15 & 55 \end{pmatrix} \\ &= \begin{pmatrix} 0.1 & -0.3 \\ -0.3 & 1.1 \end{pmatrix} \end{aligned}$$

Schritt 4: Berechnung von β

Berechnung von β:

$$\beta = \begin{pmatrix} 0.1 & -0.3 \\ -0.3 & 1.1 \end{pmatrix} \begin{pmatrix} 114.2 \\ 31.1 \end{pmatrix} = \begin{pmatrix} 2.09 \\ -0.05 \end{pmatrix}$$

Das Ergebnis lautet:

$$y = 2.09x - 0.05$$

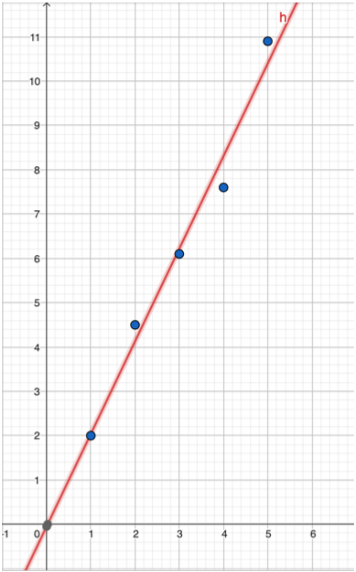

Abbildung 2

Nachteile der Normalgleichungen

- Die Verwendung der Matrixinversion zur Lösung eines Gleichungssystems ist aufwendiger und kann ungenau sein.
- Die Normalgleichung ist immer schlechter konditioniert als das ursprüngliche überbestimmte System.
- Aus diesen Gründen wird für den Großteil der Berechnungen ein Orthogonalisierungsalgorithmus verwendet.

NORMEN

Normen

Die Residuen sind die Differenzen zwischen den Beobachtungen und dem Modell:

$$r_i = y_i - \sum_{j=1}^{n} \beta_j \varphi_j(t_j, \alpha), \quad i = 1, \ldots, m$$

oder in Matrix-Vektor-Notation:

$$r = y - X(\alpha)\beta$$

Ziel: Werte für α und β finden, die die Residuen so klein wie möglich machen.

Interpolation

- Anzahl der Parameter gleich der Anzahl der Beobachtungen $\Rightarrow$ Residuen können Null sein.
- Für lineare Probleme: $m = n \Rightarrow$ Designmatrix X ist quadratisch.
- Falls X nicht singulär ist, sind die β-Werte die Lösung des quadratischen Gleichungssystems.

Kleinste Quadrate und Gewichtete kleinste Quadrate

Kleinste Quadrate: Minimierung der Summe der Quadrate der Residuen:

$$\|r\|^2 = \sum_{i=1}^{m} r_i^2$$

Gewichtete kleinste Quadrate: Wenn einige Beobachtungen wichtiger oder genauer sind als andere, erfolgt eine Verknüpfung und Minimierung von Gewichten ω_j mit den Beobachtungen:

$$\|r\|_\omega^2 = \sum_{i=1}^{m} \omega_i r_i^2$$

Beispiel: Fehler in der i-ten Beobachtung sei e_i, dann wählen wir $\omega_i = \frac{1}{e_i}$.

Eins-Norm und Unendlich-Norm

Eins-Norm: Minimierung der Summe der Absolutwerte der Residuen:

$$\|r\|_1 = \sum_{i=1}^{m} |r_i|$$

Unendlich-Norm: Minimierung des größten Residuum:

$$\|r\|_\infty = \max_{i=1,\dots,n} |r_i|$$

HOUSEHOLDER-TRANSFORMATION

Einführung in die Householder-Transformation

Householder-Transformationen sind Matrizen-Transformationen, die Grundlage für leistungsfähige und flexible Algorithmen bilden. Das Format einer Householder-Transformation lautet:

$$H = I - puu^T$$

- u ist ein beliebiger, von Null verschiedener Vektor.
- $p = \frac{2}{\|u\|^2}$
- uu^T ist eine Matrix vom Rang 1, jede Spalte ist ein Vielfaches von u und jede Zeile ein Vielfaches von u^T.

Eigenschaften und praktische Berechnungen

Die Matrix H hat folgende Eigenschaften:

$$H^T = H \quad \text{(symmetrisch)} \quad \text{und} \quad H^T H = H^2 = I \quad \text{(orthogonal)}$$

In der Praxis wird die Matrix H nie vollständig gebildet. Stattdessen wird nur die Anwendung der Transformation auf einen Vektor x berechnet:

$$\tau = pu^T x$$

$$Hx = x - \tau u$$

- τ ist ein Skalar, der das Produkt von u^T und x darstellt.
- Hx ist der transformierte Vektor, der von x abgezogen wird, um die reflektierte Richtung zu erzeugen.

Geometrische Interpretation der Householder-Transformation

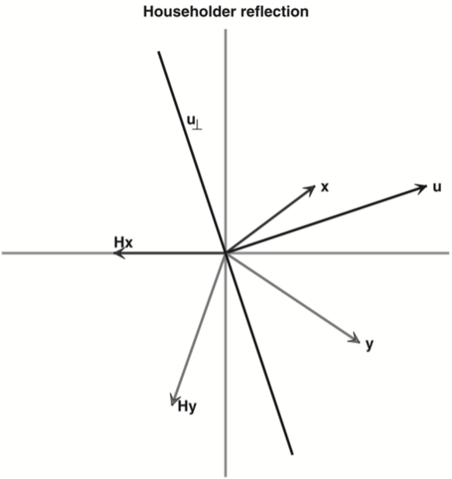

Abbildung 3

Geometrische Interpretation der Householder-Transformation

- $u_\perp$ steht senkrecht auf u.
- Matrix H spiegelt jeden Vektor an der Linie $u_\perp$.
- Für beliebigen Vektor x liegt der Punkt zwischen x und Hx, also der Vektor:
$$x - \frac{\tau}{2} u$$
tatsächlich auf der Linie $u_\perp$.

Householder-Transformation für einen gegebenen Vektor x

Für einen gegebenen Vektor x wird die Householder-Transformation, die alle bis auf die k-te Komponente von x auf null setzt, wie folgt definiert:

$$\sigma = \pm \|x\|,$$

$$u = x + \sigma e_k,$$

$$p = \frac{2}{\|u\|^2} = \frac{1}{\sigma u_k},$$

$$H = I - puu^T$$

Beispiel: Householder-Transformation

Gegeben sei der Vektor:

$$x = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$$

Ziel ist die Transformation, sodass alle bis auf die erste Komponente des resultierenden Vektors null sind. Das Ziel ist es, einen Vektor Hx der Form:

$$Hx = \begin{pmatrix} \sigma \\ 0 \\ 0 \end{pmatrix}$$

zu finden, wobei $\sigma = \pm\|x\|$ ist.

Schritt 1: Berechne σ

Berechne die Norm von x:

$$\|x\| = \sqrt{4^2 + 3^2 + 0^2} = \sqrt{16 + 9} = \sqrt{25} = 5$$

Wähle $\sigma = 5$.

Schritt 2: Berechne u

Berechne u:

$$u = x + \sigma e_1$$

wobei $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ der Einheitsvektor in der ersten Richtung ist.

$$u = \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix} + 5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \\ 0 \end{pmatrix}$$

Schritt 3: Berechne p

Berechne p:

$$\|u\| = \sqrt{9^2 + 3^2 + 0^2} = \sqrt{81 + 9} = \sqrt{90}$$

$$p = \frac{2}{\|u\|^2} = \frac{2}{90} = \frac{1}{45}$$

Schritt 4: Berechne Householder-Matrix H

Berechne die Householder-Matrix H:

$$H = I - puu^T$$

Berechne uu^T:

$$uu^T = \begin{pmatrix} 9 \\ 3 \\ 0 \end{pmatrix} \begin{pmatrix} 9 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 81 & 27 & 0 \\ 27 & 9 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Nun berechne H:

$$H = I - \frac{1}{45} \begin{pmatrix} 81 & 27 & 0 \\ 27 & 9 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1.8 & 0.6 & 0 \\ 0.6 & 0.2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} -0.8 & -0.6 & 0 \\ -0.6 & 0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Schritt 5: Transformiere x mit H

Berechne die Transformation Hx:

$$Hx = \begin{pmatrix} -0.8 & -0.6 & 0 \\ -0.6 & 0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 3 \\ 0 \end{pmatrix}$$

$$Hx = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}$$

Das Ergebnis der Transformation ist der Vektor:

$$Hx = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}$$

QR-FAKTORISIERUNG

QR-Faktorisierung

Die QR-Faktorisierung zerlegt die Matrix X in zwei Matrizen:

$$X = QR$$

- **Vollständige Version:** R hat die gleiche Größe wie X, Q ist quadratisch mit genauso vielen Zeilen wie X.
- **Gekürzte Version:** Q hat die gleiche Größe wie X, R ist eine quadratische Matrix mit genauso vielen Spalten wie X.
- R ist eine "rechteckige" (rechte Dreiecks-)Matrix.

QR-Faktorisierung - Householder-Transformationen

Householder-Transformationen werden auf die Spalten von X angewendet, um R zu erzeugen:

$$H_n H_{n-1} \ldots H_2 H_1 X = R$$

Die j-te Spalte von R ist eine Linearkombination der ersten j-Spalten von X, wobei die Elemente unter der Hauptdiagonalen null sind.

QR-Faktorisierung - Weiterer Prozess

Wenn die gleiche Anzahl an Transformationen auf die rechte Seite angewendet wird, entsteht das Gleichungssystem:

$$X\beta \approx y \quad \Rightarrow \quad R\beta \approx z$$

$$H_n \dots H_2 H_1 y = z$$

Die ersten n dieser Gleichungen bilden ein kleines, quadratisches, dreieckiges Gleichungssystem, das für β durch Rückwärtseinsetzen gelöst werden kann. Die restlichen Koeffizienten sind null.

Matrix Q in der QR-Faktorisierung

Die Matrix Q ergibt sich aus der QR-Faktorisierung:

$$Q = (H_n \dots H_2 H_1)^T$$

Wobei Q nicht unbedingt vollständig berechnet werden muss. Es gibt zwei Varianten der Faktorisierung:

- Wenn nur die ersten n Spalten berechnet werden, erhalten wir die ökonomische Faktorisierung.
- Wenn alle m Spalten berechnet werden, erhalten wir die vollständige Faktorisierung.

In beiden Fällen gilt:

$$Q^T Q = I$$

Das bedeutet, dass die Spalten von Q senkrecht zueinander stehen und die Länge 1 haben.

Modellierung des Bevölkerungswachstums mit QR-Zerlegung

Gegebene Daten:
Die Bevölkerungsgröße $P(t)$ in Millionen für die Zeitpunkte t (Jahre) ist wie folgt gegeben:

t (Jahre)	$P(t)$ (Millionen)
0	2.5
1	3.1
2	3.8
3	4.7
4	5.9

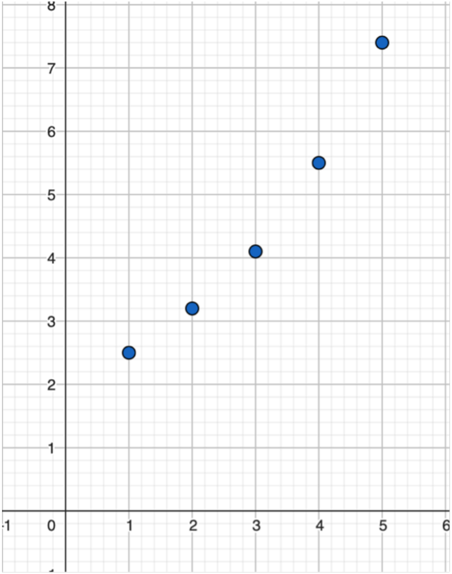

Abbildung 4

Ziel: Bestimme die Wachstumsfunktion

$$P(t) = \beta_0 e^{\beta_1 t},$$

die die Datenpunkte bestmöglich approximiert. Verwende die QR-Zerlegung und berechne:

- Die Matrix R
- Die Parameter β_0 und β_1
- Die endgültige Wachstumsfunktion

Transformation in ein lineares Modell

Die exponentielle Funktion wird linearisiert durch Logarithmieren:

$$\ln(P(t)) = \ln(\beta_0) + \beta_1 t$$

Setze:

$$y = \ln(P(t)), \quad \beta_0' = \ln(\beta_0)$$

Das lineare Modell wird:

$$y = \beta_0' + \beta_1 t$$

Transformation der Daten

Die transformierten Daten sind:

t (Jahre)	$P(t)$	$\ln(P(t))$
0	2.5	0.916
1	3.1	1.131
2	3.8	1.335
3	4.7	1.549
4	5.9	1.774

Die Zielwerte y sind:

$$y = \begin{bmatrix} 0.916 \\ 1.131 \\ 1.335 \\ 1.549 \\ 1.774 \end{bmatrix}$$

Setze die Designmatrix X:

$$X = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

QR-Zerlegung mittels Householder-Transformation

Schritt 1: Berechne u_1

$$u_1 = x_1 + \sigma e_1, \quad \sigma = \|x_1\| = \sqrt{1^2 + 1^2 + 1^2 + 1^2 + 1^2} = \sqrt{5}$$

$$u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \sqrt{5} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 + \sqrt{5} \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

QR-Zerlegung: Householder-Matrix

Schritt 2: Berechne p_1

$$\|u_1\| = \sqrt{(1+\sqrt{5})^2 + 1^2 + 1^2 + 1^2 + 1^2} = \sqrt{2\sqrt{5}+10}$$

$$p_1 = \frac{2}{\|u_1\|^2} = \frac{1}{\sqrt{5}+5}$$

Schritt 3: Householder-Matrix H

$$H = I - puu^T$$

Dabei wird die Matrix H iterativ berechnet, um die QR-Zerlegung abzuschließen.

Schhritt 4: Berechnung der Matrix $u_1 u_1^T$

Berechnung:

$$u_1 u_1^T = \begin{bmatrix} 1+\sqrt{5} \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1+\sqrt{5} & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 6+2\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \end{bmatrix}$$

Schritt 4: Berechnung der Householder-Matrix H_1

Formel für H_1:

$$H_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \frac{1}{\sqrt{5}+5} \begin{bmatrix} 6+2\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} & 1+\sqrt{5} \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \\ 1+\sqrt{5} & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 - \frac{6+2\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & 1 - \frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1 - \frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1 - \frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1 - \frac{1}{\sqrt{5}+5} \end{bmatrix}$$

Schritt 5: H von links auf X multiplizieren

Berechnung von A_1:

$$A_1 = \begin{bmatrix} 1-\frac{6+2\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & 1-\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1-\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1-\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & -\frac{1}{\sqrt{5}+5} & 1-\frac{1}{\sqrt{5}+5} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} 1-\frac{6+2\sqrt{5}}{\sqrt{5}+5}-\frac{4(1+\sqrt{5})}{\sqrt{5}+5} & 1-\frac{6+2\sqrt{5}}{\sqrt{5}+5}-\frac{14(1+\sqrt{5})}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}+1-\frac{4}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5}+2-\frac{14}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}+1-\frac{4}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+3 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}+1-\frac{4}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+4 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}+1-\frac{4}{\sqrt{5}+5} & -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+5 \end{bmatrix}$$

Erste Spalte und Zeile streichen, um A_1' zu erhalten:

$$A_1' = \begin{bmatrix} -\frac{1+\sqrt{5}}{\sqrt{5}+5}+2-\frac{14}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+3 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+4 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5}-\frac{14}{\sqrt{5}+5}+5 \end{bmatrix}$$

Schritt 5: Bildung von H'

Betrachten wir die Spalte a_1':

$$a_1' = \begin{bmatrix} -\frac{1+\sqrt{5}}{\sqrt{5}+5} + 2 - \frac{14}{\sqrt{5}+5} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 3 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 4 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 5 \end{bmatrix}$$

Berechnung von σ_2:

$$\sigma := \sqrt{\left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} + 2 - \frac{14}{\sqrt{5}+5}\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 3\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 4\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 5\right)^2}$$

Berechnung von u_2

Berechnung von u_2:

$$u_2 = a_1' + \sigma_2 e_1$$

$$= \begin{bmatrix} -\frac{1+\sqrt{5}}{\sqrt{5}+5} + 2 - \frac{14}{\sqrt{5}+5} + \sqrt{\left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} + 2 - \frac{14}{\sqrt{5}+5}\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 3\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 4\right)^2 + \left(-\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 5\right)^2} \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 3 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 4 \\ -\frac{1+\sqrt{5}}{\sqrt{5}+5} - \frac{14}{\sqrt{5}+5} + 5 \end{bmatrix}$$

Schritt 6: Berechnung von H' und H_2

Berechnung von H':

Wir berechnen H' mithilfe der folgenden Formeln:

$$p_2 = \frac{2}{\|u_2\|^2}$$

$$H = I - p \cdot uu^T$$

Berechnung von H_2:

Aus H' kann H_2 gebildet werden. Dabei setzen wir die Transformation in die Matrix ein.

$$H_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & H & H & H & H \\ 0 & H & H & H & H \\ 0 & H & H & H & H \\ 0 & H & H & H & H \end{bmatrix}$$

Dabei ist H der Wert, den wir aus der Berechnung von H' erhalten haben.

Berechnung von β_1 und β_2

Berechnung von R und A_1:

$$R = H_2 A_1 = \begin{bmatrix} \sqrt{5} & 3\sqrt{5} \\ 0 & \sqrt{10} \end{bmatrix}$$

Schritt 7: Rückwärtseinsetzen zur Berechnung von β:

$$y' = H_1 H_2 y$$

$$R\beta = y'$$

Die Werte für β_1 und β_2 können durch Rückwärtseinsetzen berechnet werden:

$$\beta_1 \approx 0.693, \quad \beta_2 \approx 0.267$$

Exponentialfunktion:
Mit diesen Werten ergibt sich die Exponentialfunktion:

$$y_t = 0.693 + 0.267t$$
$$P(t) = e^{0.693} \cdot e^{0.267t} = 1.9 \cdot e^{0.267t}$$

Fazit: Berechnung des Residuums

Berechnung des Residuums:

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Tabelle der Werte:

y_i	$\hat{y}_i$	$y_i - \hat{y}_i$
2.5	2.48	0.02
3.2	3.24	−0.04
4.1	4.23	−0.13
5.5	5.52	−0.02
7.4	7.22	0.18

Berechnung des Residuums:

$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = (0.02)^2 + (-0.04)^2 + (-0.13)^2 + (-0.02)^2 + (0.18)^2 = 0.0517$

PROBLEMLÖSUNG MIT MATLAB

Kleinste-Quadrate-Problem

Gegeben sei ein lineares System der Form:

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

mit einer $m \times n$-Matrix $\mathbf{A}$ und einem Vektor $\mathbf{b} \in \mathbb{R}^m$, wobei $m > n$. Das Ziel ist es, den Vektor $\mathbf{x} \in \mathbb{R}^n$ zu finden, der das System im Sinne der kleinsten Quadrate löst, d.h. den Fehler

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$$

zu minimieren.

QR-Zerlegung

Die QR-Zerlegung von $\mathbf{A}$ zerlegt die Matrix $\mathbf{A}$ in zwei Matrizen:

$$\mathbf{A} = \mathbf{QR}$$

wobei:

- $\mathbf{Q}$ eine orthogonale Matrix ist ($\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$),
- $\mathbf{R}$ eine obere Dreiecksmatrix ist.

Durch die QR-Zerlegung kann das System $\mathbf{Ax} = \mathbf{b}$ umgeschrieben werden zu:

$$\mathbf{QRx} = \mathbf{b}$$

Multiplizieren wir beide Seiten mit $\mathbf{Q}^T$, erhalten wir:

$$\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$$

Da $\mathbf{R}$ eine obere Dreiecksmatrix ist, kann das System durch Rückwärtseinsetzen gelöst werden.

Householder-Transformation

Die Householder-Transformation wird verwendet, um die Matrix $\mathbf{A}$ in eine obere Dreiecksmatrix $\mathbf{R}$ zu zerlegen.
Die Householder-Matrix $\mathbf{H}$ ist definiert als:

$$\mathbf{H} = \mathbf{I} - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

wobei $\mathbf{v}$ ein Vektor ist, der so gewählt wird, dass $\mathbf{HA}$ eine obere Dreiecksmatrix ergibt.

Lösung des Systems

Nachdem wir $\mathbf{A}$ in $\mathbf{QR}$ zerlegt haben, berechnen wir $\mathbf{Q}^T\mathbf{b}$, um das System

$$\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$$

zu lösen.
Die Lösung des Systems $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$ ergibt den Vektor $\mathbf{x}$, der die kleinsten Quadrate minimiert.

Beispiel: Kleinste Quadrate mit MATLAB

Gegeben sei das folgende System:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5 \\ 6 \\ 7 \\ 8 \end{pmatrix}$$

Die Schritte zur Lösung des Systems:

1. Berechne die QR-Zerlegung von $\mathbf{A}$.
2. Berechne $\mathbf{Q}^T\mathbf{b}$.
3. Löse das System $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$ mittels Rückwärtseinsetzen.

MATLAB-Code mit Beschreibung

- **Schritt 1: Definieren der Matrix A und des Vektors b**

  A = [1 2; 2 3; 3 4; 4 5]; b = [5; 6; 7; 8];

- **Schritt 2: Berechnung der QR-Zerlegung von A**

  [Q, R] = qr(A);

- **Schritt 3: Berechnung von $\mathbf{Q}^T\mathbf{b}$**

  $\mathrm{Qt}_b = Q' * b;$

- **Schritt 4: Lösung des Systems $\mathbf{Rx} = \mathbf{Q}^T\mathbf{b}$**

  $\mathrm{x} = \mathrm{R}\ \mathrm{Qt}_b;$

- **Schritt 5: Ausgabe der Lösung**

  disp(x);

Ergebnis

Nach Ausführung des MATLAB-Codes erhalten wir die Lösung des Kleinste-Quadrate-Problems:

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Dies ist die Lösung, die den Fehler $\|\mathbf{Ax} - \mathbf{b}\|_2$ minimiert.

PSEUDOINVERSE

Pseudoinverse

Die Definition der Pseudoinversen basiert auf der Frobenius-Norm einer Matrix:

$$\|A\|_F = \left(\sum_i \sum_j a_{ij}^2 \right)^{1/2}$$

Moore-Penrose-Pseudoinverse

Die Moore-Penrose-Pseudoinverse verallgemeinert und erweitert die üblichen Matrix-Inversen. Sie wird durch $Z = X^\dagger$ dargestellt.

- Wenn X quadratisch und nicht-singulär ist, dann ist die Pseudoinverse gleich der Inversen:

$$X^\dagger = X^{-1}$$

- Wenn X eine $m \times n$-Matrix ist, mit $m > n$ und vollem Rang, dann ist die Pseudoinverse die Matrix, die im Normalgleichungssystem vorkommt:

$$X^\dagger = (X^T X)^{-1} X^T$$

Eigenschaften der Pseudoinversen

Die Pseudoinverse $X^\dagger$ hat folgende Eigenschaften:

$$X^\dagger X = (X^T X)^{-1} X^T X = I$$

$$X X^\dagger = X (X^T X)^{-1} X^T$$

Die Pseudoinverse wird verwendet, wenn die Matrix X nicht quadratisch oder nicht vollrangig ist, sodass das Inverse von X nicht existiert.

Verwendung der Pseudoinversen

Für das Gleichungssystem:

$$x = X^{\dagger} b$$

wird die Pseudoinverse verwendet, wenn X rangdefizient ist oder mehr Spalten als Zeilen hat. In diesem Fall existiert das Inverse von $(X^T X)^{-1}$ nicht.

- Bei rangdefizienten Matrizen existiert keine eindeutige Lösung für das System.
- Die Pseudoinverse liefert in der Regel die beste Lösung im Sinne der kleinsten Quadrate.

RANGDEFIZIENZ

Rangdefizienz

Wenn X rangdefizient ist oder mehr Spalten als Zeilen hat, existiert das Inverse von $(X^T X)^{-1}$ nicht. Daher kann die Normalgleichung nicht zur Lösung verwendet werden, und es wird die Pseudoinverse genutzt.

- In solchen Fällen gibt es keine eindeutige Lösung für das System.
- Die Pseudoinverse bietet eine beste" Lösung im Sinne der kleinsten Quadrate.

Quellen

Literaturquellen:

- Moler, C. B. (2004). *Numerical computing with MATLAB* (Kapitel über Least Squares). SIAM.

Internetquellen:

- `https://studyflix.de/mathematik/qr-zerlegung-1786`
- `https://chatgpt.com/c/6734d4a8-3e08-8008-b887-becfb1310bc9`
- `https://www.geogebra.org/calculator`

Abbildungsverzeichnis:

- Abbildung 1: Erstellt mit GeoGebra
- Abbildung 2: Erstellt mit GeoGebra
- Abbildung 3: Moler, C. B. (2004). *Numerical computing with MATLAB* (Kapitel über Least Squares). SIAM.
- Abbildung 4: Erstellt mit GeoGebra