Vorschläge zur Anfängerausbildung

Im Akademischen Jahr 99/00 habe ich in einer 6-stündigen Anfängervorlesung Analysis und Lineare Algebra gemeinsam behandelt. Zu der Frage, ob 6- oder 8-stündig, hatte es kontroverse Diskussionen gegeben, die rückblickend nicht mehr so wichtig sind, weil der begrenzende Faktor die bei vielen Studierenden ungenügende Auseinandersetzung mit den Hausaufgaben war. Ich habe auch in 6 Stunden mehr erzählt als nachgearbeitet worden ist, obwohl die Studierenden nicht den Eindruck hatten, ich stünde unter Zeitdruck.
Verschiedene Abweichungen von der Standardliteratur sind in dieser Vorlesung gut angekommen, so daß ich darüber berichten möchte.

In der Linearen Algebra handelt es sich um zwei Änderungen, die sich auf die Beweisstruktur der Theorie nicht auswirken, wohl aber auf die Koordination mit der Analysis und auf den Transfer des Gelernten. Ich will sehr nachdrücklich dafür plädieren, die Skalarprodukte **vor** den Determinaten zu behandeln, und ich will schildern, warum es günstig war, als häufigste Beispiele Polynomvektorräume auftreten zu lassen und nicht hauptsächlich die Standardräume K^n zum Üben zu nehmen.

In der Analysis bin ich einem Konzept gefolgt, daß in der Beweisstruktur (im ersten Semester) stark vom Standardaufbau abweicht. Ich plädiere für eine geänderte Reihenfolge der Hauptbegriffe, nämlich: **Differenzierbarkeit, Vollständigkeit, Stetigkeit.**
Auch hier bekommen die Polynome ein stärkeres Gewicht, und die Zeit, die für Aufgaben zur Differentialrechnung zur Verfügung steht, wird verlängert. Daß diese Reihenfolge überhaupt möglich ist, liegt im wesentlichen an zwei mathematischen Tatsachen:

1.) Der Monotoniesatz: “$f' \geq 0 \Rightarrow f$ schwach wachsend” (und die üblichen Folgerungen) läßt sich für Polynome schon beweisen, wenn man nur die rationalen Zahlen kennt. (Der Beweis des sonst verwendeten Mittelwertsatzes – über den Maximumsatz für stetige Funktionen – ist wesentlich verwickelter.)

2.) Die Eigenschaften von Grenzfunktionen, Differenzierbarkeit vor allem, können mit gleichmäßigen Fehlerabschätzungen ebensogut hergeleitet werden, wie mit gleichmäßiger Konvergenz. (Die Argumente sind sogar einfacher und kürzer.)

Für die Vorteile einer so geänderten Reihenfolge werde ich mit Beispielen argumentieren.

Lineare Algebra Vorschläge

Warum hilft es, Skalarprodukte vor Determinanten zu behandeln?

1.) Anwendungen der Skalarprodukte in der Analysis liegen weit vor Anwendungen der Determinaten in der Analysis.

2.) Die Analysis braucht Skalarprodukte früher als die Lineare Algebra die charakteristischen Polynome.

3.) Die Bedeutung der Funktionaldeterminante ist konzeptionell schwieriger als das charakteristische Polynom. Daher scheint es mir hilfreich, die Eigenschaften der Determinante (multilinear + alternierend) durch anschauliche Volumeneigenschaften (Cavalieri) nahe zu legen. Zumindest in Anfängervorlesungen müssen dazu Einheitswürfel schon bekannt sein.

4.) Die Normalformen der selbstadjungierten, der orthogonalen oder der schiefen Endomorphismen werden häufiger benötigt und können leichter hergeleitet werden als (auch Abschwächungen) der Jordanschen Normalform. Um diese spezielleren Normalformen auch eher behandeln (und länger üben) zu können, müssen Skalarprodukte vorhergehen.

5.) In Dimension 2 ist es leicht, die Gleichung $(a_{11} - \lambda) \cdot (a_{22} - \lambda) - a_{12} \cdot a_{21} = 0$ durch Gaußelimination als notwendige und hinreichende Bedingung für rang $(a - \lambda \mathrm{id}) < 2$ auftreten zu lassen.
Ein guter Tip für höhere Dimensionen: $a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$ ist der (signierte) Flächeninhalt des Parallelogramms aus den Spaltenvektoren.

6.) Die Dreiecksungleichung für die Norm $|x|^2 = \sum (x_j)^2$ führt auf Skalarprodukte.

7.) Da Linearformen, z. B. die Auswertungsabbildungen $P \mapsto P(a)$ auf dem Vektorraum der Polynome vom Grad $\leq d$ schon bekannt sind, hat man **viele** euklidische Normen, wähle z. B. $n \geq d+1$ Stellen $a_1, \ldots, a_n$ und definiere

$$\|P\|^2 := P(a_1)^2 + P(a_2)^2 + \ldots + P(a_n)^2,$$

die entsprechenden Skalarprodukte sind:

$$\langle P, Q \rangle := P(a_1) \cdot Q(a_1) + P(a_2) \cdot Q(a_2) + \ldots \ + P(a_n) \cdot Q(a_n).$$

Im Gegensatz dazu kann man Determinanten nicht direkt hinschreiben, sondern erst nach Kenntnis des Vorzeichens von Permutationen.

Warum sind Polynomvektorräume oft bessere Beispiele als die Standardvektorräume?

1.) Basiswechsel in K^n wirken auf viele Anfänger als Verschlechterung gegenüber der Standardbasis. Demgegenüber sind bei Polynomen die verschiedenen Taylorbasen

$$\{P_k(X) := (X - a)^k; \quad 0 \leq k \leq n\}$$

viel gleichberechtigter. Auch die Interpolationsbasen

$$\{P_k(X) := \prod_{j \in \{0, \ldots, n\} \setminus \{k\}} (X - a_j)/(a_k - a_j); \quad k = 0, \ldots, n\}$$

sind optimal für naheliegende Interpolationsprobleme, aber ungünstig zum Differenzieren.

2.)Man hat mit den Auswertungen $P \mapsto P(a)$ einen großen Vorrat an Linearformen, die nicht so leicht wie bei n-Tupeln mit Elementen des Vektorraums verwechselt werden können. Aber nicht alle interessanten Linearformen sind Auswertungen, vgl. etwa die numerische Integration: $P \mapsto (P(a) + P(b)) \cdot (b - a)/2$.

3.) Die grundlegende Aufgabe:

Zeige, daß die folgenden Elemente linear (un)abhängig sind!

kann natürlicher als bei K^n-Beispielen variiert werden. Zum Beispiel sind die Polynome P_k einer Interpolationsbasis linear unabhängig, weil man nur die Auswertungen $P \mapsto P(a_j)$ auf eine Linearkombination $\sum \lambda_k \cdot P_k = 0$ anwenden muß, um $\lambda_j = 0$ zu erhalten. Auch die Frage:

Sind $\{X \cdot (X-1) \cdot (X-2),\ X \cdot (X-1) \cdot (X-3),\ X \cdot (X-1) \cdot (X-4)\}$ *linear abhängig?*

zeigt den Nutzen des Einsatzes linearer Abbildungen bei derartigen Problemen.
Hier: $L_Q(P) := Q \cdot P, \quad (Q(X) := X \cdot (X-1)\ )$.

4.) Ein ganz sicherer Umgang mit Polynomen ist so wichtig, daß sie in Analysis und Linearer Algebra geübt werden sollten.

Analysis Konzept

I. Warum könnte man wünschen die Differentialrechnung vor der Stetigkeit zu beginnen?

I.1. Historisch war das 200 Jahre lang so.

I.2. In der Differentialrechnung ist einerseits ein Kalkül einzuüben, andererseits muß Verständnis dafür geweckt werden, wie Ableitungsinformationen zu guten Informationen über f führen und warum man das überhaupt will. Bei der Stetigkeit gibt es keinen Kalkül: Linearkombinationen, Produkte, Kompositionen stetiger Funktionen sind stetig – "*wie soll es denn sonst sein??*" Auch Hauptsätze wie Zwischenwertsatz und Maximumsatz sind **viel** leichter zu glauben und anzuwenden als zu beweisen. Kurz: In der Differentialrechnung muß viel mehr Technik geübt werden.

I.3. Wenn zu Beginn der Behandlung der Stetigkeit schon Erfahrungen mit Dreiecksungleichung und konvergenten Funktionenfolgen vorliegen, kommt man erstens bei der Behandlung der Stetigkeit schneller voran, und zweitens kann man die zu dem Begriff gehörenden Beispiele von Anfang an mit auftreten lassen.

II. Warum kann man die Differentialrechnung vor der Stetigkeit beginnen?

II.1. Die Polynome erlauben bezüglich der Stetigkeit eine explizite Behandlung:

Aus $x^n - a^n = (x-a) \cdot (x^{n-1} + \ldots + a^{n-1})$ und $x, a \in [-R,\ R]$ folgt:

$$|x^n - a^n| \ \leq nR^{n-1} \cdot |x - a|.$$

Daher kann man für Polynome aus den Funktionsdaten *auf jedem Intervall Dehnungsschranken ausrechnen*:

$$\text{Aus } P(X) := \sum_0^n a_k X^k \text{ und } x, a \in [-R,\ R] \text{ folgt:}$$
$$|P(x) - P(a)| \le \left(\sum |a_k| k \cdot R^{k-1}\right) \cdot |x - a|.$$

II.2. Nicht notwendig, aber aus meiner Sicht wünschenswert, ist die Beobachtung, daß die Graphen quadratischer Funktionen **zwischen** Kreisen und deren Tangenten verlaufen. Man kann also mit demselben Recht von Tangenten sprechen wie bei Kreisen.

II.3. Das Argument aus (II.1) kann wiederholt werden, um Polynome mit quadratischen Funktionen zu vergleichen:
Aus $|x^n - a^n - na^{n-1} \cdot (x - a)| = |(x - a) \cdot (x^{n-1} + \ldots + a^{n-1}) - (x - a) \cdot na^{n-1}|$
und $x, a \in [-R,\ R]$ folgt:
$$|x^n - a^n - na^{n-1} \cdot (x - a)| \le \tfrac{n(n-1)}{2} \cdot |x - a|^2.$$
und dies setzt sich wieder auf Polynome $P(X) := \sum a_k X^k$ und $x, a \in [-R,\ R]$ fort:

$$|Px) - P(a) - P'(a) \cdot (x - a)| \le \left(\sum |a_k| \frac{k(k-1)}{2} \cdot R^{k-2}\right) \cdot |x - a|^2.$$

Die Abweichung von der Tangentenapproximation $x \mapsto P(a) + P'(a) \cdot (x - a)$ wird also durch einen quadratischen Fehler kontrolliert, *der wieder auf jedem Intervall aus den Funktionsdaten berechenbar ist.*

II.4. Die üblichen Beweise der *Differentiationsregeln* haben folgende Eigenschaft: Verschärft man die Voraussetzungen – zum Beispiel wie in (II.3): “Die Abweichung von der Tangente sei gleichmäßig quadratisch kontrolliert” –, so ergeben sich auch die entsprechend verschärften Behauptungen.

II.5. Die *rationalen Funktionen* sind wegen der Kettenregel zugänglich, weil man $x \mapsto 1/x$ als weiteres Beispiel ohne Theorie behandeln kann:

$$0 < a/2 \le x \Rightarrow \frac{1}{x} - \frac{1}{a} + \frac{(x-a)}{a^2} = \frac{(x-a)^2}{x \cdot a^2} \ge 0 \text{ und } \le \frac{2}{a^3} \cdot (x-a)^2.$$

Soweit übt man Ungleichungen und lernt die Polynome kennen, aber

III. Kann man echte Analysissätze beweisen?

III.1. Der Monotoniesatz ist beweisbar, ohne die Vollständigkeit zitieren zu müssen:
Sei $f : [\alpha, \omega] \to \mathbb{R}$ gegeben mit **gleichmäßigen** Tangentenapproximationen (wie in allen bisherigen Beispielen), also:
$$a, b \in [\alpha, \omega] \Rightarrow |f(b) - f(a) - f'(a) \cdot (b - a)| \le K \cdot |b - a|^2$$
und sei zusätzlich vorausgesetzt: $0 \le f'$, dann gilt

$$a \le b \ \Rightarrow\ f(a) \le f(b) \text{ (Monotoniesatz)}$$

Denn: Unmittelbar nach Voraussetzung gilt:
$$f(a) \le f(b) - f'(a) \cdot (b - a) + K \cdot |b - a|^2 \le f(b) + K \cdot |b - a|^2.$$

Dies gilt auch für $a,\ m = (a+b)/2$ bzw. $\ m = (a+b)/2,\ b \in [\alpha, \omega]$, also haben wir für **beliebige** $a \le b \in [\alpha, \omega]$ folgende Verbesserung:

$$\begin{aligned} f(a) &\le f(m) + K \cdot (m-a)^2 \le f(b) + K \cdot (m-a)^2 + K \cdot (b-m)^2 \\ &\le f(b) + \frac{K}{2} \cdot |b-a|^2. \end{aligned}$$

Wiederholung liefert $f(a) \le f(b) + 2^{-n} \cdot K \cdot |b-a|^2$.
Das Archimedesaxiom liefert die Behauptung $f(a) \le f(b)$.

III.2. Man hat die sonst aus dem Mittelwertsatz gezogenen Folgerungen :

$$\begin{array}{rl} f' \le g' \text{ und } a \le b \Rightarrow f(b) - f(a) \le g(b) - g(a), & \text{(verallgemeinerter Monotoniesatz)} \\ |f'| \le L \Rightarrow |f(b) - f(a)| \le L \cdot |b-a| & \text{(Schrankensatz).} \\ 0 < f, g,\ f'/f < g'/g \Rightarrow g/f \text{ ist nicht fallend} & \text{(Multiplikativer Schrankensatz)} \end{array}$$

III.3. Man kann die Anwendung des Monotoniesatzes iterieren:

$$\begin{aligned} |f''| \le B \text{ und } \ a \le x \quad &\Rightarrow f'(a) - B \cdot (x-a) \le f'(x) \le f'(a) + B \cdot (x-a) \\ \Rightarrow f'(a) \cdot (x-a) - \frac{B}{2}(x-a)^2 \le \ f(x) - f(a) \ &\le f'(a) \cdot (x-a) + \frac{B}{2}(x-a)^2. \end{aligned}$$

III.4. Diese Ungleichungen sind drastische Verbesserungen der Fehlerschranken, die in (II.3) aus den Koeffizienten berechnet wurden. Beispiel: Für $0 \le x \le 1$ betrachte die Taylorpolynome $P_n(X) := \sum_{k=0}^{n} (-1)^k/k! \cdot X^k$. Diese bilden in $[0,1]$ eine Leibnizreihe, insbesondere also $|P_n''(x)| \le 1$ für alle n. Der vorhergehende Satz liefert also für $a, x \in [0,1]$ und **alle (!)** n:

$$|P_n(x) - P_n(a) - P_n'(a) \cdot (x-a)| \le \frac{1}{2}|x-a|^2.$$

III.5. Mit nur geringen Variationen (komplexer Schrankensatz) behandelt man die komplexen Polynome und bereichert Anschauung und Handwerkszeug erheblich.

IV. Was sind die nächsten wichtigen Ziele und wie erreicht man sie?

IV.1. Meiner Meinung nach ist es ein wesentliches erstes Ziel der Analysis, differenzierbare Funktionen zu konstruieren, die andersartige Eigenschaften haben, als man sie bei rationalen Funktionen findet. Beispiele: $f' = f,\ f'' + f = 0$.

IV.2. Nun ist Konvergenz und Vollständigkeit von $\mathbb{R}$ zu diskutieren. Meine einzige Änderung: Ich war hauptsächlich an Folgen von Funktionswerten $\{f_n(x)\}$ interessiert, der **Grenzfunktionen** wegen. Damit steht der Monotoniesatz als zusätzliches Hilfsmittel zur Verfügung (Argumentationsbeispiel: $f(x) := (1+x)^n - (1 + n \cdot x) \ge f(0) = 0$ (Bernoulli), solange $f''(x) = n(n-1)(1+x)^{n-2} \ge 0$ ist. Oder: $n \to f_n(x) := (1 + x/n)^n$ ist wachsend, weil $f_n'/f_n(x) = 1/(1+x/n) \le f_{n+1}'/f_{n+1}(x))$. – Leibnizreihen und Majorisierung durch die geometrische Reihe waren meine beiden anderen Werkzeuge.

IV.3. Die **gleichmäßige** Fehlerabschätzung aus (III.4) und das Archimedesaxiom liefern für die Grenzfunktion (hier ist speziell $P_n' = -P_{n-1}$):

$$x, a \in [0,1] \Rightarrow |P_\infty(x) - P_\infty(a) + P_\infty(a) \cdot (x-a)| \le \frac{1}{2} \cdot |x-a|^2.$$

Diese Ungleichung zeigt: P_∞ **ist differenzierbar**, $P_\infty' = -P_\infty$ und auch die neue Grenzfunktion wird von ihren Tangenten weiterhin mit quadratischem Fehler approximiert.

IV.4. Diese Argumentation habe ich auf **komplexe Potenzreihen** ausgedehnt.

IV.5. Mit dem Schrankensatz folgen Fehlerabschätzungen zwischen Stammfunktionen und Riemannsummen:

$$|F(b) - F(a) - \sum_{k=1}^{n} f(\tau_k) \cdot (t_k - t_{k-1})| \le \sup |f'| \cdot |b-a| \cdot \max |t - t_{k-1}|,$$

diese führen zur *Integralrechnung.*

IV.6. Alle bisherigen Funktionen f sind gleichmäßig durch stückweise lineare (Sehnen-) Funktionen s_n approximierbar. Diese s_n haben stückweise quadratische Stammfunktionen $S_n,\ S_n' = s_n,\ S_n(a) = 0$. Wegen des Schrankensatzes ist

$$|S_n(x) - S_m(x)| \le \sup |s_n(x) - s_m(x)| \cdot |b-a|.$$

Die $\{s_n(x)\}$ waren als gleichmäßige Cauchyfolgen konstruiert, also sind auch die $\{S_n(x)\}$ Cauchyfolgen und für die Grenzfunktion gilt wie in (IV.3): $S_\infty' = s_\infty = f$, *sie ist also Stammfunktion* von f.

V. Und was ist mit der Stetigkeit?

V.1. Das erste Semester ist noch nicht zuende, aber allerhand Üben mit Analysisargumenten hat schon stattgefunden. Ich habe zunächst solche Funktionen als untersuchenswert erklärt, die *im Definitionsbereich konvergente Folgen* abbilden *auf konvergente Folgen.* Der Schrankensatz liefert alle bisherigen Funktionen als Beispiele: $|f(a_n) - f(a_m)| \le L \cdot |a_n - a_m|$. Offenbar erben Summe, Produkt und Komposition die betrachtete Eigenschaft.

V.2. Aber $f(a) > 0 \Rightarrow f > 0$ in $[a-\delta,\ a+\delta]$ ist eine neuartige Schwierigkeit. Wir kommen zur ϵ-δ-Definition.

V.3. Abschluß des Semesters sind die Hauptsätze über stetige Funktionen, mit Beispielen von Cantor, Hilbert, Koch, Weierstraß. Das macht die Hauptsätze eindrucksvoller.

Vom zweiten Semester an steht auch den so Ausgebildeten die Standardliteratur zur Verfügung. Ein 60-Seiten Manuskript liegt auf meiner homepage.