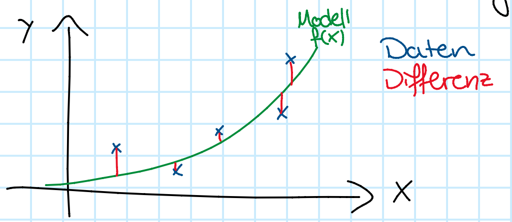
Woche 12

Freitag, 26. Mai 2023 11:21

Ausgleichsrechnung / Regression

Ziel: Finde Modellparameter, s.d. Daten möglichst gut vom Modell approximiert werden

Intuition:



Seien $y_i \in \mathbb{R}$ die gemessenen Werte für die Eingangsgrößen x_i

Wir wollen das Modell f so bestimmen, dass $\sum_{i=1}^{n} |f(x_i) - y_i|^2 = \|f(x) - y\|_2^2$ minimal

Wieso $\|\cdot\|_2^2$? Wenn wir annehmen, dass unsere gemessenen y_i fast den tatsächlichen y_i^{exakt} entsprechen, d.h. bis auf einen Messfehler übereinstimmen ($y_i = y_i^{exakt} + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2)$), dann ist $f(x)$ die bestmögliche Approximation die wir erhalten können.

Formal: Sei $\vec{a} = (a_1, \dots, a_m)$ der Modellparametervektor und $f_{\vec{a}} : \mathbb{R} \to \mathbb{R}, x \mapsto y$ die von $\vec{a}$ abhängige Modellfunktion. x_i: Messpunkte

y_i: Messwerte

Wichtig: Wir gehen davon aus, dass die x_i keinen Fehler haben

Falls $n > m \Rightarrow$ keine exakte Lösung, da nicht invertierbar

Bestimme nun $\vec{a}$, s.d. $\sum_{i=1}^{n} |y_i - f_{\vec{a}}(x_i)|^2$ minimal

Vektornotation:

$$F(\vec{a}) := \begin{pmatrix} y_1 - f_{\vec{a}}(x_1) \\ \vdots \\ y_n - f_{\vec{a}}(x_n) \end{pmatrix} \in \mathbb{R}^n \qquad \|F(\vec{a})\|_2 \text{ minimal}$$

Residuenvektor

Dann ist die optimale Wahl von $\vec{a}$ gegeben durch $\vec{a}^* = \underset{\vec{a} \in \mathbb{R}}{\operatorname{argmin}} \|F(\vec{a})\|_2$

Beispiel: Radioaktiver Zerfall

Menge N des noch nicht zerfallenem Material: $N(t) = N_0 e^{-\lambda t}$

N_0: Anfangsmenge
λ: Zerfallskonstante

Betrachten wir nun die Messzeiten t_i, $1 \le i \le n$ und die gemessenen, fehlerbehafteten noch nicht zerfallen N_i. Dann wollen wir die Parameter $\vec{a} = (N_0, \lambda)$ finden, s.d.

$$F(N_0, \lambda) = \begin{pmatrix} N_1 - N_{\vec{a}}(t_1) \\ \vdots \\ N_n - N_{\vec{a}}(t_n) \end{pmatrix} \quad \text{und} \quad \|F(N_0,\lambda)\|_2 \text{ minimal}$$

Lineare Ausgleichsrechnung:

$f_{\vec{a}}(x) = a_1 b_1(x) + \dots + a_m b_m(x)$ wobei $b_1, \dots, b_m : \mathbb{R} \to \mathbb{R}$ Basisfunktionen sind.

Intuitiv: $a_i \equiv$ linear effect of $b_i(x)$ on Y which is not explained by the other b_j

So können wir den Residuenvektor wie folgt schreiben:

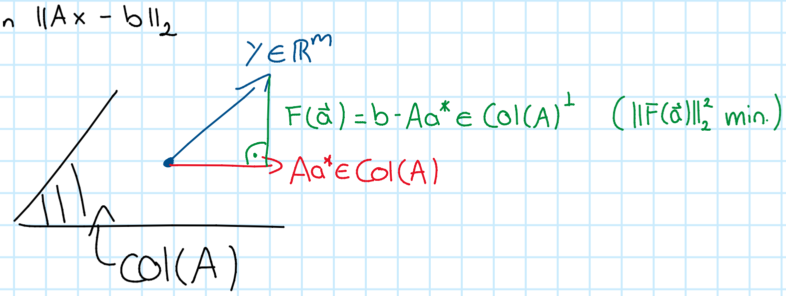
$$F(\vec{a}) = \vec{b} - A \cdot \vec{a}$$

Dann: $\|F(\vec{a})\|_2$ minimal $\iff$ $\|A\vec{a} - \vec{b}\|_2$ minimal

Wir suchen jetzt $\quad A = \begin{pmatrix} b_1(x_1) & \dots & b_m(x_1) \\ \vdots & & \\ b_1(x_m) & \dots & b_m(x_m) \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$$\vec{a}^* = \underset{x \in \mathbb{R}^n}{\operatorname{argmin}} \|Ax - b\|_2$$

Geometrisch:



Wie finden wir $\vec{a}^*$?

a) QR	b) SVD	c) Normalengleichungen
⊖ nur für $m \gg n$ und vollen Rang	⊕ kein voller Rang	⊕ für dünn besetzte A
	⊖ $m \gg n$	⊖ i.A. schlecht konditioniert

Python: `a = np.linalg.lstsq(A, b)[0]`

a) $\|Ax - b\|_2 = \|QRx - b\|_2 = \|Q^T(QRx - b)\|_2 = \|Rx - Q^T b\|_2$

$$Rx - Q^T b = \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} x - \begin{bmatrix} \hat{Q}^T \\ * \end{bmatrix} b = \begin{bmatrix} \hat{R}x - \hat{Q}^T b \\ * \end{bmatrix} \leftarrow \text{nicht von } x \text{ abh.}$$

D.h. $\|Rx - Q^T b\|_2$ minimal für $\hat{R}x - \hat{Q}^T b = 0$

$$\Rightarrow \vec{a}^* = \hat{R}^{-1} \hat{Q}^T b$$

b) Analog zu QR: $\Sigma_{1,r}' V_1^T x - u_1^T b = 0$

$\iff \quad a^* = \underbrace{V_1 \Sigma_{1,r}'^{T} U_1^T}_{A^+} b$

$A^+ \equiv$ Pseudoinverse

($A^+ \in \mathbb{R}^{n \times m}$ im Least square Sinn die beste Approximation an die Umkehrfunktion)

c) Normalengleichung

Wir wollen $f(\vec{x}) = \|Ax - b\|_2$ minimal

Idee: Minimum wenn $Df(\vec{x}) = 0$ und $D^2 f(\vec{x})$ pos. def.

Dann: $Df(\vec{x}) = 0 \iff A^* A \vec{x} = A^* \vec{b}$ $(*)$

Falls A vollen Rang: $D^2 f(\vec{x})$ automatisch pos. definit

$\vec{a}^* = \text{solve}(A^*A, A^*\vec{b})$

↳ i.A. A^*A schlecht konditioniert aka. unbrauchbar