Universität Paderborn Sommersemester 2020
INSTITUT FÜR MATHEMATIK
Prof. Dr. Bernhard Krötz, Dr. Job Kuit

Analysis 2

7. Anleitung zum Selbststudium

(KW23)

Die Themen dieser Woche sind:

- Extremstellen von Funktionen mehrerer Veränderlicher [M, 10.8] oder [Fo2, §7 p. 92 - 100].
- Der Banachsche Fixpunktsatz mit Umkehrsatz für stetige Abbildungen [S2, §3.1].

Repetitoruim: Lineare Algebra

1. Eigenwerte. Für die Untersuchung einer Funktion mehrerer Veränderlicher auf Extremstellen benötigen wir den Satz über die Hauptachsentransformation für eine symmetrische reelle Matrix. Einige Dinge werden allerdings klarer, wenn wir den Kontext zu Beginn etwas erweitern und uns erst beliebigen $n \times n$-Matrizen $A \in M_{n\times n}(\mathbb{C})$ mit komplexen Einträgen widmen. Warum $\mathbb{C}$ und nicht $\mathbb{R}$ wird schnell klar werden und wir belassen es vorerst mit der platten, aber wichtigen Anmerkung, daß vermöge der Inklusion $M_{n\times n}(\mathbb{R}) \subset M_{n\times n}(\mathbb{C})$ eine reelle Matrix als komplexe Matrix aufgefasst werden kann.
Sei $A \in M_{n\times n}(\mathbb{C})$. Dann heißt $\lambda \in \mathbb{C}$ *Eigenwert von A*, falls ein Vektor $v \in \mathbb{C}^n \setminus \{0\}$ existiert mit

$$Av = \lambda v\,.$$

Man nennt v einen *Eigenvektor* zum Eigenwert λ. Bezeichnen wir mit $E \in M_{n\times n}(\mathbb{C})$ die Einheitsmatrix, so sehen wir sofort:

Lemma. *Folgende Aussagen sind äquivalent:*

1. *λ ist Eigenwert von A.*
2. $\ker(\lambda E - A) \neq \{0\}$.

Das *Spektrum von A* ist die Menge der Eigenwerte:

$$\operatorname{spec} A := \{\lambda \in \mathbb{C} \mid \lambda \text{ ist Eigenwert von } A\}\,.$$

Gleich werden wir sehen: $\operatorname{spec} A \neq \emptyset$ und $\#\operatorname{spec} A \leq n$. Wir definieren eine Funktion

$$\chi_A : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \det(zE - A)$$

und bemerken

$$\lambda \in \operatorname{spec} A \quad \Longleftrightarrow \quad \chi_A(\lambda) = 0.$$

Für eine Matrix $X \in M_{n\times n}(C)$ mit Einträgen x_{ij}, etwas formaler $X = \sum_{i,j} x_{ij} E_{ij}$ mit E_{ij} den Elementarmatrizen, erinnern wir an die Formel von Leibniz für die Determinante

$$\det X = \sum_{\sigma \in S_n} \epsilon(\sigma) x_{1\sigma(1)} \cdot \ldots \cdot x_{n\sigma(n)} \,. \tag{L}$$

Aus der Definition von χ_A ergibt sich direkt: χ_A ist ein normiertes Polynom n-ten Grades

$$\chi_A(z) = z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0$$

mit Koeffizienten $a_j = a_j(A) \in \mathbb{C}$, die polynomial von A abhängen[1]. Man nennt χ_A das *charakteristische Polynom von A*. Für den ersten und letzten Koeffizienten a_{n-1} und a_0 von χ_A gibt es einprägsame kurze Formeln:

$$a_{n-1} = -\operatorname{tr}(A) \quad \text{und} \quad a_0 = (-1)^n \det A \,.$$

Nach dem Fundamentalsatz der Algebra (hier kommt nun der algebraisch abgeschlossene Körper $\mathbb{C}$ zum Tragen) zerfällt $\chi_A(z)$ in Linearfaktoren

$$\chi_A(z) = (z - \lambda_1)^{m_1} \cdot \ldots \cdot (z - \lambda_k)^{m_k}$$

mit paarweise verschiedenen λ_i (den Nullstellen von χ_A ohne Vielfacheit) und $\sum_{j=1}^k m_j = n$. Insbesondere gilt

$$\operatorname{spec} A = \{\lambda_1, \ldots, \lambda_k\} \neq \varnothing$$

und $\# \operatorname{spec} A = k \leq n$.

Weiterführende Anmerkungen. *(a) (Konjugationsinvarianz) Ist $B \in \mathrm{GL}(n, \mathbb{C})$, dann gilt für alle $A \in M_{n\times n}(\mathbb{C})$:*

$$\chi_{BAB^{-1}} = \chi_A \,.$$

Diese Formel ist wichtig und ihr Beweis eine gute Selbstkontrolle für Ihr derzeitiges Verständnis. Können Sie

$$\begin{aligned} \chi_{BAB^{-1}}(z) &= \det(zE - BAB^{-1}) = \det(zBB^{-1} - BAB^{-1}) \\ &= \det(B(zE - A)B^{-1}) = \det B \det(zE - A) \det B^{-1} \\ &= \det(zE - B) = \chi_A(z) \end{aligned}$$

nachvollziehen?

(b) (Jordan-Normalform aus Sicht der Invariantentheorie) Wie üblich bezeichnen wir mit $\mathbb{C}[z]$ den Vektorraum der Polynome. Mit

$$\mathbb{Y}_n = \{p \in \mathbb{C}[z] \mid \deg = n,\ p \text{ ist normiert}\}$$

wird ein affiner Unterraum der Dimension n erklärt, welcher durch

$$\mathbb{C}^n \to \mathbb{Y}_n, \quad \mathbf{a} = \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix} \mapsto p_{\mathbf{a}}; \ p_{\mathbf{a}}(z) := z^n + a_{n-1} z^{n-1} + \ldots + a_0$$

---

[1]Damit meint man, daß die $a_j : M_{n\times n}(\mathbb{R}) \simeq \mathbb{R}^{n^2} \to \mathbb{R}$ Polynomfunktionen sind. Unter einer Polynomfunktion $P : \mathbb{R}^k \to \mathbb{R}$ versteht man eine Abbildung von der Form

$$P(x) = \sum_{\substack{\alpha \in \mathbb{N}_0^k \\ |\alpha| \leq N}} c_\alpha x^\alpha \qquad (x \in \mathbb{R}^k)$$

mit Koeffifizenten $c_\alpha \in \mathbb{R}$ und $N \in \mathbb{N}$. Die Zahl

$$\deg P := \{\max |\alpha| \mid c_\alpha \neq 0\}$$

nennt man den *Grad* von P (mit der Konvention $\deg 0 = -\infty$).

affin parametrisiert wird. Wir konzentrieren uns nun auf die Abbildung

$$\Phi : M_{n\times n}(\mathbb{C}) \to \mathbb{Y}_n, \quad A \mapsto \chi_A .$$

Die Abbildung ist polynomial, da die Koeffizienten $a_j(A)$ des charakteristischen Polynoms polynomial von A abhängen. Wir behaupten nun, daß Φ surjektiv ist. Um dies einzusehen, ist eine gänzlich andere Sichtweise von Polynomen nützlich: nicht über die Koeffizienten, sondern über seine Nullstellen wollen wir ein $p \in \mathbb{Y}_n$ ausdrücken. Ergo

$$p(z) = p_\Lambda(z) = (z - \lambda_1)(z - \lambda_2)\cdot \ldots \cdot (z - \lambda_n) .$$

mit $\Lambda = (\lambda_1, \ldots, \lambda_n)^T \in \mathbb{C}^n$ einem Nullstellenvektor. Man beachte, $p_\Lambda = p_{\sigma(\Lambda)}$ für jede Permutation $\sigma \in \mathbb{C}^n$ und damit resultiert eine Bijektion

$$\mathbb{C}^n/S_n \to \mathbb{Y}_n, \quad [\Lambda] = S_n \cdot \Lambda \to p_\Lambda ,$$

wobei $\mathbb{C}^n/S_n$ der Bahnenraum für die Standardwirkung $S_n \times \mathbb{C}^n \to \mathbb{C}^n$ ist. Zurück zur Surjektivität von Φ. Sei $p \in \mathbb{Y}_n$. Dann existiert ein $\Lambda \in \mathbb{C}^n$ mit $p = p_\Lambda$ und mit der Diagonalmatrix $D = D(\Lambda) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ erhalten wir:

$$\Phi(D)(z) = \det(zE - D) = p_\Lambda(z) .$$

Was sind nun die Fasern der stetigen, surjektiven polynomialen Abbildung

$$\Phi : M_{n\times n}(\mathbb{C}) \to \mathbb{Y}_n ,$$

i.e., wir fragen nach den Urbildern (Fasern) $\Phi^{-1}(p)$. Dazu setzen wir

$$\mathcal{O}_\Lambda := \{BD(\Lambda)B^{-1} \mid B \in \mathrm{GL}(n, \mathbb{C})\} \subset M_{n\times n}(\mathbb{C})$$

und notieren: $\mathcal{O}_\Lambda$ ist die $\mathrm{GL}(n, \mathbb{C})$-Bahn von $D(\Lambda)$ bezüglich der Wirkung von $\mathrm{GL}(n, \mathbb{C})$ auf $M_{n\times n}(\mathbb{C})$ unter Konjugation:

$$\mathrm{GL}(n, \mathbb{C}) \times M_{n\times n}(\mathbb{C}) \to M_{n\times n}(\mathbb{C}), \quad (B, A) \mapsto BAB^{-1} .$$

Der Satz über die Jordan-Normalform aus Sicht der Invariantentheorie lautet dann wie folgt:

1. *Ist Λ regulär, d.h. alle Einträge sind verschieden, so gilt $\Phi^{-1}(p_\Lambda) = \mathcal{O}_\Lambda$.*
2. *Für alle Λ ist $\mathcal{O}_\Lambda$ die eindeutige abgeschlossene $\mathrm{GL}(n, \mathbb{C})$-Bahn in der Faser $\Phi^{-1}(p_\Lambda)$.*

Wenn Sie nun wissen wollen, was hinter der Formel (Isomorphie kategorischer Quotienten)

$$M_{n\times n}(\mathbb{C})//\,\mathrm{GL}(n, \mathbb{C}) \simeq \mathbb{C}^n//S_n$$

steckt, so empfehle ich Ihnen bereits jetzt den Beginn mit Kraft's schöner Lektüre:

Hanspeter Kraft: Geometrische Methoden in der Invariantentheorie

2. Spektraltheorie Hermitescher Matrizen. Sei $A \in M_{n\times n}(\mathbb{C})$. Die komplex konjugierte Matrix $\bar{A}$ ist definiert durch die konjugierten Einträge, i.e. ist $A = (a_{ij})_{ij}$, dann $\bar{A} := (\bar{a}_{ij})_{ij}$. Wir halten fest: Für $A \in M_{n\times n}(\mathbb{C})$ gilt

$$A \in M_{n\times n}(\mathbb{R}) \iff \bar{A} = A .$$

Die zu A *adjungierte Matrix* wird erklärt als $A^* = \bar{A}^T$ und man nennt A *selbstadjungiert* oder *Hermitesch*, falls $A = A^*$. Wir dehnen das Standardskalrprodukt auf $\mathbb{R}^n$ sesquilinear auf $\mathbb{C}^n$ aus, bezeichnen dieses ebenso mit $\langle \cdot, \cdot \rangle$ und notieren

$$\langle v, w \rangle = \sum_{j=1}^{n} v_j \bar{w}_j \qquad (v, w \in \mathbb{C}^n)\,. \tag{H}$$

Der Nachweis des folgenden Lemmas ist eine gute Übung:

Lemma. *Sei $A \in M_{n\times n}(\mathbb{C})$. Dann ist A^* eindeutig durch die Eigenschaft*

$$\langle Av, w \rangle = \langle v, A^* w \rangle \qquad (v, w \in \mathbb{C}^n)$$

festgelegt. Insbesondere ist A genau dann selbstadjungiert, wenn $\langle Av, w \rangle = \langle v, Aw \rangle$ für alle alle $v, w \in \mathbb{C}^n$ erfüllt ist. Deweiteren gelten die Rechenregeln:

$$(\lambda A)^* = \bar{\lambda} A^*, \quad (A+B)^* = A^* + B^*, \quad (AB)^* = B^* A^*, \quad A^{**} = A\,.$$

Zentral ist nun die kommende Proposition mit Beweis:

Proposition. *Sei $A \in M_{n\times n}(\mathbb{C})$ Hermitesch. Dann gilt:*

1. $\operatorname{spec} A \subset \mathbb{R}$.
2. *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal. Insbesondere sind Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig.*

Beweis. Sei $\lambda \in \operatorname{spec}(A)$ und $v \in \mathbb{C}^n$ ein zugehöriger normierter Eigenvektor. Dann:

$$\lambda = \lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle Av, v \rangle = \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle = \bar{\lambda}.$$

Das zeigt die erste Aussage. Seien nun $\lambda_1, \lambda_2 \in \operatorname{spec}(A)$ verschieden und v_1, v_2 entsprechend. Dann

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v, v_2 \rangle = \langle Av_1, v_2 \rangle = \langle v_1, Av_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

uns somit $\langle v_1, v_2 \rangle = 0$. □

Für $\lambda \in \operatorname{spec}(A)$ setzen wir

$$E(\lambda) := \ker(A - \lambda E) = \{v \in \mathbb{C}^b \mid Av = \lambda v\}$$

und nennen $E(\lambda)$ den *Eigenraum zum Eigenwert λ*.

Spektralsatz für Hermitesche Matrizen - 1. Version. *Sei A eine Hermitesche Matrix. Dann*

$$\mathbb{C}^n = \bigoplus_{\lambda \in \operatorname{spec} A} E(\lambda)\,,$$

und diese Zerlegung ist orthogonal, i.e. $E(\lambda) \perp E(\mu)$ für $\lambda \neq \mu$. Anders ausgedrückt: es existiert eine Orthonormalbasis von $\mathbb{C}^n$ bestehend aus Eigenvektoren für A.

Der Beweis ist einfach und läuft über vollständige Induktion nach der Dimension. Hier ist neuer Begriff nützlich: Unter einem *Hermiteschen Raum* $(V, \langle \cdot, \cdot \rangle)$ verstehen wir einen endlichdimensionalen komplexen Vektorraum V versehen mit einer Paarung

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}\,,$$

welche den folgenden Axiomen genügt:

- $\langle v, v\rangle \geq 0$ für alle $v \in V$ und $\langle v, v\rangle = 0$ genau dann, wenn $v = 0$.
- $\langle v, w\rangle = \overline{\langle w, v\rangle}$ für alle $v, w \in V$.
- $\langle \lambda v_1 + v_2, w\rangle = \lambda\langle v_1, w\rangle + \langle v_2, w\rangle$ für alle $v_1, v_2, w \in V$ und $\lambda \in \mathbb{C}$.

Das typische Beispiel eines Hermiteschen Raumes ist ein Unterraum $V \subset \mathbb{C}^n$, versehen mit der Einschränkung der (Hermiteschen) Standardform (H). Für ein $A \in \mathcal{L}(V, V)$ definiert man dann einen adjungierten Operator $A^* \in \mathcal{L}(V, V)$, eindeutig charakterisiert durch $\langle Av, w\rangle = \langle v, A^*w\rangle$ für alle $v, w \in V$. Weiter, den Spektralsatz für Hermitesche Matrizen fassen wir etwas allgemeiner für selbstadjungierte Operatoren in Hermiteschen Räumen auf. Nun zum Beweis. Sei A eine Hermitesche Matrix. Da $\operatorname{spec} A \neq \varnothing$, finden wir einen Eigenwert λ mit Eigenvektor v. Setze $W := \mathbb{C}v$ und definiere $V := W^\perp \subset \mathbb{C}^n$. Zur Erinnerung: $\mathbb{C}^n = W \oplus V$ ist eine orthogonale direkte Zerlegung. Weiter ist V, versehen mit der eingeschränkten Standardform (H), ein Hermitescher Raum und wir beachten $A(W) \subset W$. Um den Beweis mittels Induktion zu beenden, benötigen wir nur $A(V) \subset V$. Sei nun $v \in V$ fest und $w \in W$. Dann gilt

$$\langle Av, w\rangle = \langle v, Aw\rangle = \lambda\langle v, w\rangle = 0\,,$$

für jedes $w \in W$, woraus wir $A(V) \subset V$ folgern und Induktion auf den selbstadjungierten Operator $A|_V \in \mathcal{L}(V, V)$ anwenden können.

Wir definieren die unitäre und die orthogonale Gruppe via

$$\mathrm{U}(n) := \{A \in \mathrm{GL}(n, \mathbb{C}) \mid A^*A = E\}$$
$$\mathrm{O}(n) := \{A \in GL(n, \mathbb{R}) \mid A^TA = E\}\,.$$

Machen Sie sich bitte klar, daß dies in der Tat Gruppen definieren und $\mathrm{O}(n)$ eine Untergruppe von $\mathrm{U}(n)$ ist. Beachten Sie: Eine Matrix $U \in \mathrm{GL}(n, \mathbb{C})$ ist genau dann unitär, wenn ihre Spalten eine ONB bilden. Damit erhalten Sie die zweite Variante des Spektralsatzes:

Spektralsatz für Hermitesche Matrizen - 2. Version. *Sei A eine Hermitesche Matrix. Dann existiert eine unitäre Matrix $U \in \mathrm{U}(n)$ und eine Diagonalmatrix $D = \operatorname{diag}(d_1, \ldots, d_n) \in \operatorname{Diag}(n, \mathbb{R})$ mit <u>reellen</u> Einträgen, so daß*

$$UAU^{-1} = D\,.$$

Weiter gilt $\operatorname{spec} A = \{d_1, \ldots, d_n\}$.

Das Pendant für reelle Matrizen ist der

Satz über die Hauptachsentransformation. *Sei A eine symmetrische reelle Matrix. Dann existiert eine orthogonale Matrix $O \in \mathrm{O}(n)$ und eine Diagonalmatrix $D = \operatorname{diag}(d_1, \ldots, d_n) \in \operatorname{Diag}(n, \mathbb{R})$ mit*

$$OAO^{-1} = D\,.$$

Weiter gilt $\operatorname{spec} A = \{d_1, \ldots, d_n\}$.

Bemerkung. *Wie der Name schon sagt, verbirgt sich hinter diesem Satz sehr viel interessante Geometrie. Ich versuche mal einige Aspekte zu erklären. Neben den Geraden (direkt proportionale Zusammenhänge) sind die Parabeln (quadratische Zusammenhänge, z.B. der Luftwiderstand in Abhängigkeit von der Geschwindigkeit) wohl die wichtigsten Funktionen, die man in in der Schule studiert. Mit dem Scheitel im Ursprung reden wir also von Funktionen $f(x) = cx^2$. Was ist nun eine quadratische Funktion $Q : \mathbb{R}^n \to \mathbb{R}$ in mehreren Variablen? Darauf kann man vier (äquivalente) Antworten geben:*

1. *Q ist ein homogenes Polynom vom Grad* 2. [2]
2. *$Q(x) = \langle Ax, x\rangle$ für eine Matrix $A \in M_{n\times n}(\mathbb{R})$.*
3. *$Q(x) = \langle Ax, x\rangle$ für eine symmetrische Matrix $A \in M_{n\times n}(\mathbb{R})$.*
4. *$Q : \mathbb{R}^n \to \mathbb{R}$ ist homogen vom Grad 2 und die Abbildung $B(x, y) := Q(x + y) - Q(x) - Q(y)$ ist bilinear.*

Der Beweis ist nicht schwierig, doch über die Implikation 2. $\Rightarrow$ 3. will ich dennoch kurz sprechen. Eine Matrix A zerlegen wir zuerst in ihren symmetrischen und schiefsymmetrischen Anteil

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)\,.$$

Nun ist eine Matrix S schiefsymmetrisch (d.h. $S^T = -S$) genau dann wenn $\langle Sx, x\rangle = 0$ für alle $x \in \mathbb{R}^n$. Damit können Sie den Nachweis von 2. $\Rightarrow$ 3. nun selbst beenden. Sei nun $Q(x) = \langle Ax, x\rangle$ eine quadratische Form mit A symmetrisch. Nach dem Satz über die Hauptachsentranformation finden wir eine orthogonale Matrix (also eine abstandserhaltende und den Ursprung festlassende Bewegung im Raum) O mit $OAO^{-1} = D$ und einer Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Man beachte auch $O^{-1} = O^T$, da O eine orthogonale Matrix war. Ist $O = [v_1, \ldots, v_n]$, so ist mit $\mathcal{B} = \{v_1, \ldots, v_n\}$ eine ONB des $\mathbb{R}^n$ gegeben (unser neues Hauptachsensystem). Setze $y = [x]_{\mathcal{B}} = O^{-1}x$. Dann gilt

$$Q(x) = \langle ODO^{-1}x, x\rangle = \langle Dy, y\rangle = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2\,.$$

Wir sehen, dass Q in unseren neuem Koordinatensystem eine besonders schöne Gestalt hat. Sind zum Beispiel alle λ_i positiv, so beschreibt die Niveaumenge $Q^{-1}(1)$ das (im Raum verdrehte) Ellipsoid

$$E = O\{y \in \mathbb{R}^n \mid (y_1/a_1)^2 + \ldots + (y_n/a_n)^2 = 1\}$$

mit $a_i = \sqrt{\lambda_i}^{-1}$. Mehr entnehmen Sie bitte aus den Lehrbüchern oder aus der Werkbank von Soergel.

3. Positiv definite Matrizen. Eine selbstadjungierte Matrix $A \in M_{n\times n}(\mathbb{C})$ nennt man *positiv definit*, insofern $\langle Av, v\rangle > 0$ für alle $v \in \mathbb{C}^n \setminus \{0\}$ gilt.

Satz. *Sei $A = A^*$ eine selbstadjungierte Matrix. Dann sind äquivalent:*

1. *A ist positiv definit.*
2. $\mathrm{spec}\, A \subset \mathbb{R}_{>0}$.

Beweis. Nach dem Spektralsatz existiert eine ONB $\mathcal{B} = \{v_1, \ldots, v_n\}$ des $\mathbb{C}^n$ bestehend aus Eigenvektoren, i.e. $Av_i = \lambda_i v_i$ mit $\lambda_i \in \mathbb{R}$. Damit können wir jedes $v \in \mathbb{C}^n$ als Linearkombination $v = \sum_{j=1}^n c_j v_j$ schreiben. Wegen der Orthogonalität der v_i gilt

[2]Eine Funktion $Q : \mathbb{R}^n \to \mathbb{R}$ nennt man homogen vom Grad $d \in \mathbb{R}$, falls $Q(\lambda x) = \lambda^d Q(x)$ für alle $\lambda \in \mathbb{R}_{>0}$ und $x \in \mathbb{R}^n$. Z.B. ist

$$Q(x) = \begin{cases} \|x\|^d & \text{für } x \neq 0, \\ 0 & \text{sonst} \end{cases}$$

homogen vom Grad d für jede Norm $\|\cdot\|$ auf dem $\mathbb{R}^n$. Ist $d \in \mathbb{N}_0$ und Q im Zusatz ein Polynom, so ist Q genau dann homogen vom Grad d, falls $Q(x) = \sum_{|\alpha|=d} c_\alpha x^\alpha$ (Die Richtung "$\Rightarrow$" ist banal, aber für "$\Leftarrow$" muss man eine kleine Überlegung anstellen).

$$\langle Av, v\rangle = \sum_{i,j} c_i \bar{c}_j \langle Av_i, v_j\rangle = \sum_{i,j} c_i \bar{c}_j \lambda_i \langle v_i, v_j\rangle = \sum_{i=1}^{n} |c_i|^2 \lambda_i \,.$$

Den Rest des Beweises dürfen Sie selbst ausführen. □

Lokale Extrema einer Funktion mehrerer Veränderlicher

Gute Quellen sind [M, 10.8] oder [Fo2, §7 p. 92 - 100]; ich empfehle Ihnen den Abgleich beider. Ich begnüge mich hier nur mit etwas Heuristik und der Reduktion auf den eindimensionalen Fall. Aus der Schule (oder Analysis 1) wissen wir: Ist I ein offenes Intervall, $p \in I$, und hat eine differenzierbare Funktion $f : I \to \mathbb{R}$ in p ein lokales Extremum, so verschwindet die Ableitung; in Symbolen: $f'(p) = 0$. Ist darüberhinaus f zweimal differenzierbar und $f''(p) > 0$, so nimmt f in p ein lokales und isoliertes Minimum an. Das gilt es nun zu verallgemeinern.
Wieder von vorn: Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine differenzierbare Funktion, lokal extremal in $p \in U$. Gewählt sei $\epsilon > 0$, derart daß $B_\epsilon(p) \subset U$. Für eine Einheitstrichtung $v \in S^{n-1}$ betrachten wir sodann die Abbildung in einer Veränderlichen

$$g_v : (-\epsilon, \epsilon) \to \mathbb{R}, \quad t \mapsto f(p + tv)\,.$$

Da f in p lokal extremal ist, gilt dies notwendig auch für g_v bei $t = 0$ und wir schliessen:

$$0 = g_v'(0) = \partial_v f(p) = df(p)(v) = \langle \operatorname{grad} f(p), v\rangle\,.$$

Da dies für jede Einheitsrichtung $v \in S^{n-1}$ gelten muss, landen wir bei dem

Kriterium. *Hat eine differenzierbare Funktion $f : U \to \mathbb{R}$ in $p \in U$ ein lokales Extremum, so ist das Verschwinden der Ableitung in p zwingend:*

$$\operatorname{grad} f(p) = 0\,.$$

Im weiteren Verlauf nehmen wir an, daß f zweimal stetig differenzierbar ist und in p lokal extremal ist. Die Taylorapproximation von f in zweiter Ordnung schreibt sich als

$$f(p+v) = f(p) + \langle \operatorname{grad} f(p), v\rangle + \frac{1}{2}\langle \operatorname{Hess} f(p)v, v\rangle + R(v) \qquad (v \in B_\epsilon(0))$$

mit $\lim_{v\to 0} \frac{R(v)}{\|v\|^2} = 0$. Da $\operatorname{grad} f(p) = 0$, vereinfacht sich dies zu

$$f(p+v) = f(p) + \frac{1}{2}\langle \operatorname{Hess} f(p)v, v\rangle + R(v) \qquad (v \in B_\epsilon(0))\,. \tag{T}$$

Ebenso erinnern wir uns daran, daß $\operatorname{Hess} f(p)$ nach Schwarz eine relle symmetrische Matrix war.

Satz. *Sei $f \in C^2(u)$ und $p \in U$ ein kritischer Punkt, d.h.* $\operatorname{grad} f(p) = 0$. *Ist dann* $\operatorname{Hess} f(p)$ *positiv definit, so hat f in p ein isoliertes lokales Minimum.*

Beweis. Ist $\operatorname{Hess} f(p)$ positiv definit, so wissen wir aus dem Repetitorium, daß

$$\operatorname{spec} \operatorname{Hess} f(p) \subset \mathbb{R}_{>0}\,.$$

Sei $\lambda_{\min} > 0$ der kleinste Eigenwert. Schnelles Nachrechnen mit einer ONB aus Eigenvektoren bringt uns zu

$$\langle \operatorname{Hess} f(p)v, v\rangle \geq \lambda_{\min}\|v\|^2 \qquad (v \in \mathbb{R}^n)\,.$$

Wegen $\lim_{v\to 0} \frac{R(v)}{\|v\|^2} = 0$, gibt es ein $\delta > 0$ mit

$$\|R(v)\| < \frac{\lambda_{\min}}{2}\|v\|^2 \qquad (v \in B_\delta(0) \setminus \{0\})\,.$$

Mit (T) und der Δ-Ungleichung kommen wir zur somit zur finalen Abschätzung

$$f(p+v) > f(p) \qquad (v \in B_\delta(0) \setminus \{0\})\,,$$

d.h. f hat in p ein isoliertes lokales Minimum. □

Damit ist das Wesen erklärt. Den Rest über Extrema, isoliert oder nicht, nebst Sattelpunkten lesen Sie nun in [M] und [Fo2].

Der Banachsche Fixpunktsatz und der Umkehrsatz

Die Darstellung in [S2, §3.1] steht für sich. Erfreuen Sie sich an dieser Abhandlung.