Einführung in die Grundlagen der Numerik (WS 22/23)

Manuel Hinz

13. Oktober 2022

Inhaltsverzeichnis

Vorwort

Diese Mitschrift von der Vorlesung Einführung in die Grundlagen der Numerik (Dölz,WS 2022/2023) wird von mir neben der Vorlesung geschrieben und ist dementsprechend Fehleranfällig. Fehler gerne an mh@mssh.dev!

Kapitel 1

Orthogonalität

1.1 Grundlegende Definitionen

Definition 1.1. *Sei X ein $\mathbb{R}$ Vektorraum und $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$ eine Abbildung. $\langle\cdot,\cdot\rangle$ heißt **Skalarprodukt** oder inneres Produkt, falls*

$$\forall f \in X \setminus 0 : \langle f, f\rangle > 0 \tag{Positiviät}$$

$$\forall f, g \in X : \langle f, g\rangle = \langle g, f\rangle \tag{Symmetrie}$$

$$\forall \alpha, \beta \in \mathbb{R}, f, g, h \in X : \langle \alpha f + \beta g, h\rangle = \alpha\langle f, h\rangle + \beta\langle g, h\rangle \tag{Linearität im ersten Argument}$$

Bemerkung 1.2. *Symmetrie und Linearität im ersten Argument implizieren, dass $\langle\cdot,\cdot\rangle$ eine bilineare Abbildung ist.*

Definition 1.3. *Sei X ein $\mathbb{R}$-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Wir bezeichnen die zugehörige **Norm** (in Abhänigkeit von einem Vektor $f \in X$) mit*

$$\|f\| = \sqrt{\langle f, f\rangle}.$$

Lemma 1.4. *Sei X ein $\mathbb{R}$-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann gil die Cauchy-Schwarz-Ungleichung:*

$$\forall f, g \in X : \langle f, g\rangle \leq \|f\| \cdot \|g\| \tag{C.S.}$$

mit Gleichheit genau dann, wenn f und g linear abhängig sind.

Beweis. O.B.d.A. $f, g \neq 0$, da sonst offensichtlich Gleichheit gilt. Sei $\alpha \neq 0$, dann gilt mit $f, g \in X$ und $\alpha \in \mathbb{R}$:

$$0 \leq \|f - \alpha g\|^2 = \langle f - \alpha g, f - \alpha g\rangle = \|f\|^2 - 2\alpha\langle f, g\rangle + \alpha^2\|g\|^2$$

Wählen wir jetzt $\alpha = \frac{\langle f,g\rangle}{\|g\|^2}$ folgt:

$$\begin{aligned} 0 &\leq \|f\|^2 - \frac{2\langle f, g\rangle^2}{\|g\|^2} + \frac{\langle f, g\rangle^2}{\|g\|^2} \\ &\implies \langle f, g\rangle^2 \leq \|f\|^2 \cdot \|g\|^2. \end{aligned}$$

□

Eingefügte Bemerkung. *Rechnung zur Begründung von $\langle f - \alpha g, f - \alpha g\rangle = \|f\|^2 - 2\|\alpha\langle f, g\rangle + \alpha^2\|g\|^2$:*

$$\begin{aligned} &\langle f - \alpha g, f - \alpha g\rangle \\ &= \langle f, f - \alpha g\rangle - \alpha\langle g, f - \alpha g\rangle \\ &= \langle f, f\rangle - \alpha\langle f, g\rangle - \alpha\langle g, f\rangle + \alpha^2\langle g, g\rangle \\ &= \|f\|^2 - 2\|\alpha\langle f, g\rangle + \alpha^2\|g\|^2 \end{aligned}$$

Beispiel 1.5. *1.* $X = \mathbb{R}^n$ *und* $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ *(Euklidisches Skalarprodukt)*

2. $X = \mathbb{R}^n$, $\langle x, y\rangle = x^\perp A y$, *wobei* A *positiv definit und symmetrisch ist*

3. $I = [a, b], w : I \to \mathbb{R}$ *beschränkt und strikt positiv:*

$$X = \left\{ f : I \to \mathbb{R} : \int_a^b f(x)^2 w(t) dt < \infty \right\} = L^2(I, w)$$

mit

$$\langle f, g\rangle = \int_a^b f(t)g(t)w(t)dt$$

Eingefügte Bemerkung. *Die Definition von* $L^2(I, w)$ *ist hier nicht ganz richtig, man müsste natürlich noch Äquivalenzklassen, bzgl. Gleichheit bis auf Nullmengen, bilden. Dies wird hier, da Analysis 3 / Wtheo. nicht nicht vorrausgesetzt wird, ignoriert.*

Definition 1.6. *Sei* X *ein* $\mathbb{R}$*-VR mit Skalarprodukt* $\langle\cdot,\cdot\rangle$. $f, g \in X$ *heißen* ***orthogonal****, falls* $\langle f, g\rangle = 0$.

Bemerkung 1.7. *Im* $\mathbb{R}^n$ *mit dem euklidischen Skalarprodukt stimmt Definition 1.6, wegen*

$$\langle x, y\rangle = \|x\|\|y\| \cos(\theta), \theta = \angle(x, y),$$

mit unserem bisherigen Verständnis überein.

1.2 Bestapproximationseigenschaft

Definition 1.8. *Sei* V *ein* $\mathbb{R}$*-VR mit Skalarprodukt* $\langle\cdot,\cdot\rangle$ *und* U *ein Unterraum.*

$$U^\perp = \{v \in V : \langle v, u\rangle = 0, \forall u \in U\}$$

heißt das ***orthogonale Komplement*** *von* U.

Satz 1.9. *Unter den Annahmen von Definition 1.8 und der zusätzlichen Annahme, dass* U *endlich dimensional ist, gilt folgendes für* $v \in V$*:*

$$\|v - u\| = \min_{w\in U} \|v - w\|$$

genau dann, wenn $v - u \in U^\perp$.

Beispiel 1.10. $V = \mathbb{R}^2$, $U = span\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$ *mit euklidischem Skalarprodukt* $\langle\cdot,\cdot\rangle$. *Dann ist* $U^\perp = span\left\{\begin{pmatrix}1\\-1\end{pmatrix}\right\}$.

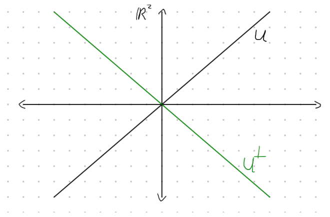

Abbildung 1.1: U und $U^\perp$

Beweis von Satz 1.9. Sei $v \in V$ und seien $u, w \in U$. Dann gilt:

$$\begin{aligned}\|v - w\|^2 &= \langle v - w, v - w\rangle = \langle (v - u) + (u - w), (v - u) + (u - w)\rangle \\ &= \|v - u\|^2 + 2\langle v - u, \underbrace{u - w}_{\in U}\rangle + \|u - w\|^2 \geq \|v - u\|^2\end{aligned}$$

mit Gleichheit genau dann, wenn $w - u = 0$ (da dann der $\|u - w\|$ Term verschwindet). $\square$

Bemerkung 1.11. *Der Satz sagt, dass es zu jedem $v \in V$ ein eindeutiges, bestmögliches $u \in U$ gibt.*

Definition 1.12. *Die Lösung aus Satz 1.9 heißt **orthogonale Projektion** von v auf U. Die Abbildung*

$$P : V \to U, v \mapsto P(v) \; \textbf{\textit{mit}} \; \|v - Pv\| = \min_{w \in U} \|v - w\|$$

*ist linear und wird **orthogonale Projektion** genannt.*

Eingefügte Bemerkung (Beweis der Linearität)**.** *Für $v_1, v_2 \in V$ und $\alpha \in \mathbb{R}$ gilt:*

$$v_1 - Pv_1 \in U^\perp$$
$$v_2 - Pv_2 \in U^\perp$$

Daher

$$\alpha(v_1 - Pv_1) + (v_2 - Pv_2) = (\alpha v_1 + v_2) - (\alpha P v_1 + P v_2) \in U^\perp.$$

Aber dann muss $\alpha P v_1 + P v_2$ schon, wegen der Eindeutigkeit, $P(\alpha v_1 + v_2)$ sein.

Bemerkung 1.13. *Satz 1.9 gilt auch, wenn U durch $W = w_0 + U$ ersetzt wird. Die orthogonale Projektion ist analog definiert..*

Frage: Die Orthogonale Projektion hat offenbar gute Eigenschaften. Aber: wie berechnen wir sie? Wie wählen wir U?

- Berechnung ist leicht
- U wählen schwierig

1.3 Orthonormalbasen

Definition 1.14. *Sei X ein $\mathbb{R}$-VR mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $X_n \subset X$ ein endlich dimensionaler Teilraum mit Basis $\{\varphi_1, \ldots, \varphi_n\}$. Die Basis heißt **Orthogonalbasis**, falls*

$$\forall i \neq j : \langle \varphi_i, \varphi_j \rangle = 0$$

gilt und Orthonormalbasis (ONB), falls zusätzlich $\|\varphi_i\| = 1$ gilt. Das impliziert:

$$\langle \varphi_i, \varphi_j \rangle = \delta_{i,j}.$$

Beispiel 1.15.

1. *$\mathbb{R}^n$ mit euklidischem Skalarprodukt und kanonischer Basis*
2. *$X = L^2(I, 1)$ mit entsprechendem Skalarprodukt und X_n der Raum der trigonometrischen Polynome bis Grad n. Dann ist folgendes eine ONB:*

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sin(x)}{\sqrt{\pi}}, \frac{\cos(x)}{\sqrt{\pi}}, \ldots, \frac{\sin(nx)}{\sqrt{\pi}}, \frac{\cos(nx)}{\sqrt{\pi}} \right\}$$

Eingefügte Bemerkung. *Trigonometrische Polynome sind Funktionen der Form*

$$f(t) = \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx).$$

Die größte Faktor vor dem x ist der Grad eine trigonometrischen Polynoms.

Satz 1.16. *Sei $\{\varphi_1, \ldots, \varphi_n\}$ eine ONB von $X_n \subset X$. Dann gilt*

1. *$f = \sum_{i=1}^{n} \langle \varphi_i, f \rangle \varphi_i$*

2. $\|f\|^2 = \sum_{i=1}^n \langle \varphi_i, f\rangle^2$
3. *Die orthogonale Projektion* f_n *von* $f \in X \setminus X_n$ *ist gegeben durch*

$$f_n = \sum_{i=1}^n \langle \varphi_i, f\rangle \varphi_i$$

4. *im Fall von 3.:*

$$\|f_n\|^2 = \sum_{i=1}^n \langle \varphi_i, f\rangle^2 \leq \|f\|$$

Beweis. 1.:

$$\begin{aligned} f \in X_n &\implies \exists \alpha_i \in \mathbb{R} : f = \sum_{i=1}^n \alpha_i \varphi_i \\ &\implies \langle \varphi_i, f\rangle = \langle \varphi_i, \sum_{j=1}^n \alpha_j \varphi_j\rangle = \sum_{j=1}^n \alpha_j \langle \varphi_i, \varphi_j\rangle = \alpha_i \end{aligned}$$

2.:

$$\begin{aligned} &\|f\|^2 = \langle f, f\rangle \\ &= \langle \sum_{i=1}^n \alpha_i \varphi_i, \sum_{j=1}^n \alpha_j \varphi_j\rangle = \sum_{i,j=1}^n \alpha_i \alpha_j \delta_{i,j} = \sum_{i=1}^n \alpha_i^2 \end{aligned}$$

3.:

$$\begin{aligned} &f \in X \setminus X_n : \\ &\|f - \underbrace{\tilde{f}_n}_{\in X_n}\| = \langle f - \sum_{i=1}^n \tilde{\alpha}_i \varphi_i, f - \sum_{i=1}^n \tilde{\alpha}_i \varphi_i\rangle \\ &\qquad = \|f\|^2 - 2\sum_{i=1}^n \tilde{\alpha}_i \underbrace{\langle \varphi_i, f\rangle}_{=:\alpha_i} + \sum_{i,j=1}^n \alpha_i \alpha_j \langle \varphi_i, \varphi_j\rangle \\ &\qquad = \|f\|^2 - \sum_{i=1}^n \tilde{\alpha}_i \alpha_i + \sum_{i=1}^n \tilde{\alpha}_i^2 \overset{\text{Quadratische Ergänzung}}{=} \|f\|^2 - \sum_{i=1}^n \alpha_i^2 + \sum_{i=1}^n \underbrace{(\alpha_i - \tilde{\alpha}_i)^2}_{\geq 0} \end{aligned} \tag{1.1}$$

Dies wird minimiert, wenn $\tilde{\alpha}_i = \alpha_i$ ist.
4.:
$f \in X_n$ wurde in 2. gezeigt. Sonst:

$$\begin{aligned} &f \notin x_n \implies \text{ mit } \alpha_i = \tilde{\alpha}_i \text{ in } (1.1): \\ &0 \leq \|f - f_n\|^2 = \|f\|^2 - \sum_{i=1}^n \underbrace{\alpha_i^2}_{\langle \varphi_i, f\rangle^2} \end{aligned}$$

Es folgt die Behauptung. □

Vorteile von Orthogonalität:

- Bestapproximation
- Einfache Basisdarstellung

——————————Ende von Vorlesung 01 am 11.10.2022——————————

Kapitel 2

Das lineare Ausgleichsproblem

2.1 Problemstellung und Normalengleichung

Gegeben seien Punkte $(t_i, b_i) \in \mathbb{R}^2$ mit $i = 1, \dots, m$. Wir nehmen an, dass es eine Gestzmäßigkeit im Sinne eines parameterabhängigen Modelles

$$b_i = b(t_i) = b(t_i; \underbrace{x_1, \dots, x_n}_{\text{Parameter}}),$$

wobei die Parameter $x_1, \dots, x_n$ unbekannt seien, gibt. In der Praxis sind die Messungen zusätzlich mit Fehlern behaftet und das Modell gilt nur approximativ. Zusätzlich gibt es oft mehr Messungen als Parameter, d.h. $m > n$.
Frage: Gegeben die Messungen, können wir zugehörige Parameter bestimmen?
Annahme: b ist linear in den Parametern, d.h. es gibt Funktionen

$$a_i : \mathbb{R} \to \mathbb{R}$$

s.d.

$$b(t; x_1, \dots, x_n) = a_1(t)x_1 + \dots + a_n(t)x_n.$$

Idee: Formuliere ein lineares Gleichungssystem:

$$b_i \approx b(t_i; x_1, \dots, x_n) = a_1(t_i)x_1 + \dots + a_n(t_i)x_n, i = 1, \dots, m$$

kurz $Ax \approx b$ mit $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$.
Problem: Durch Modell- und Messfehler gilt das Gleichungssystem nur ungefähr, und wir mehr Gleichungen als Unbekannte ("das Gleichungssystem ist überbestimmt"). Wir können unser Gleichungssystem also im Allgemeinen nicht lösen.

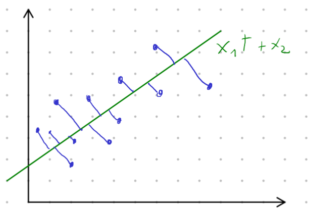

Abbildung 2.1: Datenpunkte und approximierte Gerade

Beispiel 2.1.

Idee: Finde Parameter, sodass das Modell "bestmöglich" mit den Messpunkten übereinstimmt, d.h. finde $(x_1, \dots, x_n)^t = x \in \mathbb{R}^n$ s.d.:

$$\|Ax - b\| = \min_{y \in \mathbb{R}} \|Ay - b\| \tag{2.1}$$

Definition 2.2. *Die Gleichung (2.1) heißt **lineares Ausgleichsproblem**. Der Term $Ax - b$ heißt **Residuum**.*

<u>Bemerke:</u> $V = \mathbb{R}^m, U = \text{Bild}(A) \subset V, \dim(\text{Bild}(A)) \underbrace{\leq n \leq m}_{\text{Grundannahme}}$

Statte V mit euklidischem Skalarprodukt aus.
$\overset{Satz 1.9}{\Longrightarrow}$ Es gibt genau ein $Ax \in \text{Bild}(A)$ so, dass

$$\|Ax - b\| = \min_{w \in U} \|w - b\|$$

gilt.
<u>Aber:</u> Wie berechnen wir x?

Satz 2.3. *Sei $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m \geq n, x \in \mathbb{R}^n$ ist genau dann eine Lösung von (2.1) bezüglich der euklidischen Norm, falls*

$$A^t Ax = A^t b. \tag{2.2}$$

Insbesondere ist das lineare Ausgleichproblem genau dann lösbar, falls rang$(A) = n$.

Beweis.

$$\begin{aligned} \|Ax - b\| &= \min_{y \in \mathbb{R}^n} \|Ay - b\| \\ &\overset{\text{Satz (1.9)}}{\iff} Ax - b \in U^\perp = \text{Bild}(A)^\perp \\ &\iff \forall y \in \mathbb{R}^n : \langle Ax - b, Ay \rangle = 0 \\ &\iff \forall y \in \mathbb{R}^n : \langle A^t Ax - A^t b, y \rangle = 0 \\ &\iff A^t Ax = A^t b \end{aligned}$$

Die letzte Gleichung ist genau dann invertierbar, wenn $A^t A$ vollen Rang hat, also wenn A vollen Rang (n) hat. □

Bemerkung 2.4. *Im beweis verwenden wir, dass $Ax - b$ orthogonal zu $U = \text{Bild}(A)$,*

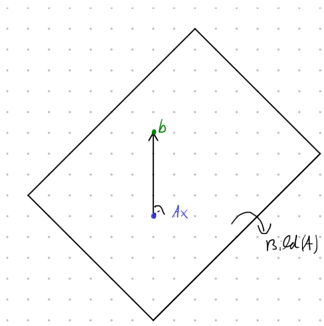

Abbildung 2.2: Hyperebene und Projektion

*d.h. eine Normale zur Hyperebene Bild(A) im R^m, ist. Deshalb heißt (2.2) auch **Normalengleichung**.*

Bemerkung 2.5. *Für $m = n$ und rang$(A) = n$ ist die Lösung des linearen Ausgleichproblems exakt (im mathematischen Sinne).*

Satz 2.6. *Für $A \in \mathbb{R}^{m \times n}$ ist $A^t A$ symmetrisch und positiv semidefinit. Falls $m \geq n$ ist $A^t A$ genau dann positiv definit, wenn rang$(A) = n$.*

Beweis.

- Symmetrisch: klar
- positiv semidefinit:

$$\forall x \in \mathbb{R}^n : x^t (A^t A) x = (Ax^t)(Ax) = \|Ax\|_2^2 \geq 0$$

- positiv definit: $\text{rang}(A) = n \implies Ax = 0 \iff x = 0 \implies \|Ax\|_2 = 0 \iff x = 0 \implies$ Behauptung.

$\square$

Einfachste Möglichkeit zur Lösung von (2.2): Berechne $A^t A, A^t b$, löse LGS mittels Cholesky. Kosten sind ungefähr:

$$\frac{n^2 m}{2} + m \cdot n + \frac{n^3}{6} + \frac{n^2}{2} + \frac{n^2}{2} \approx \frac{mn^2}{2} \text{ für } m \gg n.$$

Eingefügte Bemerkung. *Anmerkung vom Donzent: $A^t A$ eig. immer schlecht zu berechnen.*

Aber: Dieser Vorgang ist schlechter konditioniert als das lineare Ausgleichsproblem:

Eingeschobene Definition / Wiederholung

$$\text{cond}(A) = \|A\| \|A^{-1}\|$$
$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

Falls $A \in \mathbb{R}^{n \times n}$ spd (symmetrisch, positiv definit) gilt $\text{cond}_2((A^t A)) = \text{cond}_2(A)^2$.
Für $A \in \mathbb{R}^{m \times n}$ gelten ähnliche Überlegungen, siehe Deuflhard & Hohmann.

Beispiel 2.7. *Sei* $A = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix}$ *mit* $\epsilon > \underbrace{eps}_{Maschienengenauigkeit}, \epsilon^2 < eps.$

$$\implies A^t A = \begin{bmatrix} 1+\epsilon^2 & 1 \\ 1 & 1+\epsilon^2 \end{bmatrix} \overset{im\ Computer}{=} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$\implies$ *$A^t A$ ist im Computer singulär, obwohl A vollen Rang hat!*

Idee / Wunsch: Gebe einen Algorithmus an, der das lineare Ausgleichsproblem löst und nur auf A arbeitet.

2.2 Methode der Orthogonalisierung

Definition 2.8. *Eine Matrix $Q \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, wenn $Q^t Q = I$, d.h. falls die Spalten von Q eine ONB bzgl. des euklidischen Skalarprodukts bilden. Schreibe $Q \in O(n)$.*

Notation: $\langle \cdot, \cdot \rangle_2, \| \cdot \|_2$ für das euklidische Skalarprodukt / die euklidische Norm.

Lemma 2.9. *Für alle $Q \in O(n)$ gilt*

1. $\|Qx\|_2 = \|x\|_2$ *(Invarianz der Norm bzgl. orthogonaler Projektionen)*
2. $cond_2(Q) = 1$

Beweis. 1.: $\|Qx\|_2^2 = \langle Qx, Qx \rangle_2 = \langle Q^t Qx, x \rangle_2 = \langle x, x \rangle_2 = \|x\|_2^2$
2.: $\|Q\|_2 = \max_{\|x\|_2 = 1} \|Qx\| = 1$ und auch $\|Q^-1\|_2 = 1 \implies$ Behauptung. $\square$

Satz 2.10. *$A \in \mathbb{R}^{m \times n}, m \geq n, rang(A) = n$. Dann hat A eine QR-Zerlegung:*

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$$

wobei $Q \in O(m), R \in \mathbb{R}^{n \times n}$ eine obere Dreiecksmatrix ist.

Beweis. Schreibe das Gram-Schmidt-Orthogonalisierungsverfahren in Matrixform:

$$Q = \underbrace{\begin{bmatrix} \\ A_n & \dots & A_2 & A_1 \\ \\ \end{bmatrix} \underbrace{\begin{bmatrix} 1 & \dots & \dots & \dots & \frac{-\langle A_n, A_1\rangle_2}{\|A_1\|_2^2} \\ & \ddots & \dots & \dots & \vdots \\ & & 1 & \frac{-\langle A_3, A_2\rangle_2}{\|A_2\|_2^2} & \frac{-\langle A_3, A_1\rangle_2}{\|A_1\|_2^2} \\ & \mathbf{0} & & 1 & \frac{-\langle A_2, A_1\rangle_2}{\|A_1\|_2^2} \\ & & & & 1 \end{bmatrix}}_{R'}}_{\begin{bmatrix} B_n & \dots & B_1 \end{bmatrix}} \underbrace{\begin{bmatrix} \frac{1}{\|B_1\|_2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\|B_n\|_2} \end{bmatrix}}_{R''}$$

$\implies$ $Q \in R^{m\times n}, R'R''$ ist obere Dreiecksmatrix mit nicht-null Diagonaleinträgen
$\implies$ invertierbar: $R = (R'R'')^{-1}$
$\implies$ $QR = A$, wenn wir Q zu einer ONB von R^m erweitern. $\square$

———————Ende von Vorlesung 02 am 13.10.2022———————