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

Manuel Hinz

25. 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:

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

□

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. $X = \mathbb{R}^n$, $\langle x, y \rangle = x^\perp Ay$, *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* $U^\perp = span\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$.

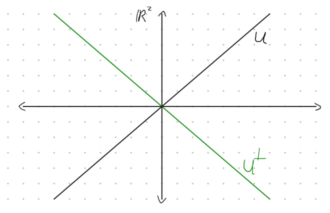

Abbildung 1.1: U und $U^\perp$

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* <u>*jedem*</u> $v \in V$ *ein* <u>*eindeutiges, bestmögliches*</u> $u \in U$ *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 $\alpha Pv_1 + Pv_2$ 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..*

<u>**Frage:**</u> 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.

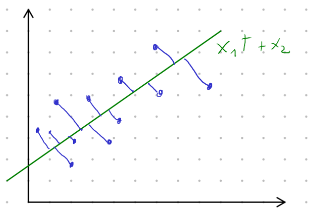

Abbildung 2.1: Datenpunkte und approximierte Gerade

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.:

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

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 $U = Bild(A)$,*

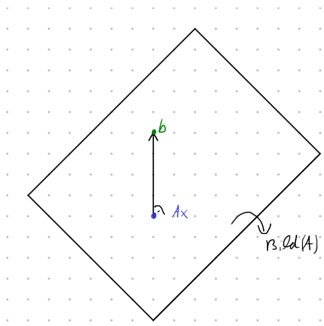

Abbildung 2.2: Hyperebene und Projektion

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.

□

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

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

Eingefügte Bemerkung. *Anmerkung vom Donzent: A^tA 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 $\text{cond}_2((A^tA)) = \text{cond}_2(A)^2$.
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 $\epsilon > \underbrace{eps}_{Maschienengenauigkeit}, \epsilon^2 < eps$.*

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

$\implies$ *A^tA 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 $Q^tQ = I$, 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.: $\|Qx\|_2^2 = \langle Qx, Qx\rangle_2 = \langle Q^tQx, x\rangle_2 = \langle x, x\rangle_2 = \|x\|_2^2$
2.: $\|Q\|_2 = \max_{\|x\|_2=1} \|Qx\| = 1$ und auch $\|Q^-1\|_2 = 1 \implies$ Behauptung. □

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. □

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

Satz 2.11. *Sei $A \in \mathbb{R}^{m\times n}, m \geq n, rang(A) = n, b \in \mathbb{R}^n$. Sei $A = QR$ eine QR-Zerlegung von A und*

$$\underbrace{Q^t A}_{=R} = Q^t b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \begin{matrix} \in \mathbb{R}^n \\ \in \mathbb{R}^{m-n} \end{matrix} .$$

Dann ist $x = R_1^{-}1 b_1$ die Lösung des linearen Ausgleichsproblems, wobei $R_1 \in \mathbb{R}^{n\times n}$ der obere Teil von R ist.

Beweis.

$$\begin{aligned} \|Ax-b\|_2^2 &\overset{\text{Lemma 2.9}}{=} \|Q^t(Ax-b)\|_2^2 \\ &= \left\| \begin{matrix} R_1 x - b \\ b_2 \end{matrix} \right\|_2^2 = \|R_1 x - b_1\|_2^2 + \|b_2\|_2^2 \\ &\geq \|b_2\|_2^2 \end{aligned}$$

$n = \text{rang}(A) = \text{rang}(R) = \text{rang}(R_1) \implies R_1$ invertierbar $\implies$ Behauptung

□

<u>Problem:</u>

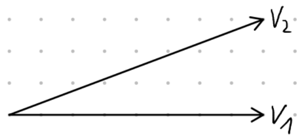

Abbildung 2.3: Problemstellung

$w_2 = v_2 - \frac{\langle v_2, v_1\rangle_2}{\langle v_1, v_1\rangle_2} v_1$ ist problematisch, falls $v_1 \approx v_2$ (Auslöschung). Beim Gram-Schmidt-Verfahren können Rundungsfehler auftreten. Es ist instabil.
<u>Ziel:</u> Stabiler Algorithmus um QR-Zerlegungen zu berechnen.

2.3 Grundüberlegungen zu Orthogonalisierungsverfahren

<u>Problemstellung:</u> Gegeben $v_1 = \alpha e_1 \in \mathbb{R}^2, v_2 \in \mathbb{R}^2$ transformiere v_2 auf $\tilde{w_2} = \beta e_2$, gebe β an.
<u>Gram-Schmidt:</u> $\beta = \|w\|_2$

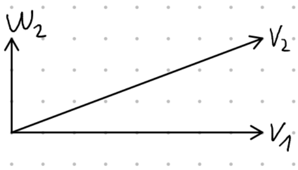

Abbildung 2.4: Gram-Schmidt

<u>**Drehungen:**</u> $\tilde{w}_2 = Qv_2$

$$Q = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

$$\beta = \|v_2\|_2$$

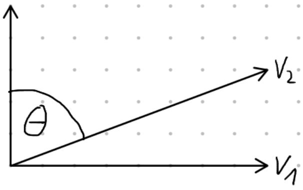

Abbildung 2.5: Drehungsansatz

<u>**Spiegelungen:**</u> $\tilde{w}_2 = Qv_2$, $Q = I - 2\frac{vv^t}{v^tv}$ und $\beta = \|v_2\|_2$

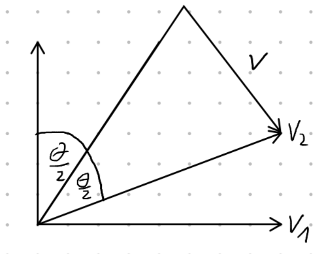

Abbildung 2.6: Spiegelungsansatz

<u>**Idee:**</u> Benutze orthogonale Transformationen $Q_1, \dots, Q_n$ um $A \in \mathbb{R}^{m\times n}$,rang$(A) = n$, sukzessive zu reduzieren.

$$A \rightsquigarrow Q_1A \rightsquigarrow Q_2Q_1A \rightsquigarrow \cdots \rightsquigarrow \begin{bmatrix} & R_1 \\ 0 & \\ & \\ & 0 \end{bmatrix}$$

Weil $\mathrm{cond}_2(Q) = 1$ ist die Vorgehensweise stabil, bzw. gut konditioniert.
<u>**Aber:**</u> Wie wählen wir $Q_1, \dots, Q_n$?

2.4 QR-Zerlegung mittels Givens-Rotationen

Definition 2.12. *Eine Matrix der Form*

$$\delta_{k,l} = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & c & & & s & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & -s & & & c & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{bmatrix}$$

, wobei die s, c Einträge in der k, lten Zeile / Spalte sind, heißen Givens-Rotationen.

Bemerke: Für $c = \cos(\theta), s = \sin(\theta)$ ist $\delta_{k,l}$ eine Drehung um θ in in der Koordinaten (k, l). $\delta_{k,l}$ ist Orthogonal.
Frage: Wie wählen wir c, s?
Gegeben $x \in \mathbb{R}^n$, elemeniere $l$te Koordinate zu 0.

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} x_k \\ x_l \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

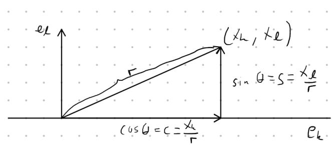

Abbildung 2.7: Trigonometriesetting

$$r^2 = x_k^2 + x_l^2 \implies \pm\sqrt{x_k^2 + x_l^2}$$

Aber: Diese Berechnungsweise ist nicht unbedingt stabil ($x_k \gg x_l$)
Stabile Variante:

$$\begin{gathered} \text{Falls } |x_l| > |x_k| \implies \tau = \tfrac{x_k}{x_l}, s = \tfrac{1}{\sqrt{1+\tau^2}}, c = s\tau \\ \text{Sonst: } \tau = \tfrac{x_l}{x_k}, c = \tfrac{1}{\sqrt{1+\tau^2}}, s = c\tau \end{gathered} \tag{2.3}$$

Beispielprozess:

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \rightsquigarrow \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ 0 & * & * \end{bmatrix} \rightsquigarrow \begin{bmatrix} * & * & * \\ * & * & * \\ 0 & * & * \\ 0 & * & * \end{bmatrix} \rightsquigarrow \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{bmatrix}$$

Algorithm 2.13

Input: $A \in \mathbb{R}^{m \times n}, m \geq n$
Output: R von der QR-Zerlegung (A wird zerstört "in place")

for $j = 1, \ldots, n$ **do**
  for $i = m, m-1, \ldots, j+1$ **do**
    Berechne c, s wie in (2.3)
    $A[i-1:i, j:n] = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}^t A[i-1:i, j:n]$
  end for
end for

$\underline{m \approx n}$**:**
c, s: In jedem Eintrag einmal Wurzeln ziehen: $\implies \frac{n^2}{2}$ Quadratwurzeln und $\frac{4n^3}{3}$ Multiplikationen
$\underline{m \gg n}$**:** $m \cdot n$ Quadratwurzeln und $2m \cdot n^2$ Multiplikationen

Bemerkung 2.14. *Der Algorithmus 2.4 berechnet nur R von der QR-Zerlegung. Zur Berechnung von Q müssten zusätliche Operationen investiert werden um die Givens-Rotation auf I anzuwenden.*
Für das lineare Ausgleichsproblem benötigen wir $Q^t b$, weshalb wir den Algorithmus auf $\left[\begin{array}{c|c} A & b \end{array}\right]$ anwenden können (da $R = Q^t A$).

Bemerkung 2.15. *Für $m = n$ ist die QR-Zerlegung eine (teure) ALternative zur LR-Zerlegung.*

2.5 QR-Zerlegung mittels Householder-Transformationen

Definition 2.16. *Für $v \in \mathbb{R}^n, v \neq 0$, heißt*

$$Q = I - 2\frac{\overbrace{vv^t}^{\in \mathbb{R}^{n \times n}}}{\underbrace{v^t v}_{\in \mathbb{R}}}$$

Householder-Transformation / Reflexion / Spiegelung.

> **Wichtig!**
>
> Nicht vv^t berechnen, das ist sehr uneffizient!

Für $a, v \in \mathbb{R}^n, v \neq 0$ ist $Qa = \left(I - 2\frac{vv^t}{v^t v}\right) a = a - 2\frac{\langle v,a\rangle_2}{\langle v,v\rangle_2} v$

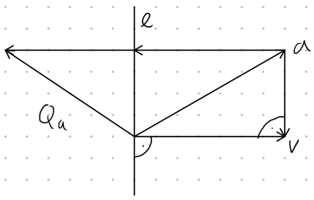

Abbildung 2.8: Householder-Transformationssetting

Qa ist a an l gespiegelt.

Lemma 2.17. *Für eine Householder-Transformation $Q \in \mathbb{R}^{n \times n}$ gilt:*

1. *Q ist symmetrisch*
2. *Q ist orthogonal*
3. *Q ist involutionisch (eine Involution), d.h. $Q^2 = I$*

Beweis. Nachrechnen. □

<u>Frage:</u> Gegeben $a \in \mathbb{R}^n$, wie müssen wir v wählen, so dass $Qa = \alpha e_1$ für $\alpha \in \mathbb{R}$?
<u>Beobachte:</u>

1. $|\alpha| = \|\alpha e_1\|_2 = \|Qa\|_2 = \|a\|_2$

2. $\underbrace{a - 2\frac{\langle v, a, \rangle}{\langle v, v\rangle}}_{\in\mathbb{R}} v = Qa$

$\implies v \in \operatorname{span}(\alpha e_1 - a) \implies \alpha = \pm \|a\|_2$

Vermeide Auslöschung $\implies \alpha = -\operatorname{sign}(a_1) \cdot \|a\|_2$

Effiziente Berechnung: Beobachte:

$$\begin{aligned} \|v\|_2^2 = \langle v, v\rangle_2 &= \langle a - \alpha e_1, a - \alpha e_1\rangle_2 \\ &= \|a\|_2^2 - 2\alpha a_1 + \alpha^2 \\ &= -2\alpha(a_1 - \alpha) \\ \implies Qa = a - 2\frac{\langle v, a\rangle_2}{\|v\|_2^2} &= a + \frac{\langle v, a\rangle_2}{\alpha(a_1 - \alpha)} v \end{aligned}$$

Ende von Vorlesung 03 am 18.10.2022

Lemma 2.18. *Sie $\alpha \in \mathbb{R}^n, a \neq 0, a \notin span\{e_1\}$. Sei*

$$v = a - \alpha e_1, \alpha = -sign(a_1) \cdot \|a\|_2 \tag{2.4}$$

Dann ist

$$\left(I - 2\frac{vv^t}{v^t v}\right) a = a + \frac{v^t a}{\alpha(a_1 - \alpha)} v = \alpha e_1. \tag{2.5}$$

Beweis. Siehe oben. □

Algorithm 2.19

Input: $A \in \mathbb{R}^{m\times n}, m \geq n$ "Mehr Zeilen als Spalten"
Output: $A \in \mathbb{R}^{m\times n}$, obere rechte Dreiecksmatrix R, Rest Householder-Transformationen

for $j = 1, \ldots, n$ **do** ▷ Iterieren über die Spalten
  Berechne v, α wie in (2.4) ,mit $a = A[j : m, j] \in \mathbb{R}^{m-j+1}$
  $v = \frac{1}{v_1} v$ ▷ Erster Eintrag wird nicht gespeichert, daher normalisieren wir
  Berechne $A[j : m, j : n] = \left(I - \frac{vv^t}{v^t v}\right) A[j : m, j : n]$ wie in (2.5)
  if $j < m$ **then**
    A[j+1:m,j]=v[2:m-j+1] ▷ Index startet von 1
  end if
end for

Bemerkung 2.20. *Die Skalierung $v = \frac{1}{v_1} v$ stellt sicher, dass die der erste Eintrag von v nicht gespeichert werden muss.*

Aufwand: $m \sim n \rightsquigarrow \frac{2}{3} n^3$ Multiplikationen
$m \gg n \rightsquigarrow 2n^2 m$ Multiplikationen
Schneller als Givensrotationen, stabiler als Normalengleichungen

2.6 Pseudoinverse

Ausgangspunkt: Wir wollen ein stabiles numerisches Verfahren, dass

$$Ax = b, A \in \mathbb{R}^{m\times n}, m \geq n, \operatorname{rang}(A) = n, b \in \mathbb{R}^n$$

"lösen" kann, d.h. es gilt

$$\|Ax - b\|_2 = \min_{y\in\mathbb{R}^n} \|Ay - b\|_2$$

Mathematisch können wir die Abbildung $b \mapsto x$, wegen der Normalengleichung (2.2), schreiben als

$$x = \underbrace{(A^tA)^{-1}A^t}_{:=A^\dagger} b = A^\dagger b$$

$A^\dagger \in \mathbb{R}^{n\times m}$. Wegen $A^\dagger A = I$ heißt $A^\dagger$ auch **Pseudoinverse**.
Frage: Können wir den Begriff der Inversen noch weiter verallgemeinern? Auf beliebige Matrizen?
Satz 1.9: $A \in \mathbb{R}^{m\times n}, U = \mathrm{Bild}(A)$

$$\begin{aligned}
&\implies \|Ax-b\|_2 = \min_{y\in\mathbb{R}^n} \|Ay-b\|_2 \overset{\text{Satz 1.9}}{\iff} Ax-b \in \mathrm{Bild}(A)^\perp \\
&\iff Ax - Pb - \underbrace{(b-Pb)}_{\in U^\perp\text{: Satz 1.9}} \in \mathrm{Bild}(A)^\perp, Pb \text{ ist die orthogonale Projektion von } b \text{ auf } U \\
&\iff \underbrace{\underbrace{Ax}_{\in U} - \underbrace{Pb}_{\in U}}_{\in U} \in \mathrm{Bild}(A)^\perp \\
&\iff Ax = Pb
\end{aligned}$$

Falls $\mathrm{rang}(A) < n$ (z.B., falls $m < n$) ist $Ax = Pb$ nicht eindeutig lösbar (aber es existiert immer eine Lösung).
Für $\tilde{x} \in \mathbb{R}^n$ mit $A\tilde{x} = Pb, x' \in \ker(A)$ ist $A(\tilde{x} + x') = Pb$.

$$\begin{aligned}
L(b) &= \left\{ x \in \mathbb{R}^n : \|Ax-b\|_2 = \min_{y\in\mathbb{R}^n} \|Ay-b\|_2 \right\} \\
&= \{x \in \mathbb{R}^n : Ax = Pb\} \\
&= \tilde{x} + \ker(A)
\end{aligned}$$

Sind gewisse Lösungen sinnvoller als andere?
Wähle: $x \in \tilde{x} + \ker(A)$ mit minimaler Norm als "eindeutige" Lösung von $Ax = b$.

$$\begin{aligned}
\overset{\text{Bem. 1.13}}{\implies} \|x-0\|_2 = \min_{y\in\tilde{x}+\ker(A)} \|y-0\|_2 &\iff x \in (\tilde{x} + \ker(A))^\perp \\
&\iff x \in \ker(A)^\perp
\end{aligned}$$

Anmerkung

Hier ist nicht ganz klar, was mit $(\tilde{x} + \ker(A))^\perp$ gemeint ist, da dies z.B. für $\ker(A) = \mathrm{span}\{(0,1)^t\}$ und $\tilde{x} = (1,0)^t$ nur $\{0\}$ ist, was natürlich nicht der Intuition entspricht!
Statt der ursprünglichen Definition müssen wir hier wieder zurück schieben ($-\tilde{x}$ rechnen), was kein Problem ist, da wir o.B.d.A $\tilde{x} \perp \ker(A)$ vorraussetzen dürfen, bevor wir das Skalarprodukt berechnen!
Zum Beispiel ist also $v = (1,0)^t$ im obigen Beipspiel doch im orthogonalen Komplement, da $\langle v, \tilde{x}+u-\tilde{x}\rangle_2 = 0$ für $u \in \ker(A)$

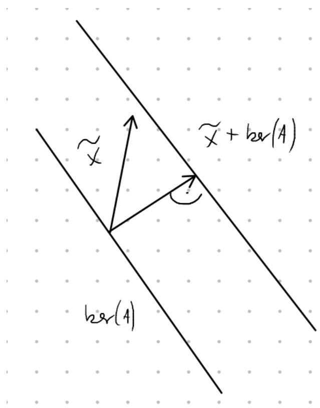

Abbildung 2.9: Setting

Bemerkung 2.21. *Diese Wahl von x für $b \mapsto x$ ist linear: Für $b_1, b_2 \in \mathbb{R}^m$ ist:*

$$\left.\begin{array}{ll} Ax_1 = b_1 & x_1 \in ker(A)^\perp \\ Ax_2 = b_2 & x_2 \in ker(A)^\perp \end{array}\right\} \implies P(x_1 + x_2) = P(x_1) + P(x_2) = Ax_1 + Ax_2 = A(x_1 + x_2), x_1 + x_2 \in \ker(A)^\perp$$

Definition 2.22. *Sei $A \in \mathbb{R}^{m\times n}$. Die Abbildungsmatrix $A^\dagger \in \mathbb{R}^{n\times m}$ von $b \mapsto x$ heißt **Pseudoinverse** oder **Moore-Pensore-Inverse** von A. D.h. gegeben $b \in \mathbb{R}^n$, dann ist $x = A^\dagger b$ die eindeutige Lösung von*

$$\min_{y\in\ker(A)^\perp} \|Ay - b\|_2 = \|Ax - b\|_2 \,.$$

Satz 2.23. *$A \in \mathbb{R}^{m\times n}$. Dann ist $A^\dagger \in \mathbb{R}^{m\times n}$ eindeutig über die Moore-Penrose-Axiome definiert:*

1. $(A^\dagger A)^t = AA^\dagger$
2. $(AA^\dagger)^t = A^\dagger A$
3. $A^\dagger AA^\dagger = A^\dagger$
4. $AA^\dagger A = A$

Beweis. Siehe Literatur oder später □

Frage: Wie berechnen wir $x = A^\dagger b$?
Sei $A \in \mathbb{R}^{m\times n}, \operatorname{rang}(A) = p \leq \min(m, n)$. Bringe A mittels orthogonaler Transformationen (z.B. Householder) auf obere Dreiecksgestalt, d.h.:

$$Q^t A = \begin{bmatrix} & R & S \\ * & & \\ & 0 & 0 \end{bmatrix} \tag{2.6}$$

wobei $S \in \mathbb{R}^{p\times(n-p)}$. Setze Analog $x = \begin{bmatrix} x_1 \in \mathbb{R}^p \\ x_2 \in R^{n-p} \end{bmatrix}, Q^t b = \begin{bmatrix} b_1 \in \mathbb{R}^p \\ b_2 \in \mathbb{R}^{m-p} \end{bmatrix}$

Lemma 2.24. *Mit obigen Bezeichungen ist $x = A^\dagger b$ genau dann, wenn*

$$x_1 = R^{-1}b_1 - R^{-1}Sx_2.$$

Beweis.

$$\begin{aligned}\|Ax-b\|_2^2 &= \left\|Q^t(Ax-b)\right\|_2^2 \\ &= \left\|\begin{pmatrix} Rx_1+Sx_2-b \\ -b_2 \end{pmatrix}\right\|_2^2 \\ &= \|Rx_1+Sx_2-b_1\|_2^2 + \|b_2\|_2^2\end{aligned}$$

ist minimal, falls $Rx_1 = b_1 - Sx_2$. □

Wir sehen $p = \text{rang}(A) = n \implies$ wie vorher, lineares Ausgleichsproblem!
Sonst: $x_2 = ?$

Lemma 2.25. *Sei $p < n, V = R^-1S \in \mathbb{R}^{n\times(n-p)}$ und $u = R^{-1}b_1 \in \mathbb{R}^p$. Dann ist*

$$\begin{aligned} & x = A^\dagger b \\ \iff & (I+V^tV)x_2 = V^tu \\ & x_1 = u - Vx_2\end{aligned}$$

Beweis.

$$\begin{aligned}\|x\|_2^2 &= \|x_1\|_2^2 + \|x_2\|_2^2 \\ &\overset{\text{Lemma 2.24}}{=} \|u-Vx_2\|_2^2 + \|x_2\|_2^2 \\ &= \|u\|_2^2 - 2\langle u, Vx_2\rangle_2 + \langle Vx_2, Vx_2\rangle_2 + \langle x_2, x_2\rangle_2 \\ &= \|u\|_2^2 + \langle x_2, (I+V^tV)x_2 - 2V^tu\rangle_2 = \varphi(x_2)\end{aligned}$$

Minimiere $\varphi(x_2)$:

$$\begin{aligned}\varphi'(x_2) &= -2V^tu + 2(I+V^tV)x_2 \\ \varphi'(x_2) &= 2(I+V^tV) \implies \text{spd}\end{aligned}$$

φ minimal $\iff \varphi'(x_2) = 0 \implies$ Behauptung. □

Algorithm 2.26

Input: $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$
Output: $x = A^\dagger b$

Berechne QR-Zerlegung (2.6) von A
$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = Q^tb$
$V = R^{-1}S$ mittels Rückwertssubstitution
$u = R^{-1}b_1$ mittels Rückwertssubstitution
Löse $(I+V^tV)x_2 = V^tu$ mittels Cholesky-Zerlegung
$x_1 = u - Vx_2$
$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Ende von Vorlesung 04 am 20.10.2022

Kapitel 3

Iterative Verfahren für große, dünn besetzte, Gleichungssysteme

3.1 Motivation

Sei $\Omega \subset \mathbb{R}^d, d \in \mathbb{N}$. Betrachte die stationäre Wärmeleitungsgleichung, eine partielle Differenzialgleichung

$$\begin{cases} -\Delta u(x) = f(x) & \in \Omega \\ u(x) = 0 & x \in \delta\Omega \end{cases} \tag{3.1}$$

mit Wärmequelle $f \in C(\Omega)$ und dem Laplace-Operator:

$$\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u(x)}{\partial x_i^2}. \tag{3.2}$$

Die Lösung $u \in C^2(\Omega)$, falls existent, beschreit die Temperaturverteilung im Raum Ω.
Diese Gleichung ist i.A. nicht von Hand lösbar!
<u>Idee:</u> Berechne approximative Lösung im Computer.
<u>Ansatz:</u> Für $g \in C^2(\mathbb{R})$ ist

$$\begin{aligned} g''(x) = \lim_{h \searrow 0} \frac{g'(x+h) - g(x)}{h} &\approx \frac{g'(x+h) - g(x)}{h} \\ &\approx \frac{\frac{g(x+h)-g(x)}{h} - \frac{g(x)-g(x-h)}{h}}{h} \\ &\approx \frac{g(x+h) - 2g(x) + g(x-h)}{h^2} \end{aligned}$$

$\rightsquigarrow$ Ersetze $\frac{\partial^2 u}{\partial x_i^2}$ in (3.2)

$\rightsquigarrow$ Überziehe Ω mit einem regelmäßigen Gitter mit Maschenweite $h = \frac{1}{n}, n \in \mathbb{N}$.

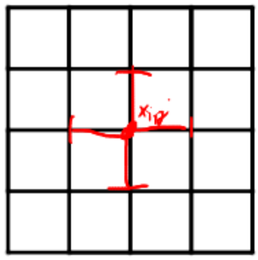

Abbildung 3.1: Gitter

Bezeichne die Gitterpunkte mit x_{ij} und $u_{ij} = u(x_{ij})$.

$$\overset{d=2}{\Longrightarrow} \frac{1}{h^2}\left(4u_{ij} - u_{i+1j} - u_{i-1j} - u_{ij+1} - u_{ij-1}\right) = f_{ij} : i,j \in 1,\dots,n-1$$

$$u_{ij} = 0, i \in \{0,n\} \text{ oder } j \in \{0,n\}$$

Wir erhalten ein lineares Gleichungssystem mit $N = (n-1)^2$ Unbekannten und $O(1)$ Einträgen pro Zeile.
$\Longrightarrow$ $A \in \mathbb{R}^{n\times n}$ hat $O(N)$ Einträge. Wir haben das Lösen eines (linearen) partiellen Differentialgleichung durch das Lösen eines linearen Gleichungssystems ersetzt.

Beispiel 3.1. $\Omega = (0,1)^2, n = 4 \implies h = \frac{1}{4}$. *Erhalte:*

$$\left[\begin{array}{ccc|ccc|ccc}
4 & -1 & & -1 & & & & & \\
-1 & 4 & -1 & & -1 & & & & \\
 & -1 & 4 & & & -1 & & & \\
\hline
-1 & & & 4 & -1 & & -1 & & \\
 & -1 & & -1 & 4 & -1 & & -1 & \\
 & & -1 & & -1 & 4 & & & -1 \\
\hline
 & & & -1 & & & 4 & -1 & \\
 & & & & -1 & & -1 & 4 & -1 \\
 & & & & & -1 & & -1 & 4
\end{array}\right]
\begin{bmatrix} u_{11}\\ u_{12}\\ u_{13}\\ u_{21}\\ u_{22}\\ u_{23}\\ u_{31}\\ u_{32}\\ u_{33} \end{bmatrix}
=
\begin{bmatrix} f_{11}\\ f_{12}\\ f_{13}\\ f_{21}\\ f_{22}\\ f_{23}\\ f_{31}\\ f_{32}\\ f_{33} \end{bmatrix}$$

Aber: Um die Lösung von (3.1) gut zu approximierenm ist oft $N \gg 1$ erforderlich. Für kleine bis mittlere N, d.h. in 2022 je nach Modell $\sim$ 10 Millionen, sind graphenbasierte Löser eine Option.
Was tun für große N?
Beobachtung: Matrix-Vektor-Multiplikation sind für dünn besetze Matrizen in $O(N)$ berechenbar.
Frage: Wie bauen wir gute Löser für LGS (lineare Gleichungssysteme) nur unter Anwendung von Matrix-Vektor-Multiplikationen?
Idee: Benutze Orthogonalität um eine Bestapproximationseigenschaft zu erhalten.

3.2 Grundidee von Projektionsmethoden

Sei $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n$ und K, L Unterräume vom $\mathbb{R}^n$.
Idee: Finde eine approximative Lösung $\tilde{x}$ zu $Ax = b$ mit

$$\tilde{x} \in K \text{ und } b - A\tilde{x} \perp_2 L$$

Kanonische Wahl: $L = AK$.
Falls wir eine Startnährung x_0 zu x kennen, können wir $\tilde{x}$ in $x_0 + K$ suchen:
Finde $\tilde{x} \in x_0 + K$ mit $b - A\tilde{x} \perp_2 L$
Beobachtung: $\tilde{x} \in x_0 + K \implies \exists d \in K : \tilde{x} = x_0 + d$

$$\implies \underbrace{b - A(x_0}_{r} + d) \perp_2 L$$

$$\iff r_0 - Ad \perp_2 L$$

Eine approximative Lösung $\tilde{x} = x_0 + d$ muss also erfüllen:

$$\begin{cases} \tilde{x} = x_0 + d \\ \langle r_0 - Ad, w \rangle_2 = 0 \quad \forall w \in L \end{cases} \tag{3.3}$$

<u>**Idee:**</u> Wähle x_0, K, L, berechne $d \in K$ durch Lösen eines Unterproblems. Setze $x_1 = x_0 + d$, wähle neue Unterräume, beginne von vorne.
Wie implementieren wir diese Idee im Computer?
Sei $K = \operatorname{span}\{v_1, \dots, v_n\}, L = \{w_1, \dots, w_n\}$
$V = [v_1 | \dots | v_n]$ und $W = [w_1 | \dots | w_n]$
(3.3) ist äquivalent zu

$$\begin{cases} \tilde{x} = x_0 + Vy & y \in \mathbb{R}^m \\ W_i^t AVy = W_i^t r_0 & i = 1, \dots n \end{cases} \iff \underbrace{W^t A}_{m \times m} Vy = W^t r_0 \tag{3.4}$$

$$\implies \tilde{x} = x_0 + V(W^t AV)^{-1} W^t r_0$$

Algorithm 3.2 Prototyp einer interativen Projektionsmethode

Input: $A \in \mathbb{R}^{n \times n}, b \in \mathbb{R}^n$, Fehlertoleranz α
Output: Nährung $x_{i+1} \approx x$

```
i=0
while Fehlertoleranz noch nicht erreicht do
    Wähle K_i, L_i
    Wähle Basen V, W von K_i, L_i
    r_1 = A x_i
    y = (W^t A V)^{-1} W^t r_i
    x_{i+1} = x_i + V y_i
    i = i + 1
end while
```

<u>**Aber:**</u> $W^t AV$ ist nicht notwendigerweise invertierbar:

Beispiel 3.3.

$$A = \left[\begin{array}{c|c} 0 & I \\ \hline I & I \end{array}\right] \in \mathbb{R}^{2m \times 2m}$$

$$K = L = span\{e_1, \dots, e_m\} \implies V = W = \left[\begin{array}{c} I_m \\ \hline 0 \end{array}\right] \in K^{2m \times m}$$

$$\implies W^t AV = 0 \textit{ ist nicht invertierbar.}$$

Lemma 3.4. *Sei einer der folgenden Bedingungen erfüllt:*

1. *A ist spd, $K = L$*
2. *A invertierbar, $L = AK$*

Dann ist $W^t AV$ für alle Basen von K, L invertierbar.

Beweis. **1.:** $L = K \implies W = V\delta$ mit $\delta \in R^{m\times m}$ invertierbar.
$\implies B = W^t AV = \delta^t V^t AV$

$$0 < \underbrace{\underbrace{y^t Ay}_{= x^t V^t AVx}}_{\text{spd, invertierbar}} \quad , y = Vx$$

2.: $L = AK \implies W = AV\delta, \delta \in \mathbb{R}^{m\times m}$ invertierbar

$$\implies B = W^t AV = \delta^t \underbrace{V^t A^t AV}_{\text{spd}} \implies \text{ invertierbar } \implies \text{ Beh.}$$

□

Ende von Vorlesung 05 am 25.10.2022