Approximation

Lösungshinweise

Gerhard Opfer SS 2000

Bogen 3: Ausgabetermin: am 20. April 2000
Abgabetermin: am 27. April 2000 vor den Übungen

Präsenzaufgaben:

Aufgabe P7. Wie lautet das Kolmogoroff-Kriterium in verschiedenen Fällen: (a) allgemeiner Fall, $Q = (X, || \quad ||, f, V)$, V Teilvektorraum, $f \in X \backslash \overline{V}$, X^* ist der Dualraum von X und $x^* \in X^*$. (b) spezieller Fall $X = \mathrm{C}(B), || \quad || = || \quad ||_\infty$.

Lösung. (a) In allen Fällen ist $\hat{v}$ genau dann eine b. A. von f, wenn die angegebenen Bedingungen gelten:

Satz 1 : Es gibt x^* mit $x^*(f - \hat{v}) = ||f - \hat{v}||, \quad ||x^*|| = 1, \quad x^* \perp V.$

Satz 2 : Zu jedem $v \in V$ existiert Extremalpunkt x^* mit $\Re(x^*(v - \hat{v})) \le 0$, $|x^*(f - \hat{v})| = ||f - \hat{v}||$.

Satz 3 :
Zu jedem $v \in V$ existiert Extremalpunkt x^* mit $\Re\{(\overline{x^*(f - \hat{v})} x^*(v - \hat{v}))\} \le 0$, $|x^*(f - \hat{v})| = ||f - \hat{v}||$.

(b) Kommt in der nächsten Aufgabe P8 (a) vor.

Aufgabe P8. (a) Skizzieren Sie den Beweis des speziellen Kolmogoroff-Kriteriums für den konvexen Fall. (b) Wie unterscheidet sich der angegebene Beweis von dem allgemeinen, linearen Fall?

Lösung. (a) Sei $X = \mathrm{C}(B)$ und $V \subset X$ konvex und immer $|| \quad || = || \quad ||_\infty$. Bei gegebenen $f \in X, \hat{v} \in V$ definieren wir

(L3.1) $$\mu(t, v) := \Re\{(\overline{f(t) - \hat{v}(t)})(v(t) - \hat{v}(t))\}, t \in B, \ v \in V.$$

Wir definieren auch noch die (nie leere und kompakte) Menge der Extremalstellen von $f - v$:

(L3.2) $$E_v := \{t \in B : |f(t) - v(t)| = ||f - v||\}.$$

Satz L3.1. (Kolmogoroff-Kriterium für stetige Funktionen, konvexer Fall) Das obige $\hat{v}$ ist genau dann b. A, von f wenn

(L3.3) $$\min_{t \in E_{\hat{v}}} \mu(t, v) \le 0 \text{ für alle } v \in V.$$

Beweis: (a) Gilt das Kolmogoroff-Kriterium (L3.3), so zeigen wir an späterer Stelle, daß $\hat{v}$ eine b. A, von f ist. (b) Gelte (L3.3) nicht. Dann gibt es ein $v_0 \in V$ und eine Konstante $\mu_0 > 0$ mit

$$\mu(t, v_0) \ge \mu_0 \text{ für alle } t \in E_{\hat{v}}.$$

Der restliche Beweis besteht darin zu zeigen, daß ein $\lambda \in]0, 1]$ gefunden werden kann, so daß

(L3.4) $$\hat{\hat{v}} := \hat{v} + \lambda(v_0 - \hat{v}) \in V$$

eine bessere Approximation als $\hat{v}$ ist. Wir definieren für $t \in B$ und λ aus (L3.4) die Größe

(L3.5) $$p(\lambda, t) := |(f - \hat{\hat{v}})(t)|^2 - ||f - \hat{v}||^2 + \lambda\mu_0.$$

Können wir $p(\lambda,t)\le 0$ für ein positives λ und alle $t\in B$ zeigen, so folgt $|(f-\hat{\hat{v}})(t)|^2\le ||f-\hat{v}||^2-\lambda\mu_0$ und daher auch die Behauptung $||f-\hat{\hat{v}}||^2<||f-\hat{v}||^2$. Setzen wir $\hat{\hat{v}}$ aus (L3.4) in (L3.5) ein, so erhalten wir nach einiger Rechnung einen Ausdruck der Form

(L3.6)

$$p(\lambda,t):=a(t)\lambda^2-2b(t)\lambda-c(t) \text{ mit } a(t)=|(v_0-\hat{v})(t)|^2, b(t)=\mu(t,v_0)-\frac{\mu_0}{2}, c(t)=||f-\hat{v}||^2-|(f-\hat{v})(t)|^2.$$

Für alle t mit $a(t)>0$ hat das quadratische Polynom $p(\lambda,t)$ zwei reelle Nullstellen $\lambda_{-,+}$, mit $\lambda_-\le -c(t)\le 0$, $\lambda_+>0$, zwischen den beiden Nullstellen ist $p(\lambda,t)<0$. Für $a(t)=0$ ist die einzige Nullstelle $\lambda_+>0$ und links von dieser Nullstelle ist $p(\lambda,t)<0$. Global ist $\lambda_+>0$ stetig von t abhängig (man verwende de l'Hospital für den Übergang $a(t)\to 0$ in der Formel für die Wurzel λ_+). In allen Fällen $t\in B$ gibt es also ein Intervall $]0,\lambda_+(t)]$ auf dem $p(\lambda,t)<0$ ist und also $\lambda_0:=\min_{t\in B}\lambda_+(t)>0$ (wegen der Kompaktheit von B und der Stetigkeit von $\lambda_+(t)$) und $p(\lambda_0,t)<0$ für alle $t\in B$. Ist $\lambda_0>1$, so kann $\lambda=1$ in (L3.4) gesetzt werden.
(b) Er ist (partiell) konstruktiv und macht keinen Gebrauch von den bisher hergeleiteten Sätzen. Im allgemeinen Fall werden funktionalanalytische Methoden verwendet.

Aufgabe P9. (a) Haben Charakterisierungssätze etwas mit Eindeutigkeits- oder Existenzsätzen zu tun? (b) Mit welchen Techniken kann man untere Schranken für $\operatorname{dist}(f,V)$ gewinnen.

Lösung. (a) Nein. Charakterisierungssätze können im Prinzip auf der leeren Menge operieren. (b) Ist V ein Vektorraum, so können geeignete Funktionale x^* verwendet werden. Die Lösung von Präsenzaufgabe P6 enthält weitere Hinweise. Eine andere Technik wurde in der Vorlesung am 20. April besprochen.

Hausaufgaben:

Aufgabe 5. Sei x^* ein Extremalpunkt der Einheitskugel $B^*:=\{x^*\in X^*: ||x^*||\le 1\}$. (a) Zeigen Sie, daß auch γx^* ein Extremalpunkt ist für jedes $\gamma\in\mathbb{K}$ mit $|\gamma|=1$. (b) Zeigen Sie, daß für $X=\mathrm{C}(B)$ die Elemente $x^*\in X^*$, definiert durch $x^*(x):=x(t_0)$ für ein festes $t_0\in B$ Extremalpunkte der Einheitskugel B^* sind. (c) Zeigen Sie (in der Situation (b)), daß es keine anderen Extremalpunkte als $x^*(x):=\gamma x(t_0)$ mit $|\gamma|=1$ und $t_0\in B$ gibt. (d) Zeigen Sie, daß die Einheitskugel von $X:=L_1[0,1]$ keine Extremalpunkte besitzt.

Lösung. (a) Eine Darstellung $x^*=(1-\lambda)y^*+\lambda z^*$ mit $y^*\ne z^*$ und $x^*\ne y^*, x^*\ne z^*$ mit $\lambda\in]0,1[$ gibt es nach Voraussetzung nicht. Daraus folgt, daß eine entsprechende Darstellung für γx^*, $|\gamma|=1$ auch nicht existieren kann.
(b) Mit ∂B^* bezeichnen wir die Funktionale $x^*\in X^*$ mit $||x^*||=1$. Ein Funktional x^* heißt *positiv*, wenn $x^*(v)\ge 0$ für $v\ge 0$. Für positive Funktionale $v^*\in\partial B^*$ gilt $v^*(1)=1$, dabei ist die linke 1 die konstante Funktion Eins. Seien zuerst alle auftretenden Funktionale positiv und sei $x^*=(1-\lambda)x_1^*+\lambda x_2^*$ mit $\lambda\in]0,1[$, $x^*,x_1^*,x_2^*\in\partial B^*$. Sei $x^*(v)=v(t_0)$. Wir wollen $x_1^*(v)=x_2^*(v)=0$ zeigen. Jede stetige Funktion v ist die Differenz von zwei nicht-negativen Funktionen. Es genügt also den Beweis für nichtnegative Funktionen zu führen. Wir wählen ein v mit $v(t_0)=0$. Also ist $x^*(v)=0=(1-\lambda)x_1^*(v)+\lambda x_2^*(v)$, und damit $x_1^*(v)=x_2^*(v)=0$. Jedes $w\in\mathrm{C}(B)$ hat die Form $w=w(t_0)\cdot 1+v$ mit $v(t_0)=0$. Also $x_1^*(w)=w(t_0)x_1^*(1)+x_1^*(v)=w(t_0)=x^*(w)$, somit $x^*=x_1^*=x_2^*$. Damit ist x^* Extremalpunkt. Geht man zu reellen Funktionalen, dann zu komplexen Funktionalen über erkennt man, daß auch γx^* mit $|\gamma|=1$ Extremalpunkte sind.
(c) Köthe [1969, S. 334]. [kein Punktabzug, wenn dieser Teil nicht vorkommt]
(d) Sei $x\in L_1[0,1]$ und $||x||_1=1$, also $\int_0^1|x(t)|\,dt=1$. Dann gibt es ein $0<c<1$ mit $\int_0^c|x(t)|\,dt=1/2$. Wir setzen $x_1(t):=2x(t)$ auf $[0,c[$ und Null im Rest und $x_2(t):=2x(t)$ auf $[c,1]$ und Null im Rest. Dann ist $||x_1||_1=||x_2||_1=1$ und $x=\frac{1}{2}(x_1+x_2)$, also gibt es keine Extremalpunkte.

Aufgabe 6. (a) Beweisen Sie den Satz von Carathéodory: Im $\mathbb{K}^n$ besteht die konvexe Hülle einer Menge M aus allen Konvexkombinationen mit höchstens $m+1$ Summanden mit

$$m=\begin{cases} n & \text{im Fall } \mathbb{K}=\mathbb{R},\\ 2n & \text{im Fall } \mathbb{K}=\mathbb{C}.\end{cases} \tag{3.2}$$

(b) Zeigen Sie an Beispielen mit $n=1$ und $n=2$, daß die angegebene Schranke $m+1$ nicht verkleinert werden kann. Hinweis zu (a): Zeigen Sie: Ist x eine Konvexkombination mit $r>m+1$ Summanden, so kann x als Summe von $r-1$ Summanden dargestellt werden.

Lösung. (a) Sei $M \subset \mathbb{K}^n$, $k \in K := \operatorname{conv}(M)$ mit $k = \sum_{j=1}^r \lambda_j m_j$, $m_j \in M$ und $\lambda_j > 0$, $\sum_{j=1}^r \lambda_j = 1$. Ist $r \le m+1$, so ist nichts zu zeigen. Sei $r > m+1$. Wir wollen zeigen, daß in diesem Fall die Zahl der Summanden r um Eins reduziert werden kann. In diesem Fall sind die $r-1 > m$ Differenzen $m_j - m_r, j = 1,2,\ldots,r-1$ linear abhängig. Also ist $\sum_{j=1}^{r-1} \alpha_j(m_j - m_1) = 0$ mit $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_{r-1}) \neq \mathbf{0}$ und $\alpha \in \mathbb{R}^{r-1}$. Ist $\mathbb{K} = \mathbb{R}$, so ist das klar. Ist aber $\mathbb{K} = \mathbb{C}$, so können die Vektoren m_j als Vektoren im $\mathbb{R}^{2n}$ aufgefaßt werden, und dann sind die $r-1 > m = 2n$ Vektoren $(m_j - m_1)$ auch als Vektoren im $\mathbb{R}^{2n}$ linear abhängig. Wir setzen $\alpha_r = -\sum_{j=1}^{r-1} \alpha_j$, dann ist $\sum_{j=1}^r \alpha_j = 0$ und $\sum_{j=1}^r \alpha_j m_j = \mathbf{0}$. Aus der Definition der α_j folgt, daß es einen Index j mit $\alpha_j > 0$ geben muß. Sei $\mu_j := \lambda_j - \tau\alpha_j$, $j = 1,2,\ldots,r$, und $1/\tau = \max_j \frac{\alpha_j}{\lambda_j} = \frac{\alpha_{j_0}}{\lambda_{j_0}} > 0$. Dann ist $\mu_j \ge 0$, $\mu_{j_0} = 0$ und $\sum_{j=1}^r \mu_j = \sum_{j=1}^r \lambda_j - \tau \sum_{j=1}^r \alpha_j = 1$. Und schließlich $k = \sum_{j=1}^r \lambda_j m_j = \sum_{j=1}^r (\mu_j + \tau\alpha_j) m_j = \sum_{\substack{j=1 \\ j \neq j_0}}^r \mu_j m_j$. Damit ist die Summe um einen Summanden reduziert.

(b) Wir betrachten für $n = 1$ die beiden Fälle $M := \{0,1\} \in \mathbb{R}$ und $M := \{0,1,\mathrm{i}\} \in \mathbb{C}$. Im ersten Fall brauchen wir zur Bildung der konvexen Hülle $n+1 = 2$ Summanden und im zweiten Fall $2n+1 = 3$ Summanden. Ist $n = 2$, so betrachten wir $M := \{(0,0),(1,0),(0,1)\} \in \mathbb{R}^2$ und $M := \{(0,0),(1,0),(0,1),(\mathrm{i},0),(0,\mathrm{i})\} \in \mathbb{C}^2$. Im ersten Fall brauchen wir $n+1 = 3$ im zweiten Fall $2n+1 = 5$ Summanden. Mit weniger kommt man nicht aus.