Einführung in die Grundlagen der Numerik

Vorlesungsmitschriften im Wintersemester 2018/19

INHALTSVERZEICHNIS

VORWORT

Diese Vorlesungsmitschriften werden in der Vorlesung *Einführung in die Grundlagen der Numerik* von Prof. Ira Neitzel im Wintersemester 2018/19 an der Universität Bonn angefertigt.

Wir versuchen, diese immer unter `https://pankratius.github.io` zu aktualisieren.

Teile, die von der Vorlesung abweichen, sind in violett markiert.

Betrachte $A \in \mathrm{GL}_n(\mathbb{R})$, wobei A schlecht konditioniert sein kann. Wir wollen ein Gleichungssystem der Form $Ax = b$, mit $b \in \mathbb{R}^n$ gegeben, lösen. Dazu suchen wir eine Orthogonalmatrix $Q \in \mathrm{O}_n(\mathbb{R})$ und eine obere Dreiecksmatrix $R \in \mathrm{M}_n(\mathbb{R})$ mit $A = QR$. Diese Zerlegung von A nennt man **Orthogonalzerlegung**. Dann erhalten wir das äquivalente Problem

$$Ax = b \iff QRx = b \iff Rx = Q^T b.$$

1.1 Eigenschaften orthogonaler Matrizen

Lemma 1.1.1. *Sei $Q \in \mathrm{O}_m(\mathbb{R})$ orthogonal. Dann ist auch Q^T orthogonal und es gilt*

$$||Qx||_2 = \left|\left|Q^T x\right|\right|_2 = ||x||_2$$

Beweis. Es gilt

$$||Qx||_2^2 = x^T Q^T Q x = x^T x = ||x||_2 .$$

Genauso für Q^T. □

Lemma 1.1.2. *Sei $A \in \mathrm{GL}_n(\mathbb{R})$ regulär und $Q \in \mathrm{O}_n(\mathbb{R})$ orthogonal. Dann gilt*

$$\kappa_2(QA) = \kappa_2(A)$$

Beweis. Die Matrixnorm $||A||_2$ ist durch die euklidsche Norm induziert, i.e.

$$||A||_2 = \max_{x \neq 0} \frac{||Ax||_2}{||x||_2}.$$

Also folgt aus lemma 1.1.1, dass $||(||_2\, QA) = ||(||_2\, A)$ gilt.
Betrachte jetzt

$$\left|\left|A^{-1}Q^T\right|\right|_2 = \max_{x \neq 0} \frac{\left|\left|A^{-1}Q^T x\right|\right|_2}{||x||_2} \quad = \max_{x \neq 0} \frac{\left|\left|A^{-1}Q^T x\right|\right|_2}{||Q^T x||_2} \overset{y := Q^T x}{=} \max_{y \neq 0} \frac{\left|\left|A^{-1}\right|\right|_2}{||y||_2} = \left|\left|A^{-1}\right|\right|_2$$

□

Also ist für das LGS $Rx = Q^T b : \kappa_2(R) = \kappa_2(A)$. Also hat sich die Kondition des Problems nicht verschlechtert.

1.2 Anwendung: Lineare Ausgleichsgeraden

Betrachte für gegebenes $b \in \mathbb{R}^n$ und $A \in \mathrm{M}_n(\mathbb{R})$ das Optimierungsproblem

$$\min_{x \in \mathbb{R}^n} ||Ax - b||_2 . \tag{O}$$

Dieses Problem ist äquivalent zur Optimierung von $||Ax-b||_2^2$.

Seien nun m Tupel $(y_i, f_i) \in \mathbb{R}^2$ $(1 \leq i \leq m)$ gegeben. Gesucht ist diejenige affine Gerade $c + dy$ in $\mathbb{R}^2$, so dass die Summe der Quadrate der Punkte von der Gerade minimal ist. Wir erhalten also das Optimierungsproblem

$$\min_{(c,d)\in\mathbb{R}^2} \left(\sum_{i=1}^{m} (c + dy_i - f_i)^2 \right) = \min_{(c,d)\in\mathbb{R}^2} \left\| \begin{pmatrix} 1 & y_1 \\ \vdots & \vdots \\ 1 & y_m \end{pmatrix} \cdot \begin{pmatrix} c \\ d \end{pmatrix} - \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \right\|_2 .$$

Betrachte allgemeiner das Polynom

$$p(y) = \sum_{k=0}^{n-1} a_k y^k .$$

Gesucht sind jetzt die Koeffizienten $a_0, ..., a_{n-1}$ mit

$$\sum_{j=1}^{m} \left(p(y_j) - f_j\right)^2$$

ist minimal. Schreibe dies ebenfalls als Optimierungsproblem:

$$\min_{a_0,\ldots,a_{n-1}} \left\| \begin{pmatrix} y_1^0 & \cdots & y_1^{n-1} \\ \vdots & \ddots & \vdots \\ y_m^0 & \cdots & y_m \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix} - \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \right\|_2^2 .$$

End of Lecture 1

Die Existenz der Lösung des Optimierungsproblems folgt aus

$$\lim_{x\to\infty} ||Ax - b||_2 \to \infty,$$

und einer anschließenden Anwendung des Satzes von Weierstraß auf die kompakten Niveaumengen der Abbildung.

Theorem 1.2.1. *(Weierstraß) Sei X ein kompakter metrischer Raum und $f : X \to \mathbb{R}$ eine stetige Abbildung. Dann nimmt f auf X sowohl ein Maximum als auch ein Minimum an.*

Sei $f : X \to \mathbb{R}$, mit $X \subseteq \mathbb{R}$. Dann heißt $x_0 \in X$ ein **lokales Maximum** bzw. **lokales Minimum**, falls es eine Umgebung $x \in V$ gibt, so dass $f(x) \leq f(x_0)$ bzw. $f(x) \geq f(x_0)$ für alle $x \in V$ gilt.

Lemma 1.2.2. *Sei $U \subset \mathbb{R}^n$ offen, und $f : U \to \mathbb{R}$. Angenommen, f hat in $x_0 \in U$ eine Extremstelle und ist in x_0 partiell differenzierbar. Dann gilt*

$$\nabla f(a) = 0.$$

Beweis. Betrachte für ein hinreichend kleines $\varepsilon > 0$ und $1 \leq j \leq n$ die Funktion

$$F : (-\varepsilon, \varepsilon) \to \mathbb{R}, \ t \mapsto F(x_0 + te_j).$$

Dann hat F in $t = 0$ eine Extremstelle, und es gilt

$$0 = F'(0) = \partial_j f(x_0)$$

□

Betrachte zur Lösung des Optimierungsproblems nun immer die Funktion

$$f : \mathbb{R}^n \to \mathbb{R},\ x \mapsto \frac{1}{2} \left|\left|Ax - b\right|\right|_2^2.$$

Dann gilt für beliebiges $\overline{x} \in \mathbb{R}^n$

$$\begin{aligned} \langle \nabla f(\overline{x}), h\rangle &= \langle A\overline{x} - b, Ah\rangle \\ &= \langle A^T(A\overline{x} - b), h\rangle \end{aligned}$$

Also gilt für eine Extremstelle $\overline{x}$

$$\nabla f(\overline{x}) = A^T(A\overline{x} - b) \overset{!}{=} 0 \iff \boxed{A^T A\overline{x} = A^T b}. \tag{NE}$$

Zur Lösung des Optimierungsproblems müssen wir also ebenfalls ein Gleichungssystem lösen. Die Gleichung (NE) heißt **Normalengleichung**.

Für $A \in \mathbb{R}^{m,n}$ ist $B := AA^T$ symmetrisch. Weiterhin ist B positiv semi-definit, denn es gilt

$$\langle x, Bx\rangle = x^T Bx = x^T AA^T x = \langle A^T x, A^T x\rangle = \left|\left|A^T x\right|\right|_2 \geq 0.$$

Weiterhin impliziert dies, dass alle Eigenwerte von x größer gleich null sind. Sei dazu λ ein Eigenwert und x ein korrespondierender Eigenvektor von B,

$$0 \leq \langle Bx, x\rangle = \langle \lambda x, x\rangle = \lambda \left|\left|x\right|\right|_2^2.$$

Also ist B positiv semi-definit. Angenommen, A hat nun vollen Rang. Dann ist A^T injektiv. Sei x ein Eigenvektor von B zum Eigenwert 0. Dann gilt

$$0 = \langle x, Bx\rangle = \left|\left|A^T x\right|\right|_2^2 \implies A^T x = 0 \implies x = 0.$$

Also hat B nur positive Eigenwerte. Nach dem euklidschen Spektralsatz [Sch18, S. 20.25] ist B aber diagonalisierbar. Also ist B sogar invertierbar, und die Normalengleichung hat hier jeweils eine eindeutig bestimmte Lösung. Im folgende habe A also immer maximalen Rang. Weil AA^T symmetrisch positiv-definit ist, kann man (NE) mit der Choleskyzerlegung lösen. Es gilt aber $\kappa_2(AA^T) = \kappa_2(A)^2$,

Wie ist die Kondition von A im nicht-quadratischen Fall definiert?

weshalb weitere Lösungsverfahren betrachtet werden müssen.

Dazu definieren wir die **erweiterte Orthogonalzerlegung** von $A \in \mathbb{R}^{m,n}$, mit $m \geq n$ durch

$$A = QR,\ \text{mit } Q \in \mathrm{O}_m\,\mathbb{R} \text{ und } R = \left[\begin{array}{c} \hat{R} \\ \hline 0 \end{array}\right],$$

wobei $\hat{R} \in \mathbb{R}^{n,n}$ eine obere Dreiecksmatrix ist. R heißt in diesem Fall **erweiterte obere Dreiecksmatrix**
Angenommen, eine solche Zerlegung existiert. Dann gilt

$$\begin{aligned} ||Ax - b||_2^2 &= ||QRx - b||_2^2 \\ &= \left|\left|QRX - QQ^T b\right|\right|_2^2 \\ &= \left|\left|Q(Rx + Q^T b)\right|\right|_2 \\ &= \left|\left|\begin{bmatrix} \hat{R} \\ \hline 0 \end{bmatrix} x - \begin{pmatrix} y_1 \\ \hline y_2 \end{pmatrix}\right|\right|_2 \\ &= \left|\left|\hat{R}x' - y_1\right|\right|_2^2 + ||y_2||_2^2, \end{aligned}$$

für passend gewählte y_1, y_2. Weil y_2 aber fix ist, reicht es,

$$\left|\left|\hat{R}x - y_1\right|\right|_2^2$$

zu minimieren.

Theorem 1.2.3. *Sei $1 \leq n \leq m$ und $A \in \mathbb{R}^{m,n}$, mit $\deg A = n$. Angenommen, A hat eine erweiterte Orthogonalzerlegung. Sei*

$$Q^T =: \begin{pmatrix} y_1 \\ \hline y_2 \end{pmatrix},$$

mit $y_1 \in \mathbb{R}^n$ und $y_2 \in \mathbb{R}^{m-n}$. Dann sind äquivalent

i) $\overline{x} \in \mathbb{R}^n$ *löst das Optimierungsproblem* (O).

ii) $\hat{R}x = y_1$.

1.3 Gram-Schmidt-Verfahren

Wir wollen die Existenz einer Orthogonalzerlegung von A zeigen. Dazu verwenden wir [Sch18, S. 11.32]

Proposition 1.3.1. *(Gram-Schmidtsches Orthogonalisierungsverfahren) Sei V ein n-dimensionaler euklidischer Vektorraum und $(b_1, ..., b_n)$ eine geordnete Basis von V. Für $1 \leq i \leq n$ sei*

$$V_i := \mathrm{Lin}(b_1, ..., b_i),$$

die also eine Flagge

$$0 = V_0 \subset V_1 \subset ... \subset V_n = V$$

bilden. Dann gibt es eine geordnete Orthogonalbasis $(\hat{b}_n, ..., \hat{b}_n)$ von V, so dass

$$V_i = \mathrm{Lin}(\hat{b}_1, ..., \hat{b}_n) \tag{$*$}$$

gilt.

Theorem 1.3.2. *Für jede reguläre Matrix $A \in \mathrm{GL}_(\mathbb{R})$ existiert eine orthogonale Matrix $Q \in \mathrm{O}_n(\mathbb{R})$ und eine obere Dreiecksmatrix $R \in \mathrm{M}\,n$, so dass $A = QR$ gilt.*

Beweis. Aus (∗) folgt, dass für die Abbildung

$$g : V \to V,\ b_i \mapsto \hat{b}_i$$

oben die Koordinatenmatrix bezüglich B in oberer Dreiecksform sein muss. Weiterhin ist g per Definition ein Isomorphismus.
Betrachte nun den konkreten Fall für $A \in \mathrm{GL}_n(\mathbb{R})$. Dann bilden die Spalten von A^{-1} eine geordnete Basis von $\mathbb{R}^n$. Also gibt es eine obere Dreiecksmatrix g, so dass

$$gA^{-1} = Q,$$

wobei $Q \in \mathrm{O}_n(\mathbb{R})$ orthogonal ist. Dabei haben wir benutzt, dass eine Matrix Q genau dann orthogonal ist, wenn ihre Spalten eine Orthonormalbasis von $\mathbb{R}^n$ bilden [**schroer**]. Damit erhalten wir aber

$$A = Q^{-1}g,$$

und $Q^{-1} \in \mathrm{O}_n(\mathbb{R})$, weil die orthogonalen Matrizen eine Gruppe bilden. □

1.4 Householder-Transformationen

Betrachte für ein fixes $w \in \mathbb{R}^s$ mit $w^T w = 1$ die Abbildung

$$H : \mathbb{R}^s \to \mathbb{R}^s, x \mapsto x - 2ww^T x.$$

Die Abbildung H heißt **Householder-Spiegelung**.

Proposition 1.4.1. *Für ein solches H gilt*

i) $H^T = H$

ii) $H^2 = E_s$,

also ist H insbesondere eine orthogonale Matrix, $H \in \mathrm{O}_s(\mathbb{R})$.

Beweis. i) $H^T = (E_s - 2ww^T)^T = E_s^T - 2(w^T)(w^T) = E_s - 2ww^T = H$

ii) Es gilt $\mathbb{R}^s = \mathrm{Lin}(w) \perp (\mathrm{Lin}(w))^\perp$. Weil H linear ist, genügen die folgenden beiden Ergebnisse

$$H(w) = w - 2(ww^T)w = w - 2w(w^T w) = w - 2w = -w,$$

und für $v \perp w$

$$H(v) = v - 2(ww^T)v = v - 2w(w^T v) = v.$$

□

End of Lecture 2

A.0 Sheet 0

Definition A.0.1. Die **Frechet-Ableitung** bezeichnet die gewöhnliche totale Ableitung einer Funktion $\mathbb{R}^n \to \mathbb{R}$.

Definition A.0.2. Sei $A \in \mathrm{M}_n(\mathbb{K})$. Der **Spektralradius** von A ist definiert als

$$\rho(A) := \max\{|\lambda_1|, ..., |\lambda_n|\},$$

wobei $\lambda_1, ..., \lambda_n$ die (möglicherweise komplexen) Eigenwerte von A darstellen.

Proposition A.0.3. *Sei $M \in \mathrm{M}_n(\mathbb{R})$, und*

$$\varphi : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Mx + c.$$

Dann sind äquivalent:

i) Für den Spektralradius p gilt $p(M) < 1$.

ii) Die Fixpunktiteration

$$x_{k+1} := \varphi(x_k)$$

konvergiert für ein beliebiges $x_0 \in \mathbb{R}^n$.

Solution A.0.2. Es gilt für das in der Aufgabenstellung spezifizierte ψ:

$$\{\text{Eigenwerte von } \psi\} = \lambda + (1 - \lambda) \cdot \{\text{Eigenwerte von } \varphi\},$$

wobei jeweils nur der lineare Teil betrachtet wurde. Nutze nun proposition A.0.3.

Proposition A.0.4. *Das Jacobi-Verfahren für die Matrix*

$$A = D - L - R$$

konvergiert, falls für die Matrix

$$I_{Jac.} := D^{-1}(L + R)$$

der Spektralradius größer 1 ist. Es konvergiert nicht, wenn der Spektralradius kleiner 1 ist.

Definition A.0.5. Die **Newton-Iteration** ist gegeben durch

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}$$

[Sch18] Jan Schrøer. **Lineare Algebra**. Vorlesungsskript, 2018.