

Seminarvortrag

Der Positivstellensatz

Christian Christensen

19. April 2005

Ziel des Vortrag ist es eine Bedingung zu finden, mit der man entscheiden kann, ob ein Polynom positiv ist.

1 Präordnungen

In diesem Abschnitt sei R ein kommutativer Ring mit 1 und $\sum R^2$ bezeichne die Menge aller endlichen Summen $\sum r_i^2$ mit $r_i \in R$. Das heißt

$$\sum R^2 = \left\{ \sum_{i=1}^n r_i^2 \mid n \in \mathbb{N} \text{ und } r_i \in R \text{ für } i = 1, \dots, n \right\} .$$

Diese Menge wird auch SOS-Menge (sum of squares) genannt.

1.1 Definition

Eine Präordnung in R ist eine Teilmenge T von R , so daß

$$T + T \subseteq T, \quad TT \subseteq T \quad \text{und} \quad r^2 \in T \quad \forall r \in R.$$

Das heißt eine Präordnung ist additiv und multiplikativ abgeschlossen und enthält alle Quadrate.

1.2 Bemerkung

Offensichtlich ist $\sum R^2$ eine Präordnung in R und $\sum R^2 \subseteq T$ für jede Präordnung T in R , das heißt eine Präordnung in R enthält alle Quadratsummen.

1.3 Bemerkung/Definition

Wir bezeichnen die kleinste Präordnung in R , die die endliche Menge $S = \{g_1, \dots, g_s\}$ enthält, mit $\sum R^2[S]$. Sie besteht also aus Termsummen der Form

$$\sigma g_1^{e_1} \cdots g_s^{e_s}, \quad \sigma \in \sum R^2, \quad e_i \in \{0, 1\} \text{ für } i = 1, \dots, s .$$

Mit der Standard Notation für Multiindizes, das heißt g^e für $g_1^{e_1} \cdots g_s^{e_s}$, wobei $e = (e_1, \dots, e_s)$, ist also

$$\sum R^2[S] = \left\{ \sum_{e \in \{0,1\}^s} \sigma_e g^e \mid g \in S \text{ und } \sigma_e \in \sum R^2 \quad \forall e \in \{0,1\}^s \right\} .$$

2 Der Positivstellensatz

Für den restlichen Vortrag gelten folgende Bezeichnungen, sofern nicht ausdrücklich anders angegeben.

- $\mathbb{R}[X] := \mathbb{R}[X_1, \dots, X_n]$ bezeichnet den Polynomring in \mathbb{R} .
- $S = \{g_1, \dots, g_s\}$ endliche Teilmenge von $\mathbb{R}[X]$.
- $K_S := \{x \in \mathbb{R} \mid \forall g \in S : g(x) \geq 0\} \stackrel{\text{hier}}{=} \{x \in \mathbb{R}^n \mid g_i(x) \geq 0 \text{ für } i = 1, \dots, s\}$
- $T_S := \sum \mathbb{R}[X]^2[S]$ die von S in $\mathbb{R}[X]$ erzeugte Präordnung.

2.1 Positivstellensatz

Mit den obigen Bezeichnungen gilt für alle $f \in \mathbb{R}[X]$:

- (1) $f > 0$ in $K_S \Leftrightarrow \exists p, q \in T_S : pf = 1 + q$
- (2) $f \geq 0$ in $K_S \Leftrightarrow \exists m \in \mathbb{N}_0, p, q \in T_S : pf = f^{2m} + q$
- (3) $f = 0$ in $K_S \Leftrightarrow \exists m \in \mathbb{N}_0 : -f^{2m} \in T_S$

2.2 Bemerkungen

- (i) Der Positivstellensatz ist auch erfüllt, wenn \mathbb{R} durch ein beliebigen reellen abgeschlossenen Körper ersetzt wird.
- (ii) In allen drei Fällen (1), (2), (3) ist die Implikation " \Leftarrow " trivial.

- Für (1):
Wir wissen, daß $pf = 1 + q$ für $p, q \in T_S$. Dann gilt für alle $x \in K_S$, das heißt $p(x) \geq 0$ und $q(x) \geq 0$,

$$p(x)f(x) = 1 + q(x) > 0 .$$

Dies erzwingt aber $p(x) \neq 0$, das heißt $p(x) > 0$, und $f(x) > 0$.

- Für (2):
Wir wissen, daß $pf = f^{2m} + q$ für $p, q \in T_S$ und irgendein $m \in \mathbb{N}_0$. Dann gilt für alle $x \in K_S$, das heißt $p(x) \geq 0$ und $q(x) \geq 0$,

$$p(x)f(x) = f^{2m}(x) + q(x) \geq 0 .$$

Für $p(x) \neq 0$ folgt direkt $f(x) = \frac{f^{2m}(x)+q(x)}{p(x)} \geq 0$ wie gewünscht.

Für $p(x) = 0$ gilt $f^{2m}(x) = -q(x)$. Da aber $q(x) \geq 0$ und die linke Seite ein Quadrat ist (, das heißt ≥ 0), folgt $f^{2m}(x) = 0$. Also $f(x) \geq 0$.

- Für (3):
Wir wissen $-f^{2m} \in T_S$ für irgendein $m \in \mathbb{N}_0$, das heißt $-f^{2m}(x) \geq 0 \Leftrightarrow f^{2m}(x) \leq 0$ für $x \in K_S$. Somit folgt, da es sich um ein Quadrat handelt: $f(x) = 0$.

(iii) Wir erhalten zwei berühmte Sätze als Spezialfälle des Positivstellensatzes.

- (a) Wenn wir in (2) $S = \emptyset$ setzen, erhalten wir $K_S = \mathbb{R}$ und $T_S = \sum \mathbb{R}[X]^2$. Auf diesem Weg liefert (2) einen alternativen Beweis für Hilbert's 17th Problem.

Hilbert's 17th Problem: Wenn $f \in \mathbb{R}[X]$, $f \geq 0$ in \mathbb{R}^n , dann gilt $f \in \sum \mathbb{R}(X)^2$.

Beweis:

Nach (2) gilt $pf = f^{2m} + q$ für $p, q \in \sum \mathbb{R}[x]^2$ und irgendein $m \in \mathbb{N}_0$. Wenn $f \neq 0$, folgt $f^{2m} + q \neq 0$, also $p \neq 0$. Somit gilt:

$$f = \frac{1}{p}(f^{2m} + q) = \left(\frac{1}{p}\right)^2 p(f^{2m} + q) \in \sum \mathbb{R}(x)^2 .$$

Für $f = 0$ ist das Ergebnis klar. \square

- (b) (3) enthält als Spezialfall den reellen Nullstellensatz.

Reeller Nullstellensatz:

Voraussetzung: Für $S = \{g_1, \dots, g_s\} \subseteq \mathbb{R}[X]$ sei $Z_S := \{x \in \mathbb{R}^n \mid g_i(x) = 0 \text{ für } i = 1, \dots, s\}$ und I bezeichne das von g_1, \dots, g_s erzeugte Ideal in $\mathbb{R}[X]$.

Behauptung: $\forall f \in \mathbb{R}[X] : f = 0 \text{ in } Z_S \iff \exists m \in \mathbb{N} \text{ und } \sigma \in \sum \mathbb{R}[X]^2 \text{ so das } f^{2m} + \sigma \in I$

Beweis:

" \Leftarrow ": Wenn $f^{2m} + \sigma \in \mathfrak{a}_S$, dann gilt für alle $x \in Z_S : f(x)^{2m} + \sigma(x) = 0$.

Weil $\sigma(x) \geq 0$ impliziert dies $f(x) = 0$.

" \Rightarrow ": Verwende (3) mit der Menge $S' = S \cup -S$. Dann gilt $Z_S = K_{S'}$ und wir erhalten $-f^{2m} = p$ für $p \in T_{S'}$ und irgendein $m \in \mathbb{N}_0$. Jedes $p \in T_{S'} = \sum \mathbb{R}[X][S']$ zerfällt aber in der Form $p = \sigma + \tau$ mit $\sigma \in \sum \mathbb{R}[X]^2$ und $\tau \in I$. Also $-f^{2m} = \sigma + \tau \Leftrightarrow f^{2m} + \sigma = -\tau \in I$. \square

3 Der Beweis

Wir zeigen als erstes, daß (1), (2) und (3) äquivalent sind. Danach beenden wir den Beweis, indem (1) zeigen.

3.1 Beweis (1) \Rightarrow (2):

Annahme: $f \geq 0$ in K_S .

Jetzt benutzen wir den Rabinowitschtrick.

Wir wechseln in eine um eins höhere Dimension mit der Notation

$$(x, y) = (x_1, \dots, x_n, y) \in \mathbb{R}^{n+1}, \quad \mathbb{R}[X, Y] = \mathbb{R}[X_1, \dots, X_n, Y]$$

Wir definieren

$$S' := S \cup \{Yf - 1, -Yf + 1\} = \{g_1, \dots, g_s, Yf - 1, -Yf + 1\} .$$

Dann ist

$$K_{S'} = \{(x, y) \in \mathbb{R}^{n+1} \mid yf(x) = 1 \text{ und } g_i(x) \geq 0 \text{ für } i = 1, \dots, s\}$$

Somit gilt auf $K_{S'}$ $f(x, y) = f(x) > 0$, da sonst $yf(x) = 0$, und mit (1)

$$p'(X, Y)f(X) = 1 + q'(X, Y) \text{ für } p', q' \in T_{S'} .$$

Ersetzen wir $Y = \frac{1}{f(X)}$ in dieser Gleichung und beseitigen wir den Nenner durch multiplizieren beider Seiten mit $f(X)^{2m}$, für m ausreichend groß, das heißt größer als die höchste Y Potenz von p' und q' , erhalten wir

$$p(X)f(X) = f(X)^{2m} + q(X)$$

mit

$$p(X) = f(X)^{2m}p' \left(X, \frac{1}{f(X)} \right), \quad q(X) = f(X)^{2m}q' \left(X, \frac{1}{f(X)} \right) .$$

Um den Beweis zu beenden genügt es jetzt zu prüfen, daß $p, q \in T_S$ für ausreichend großes m . Nach Definition von $T_{S'}$ ist $p'(X, Y)$ eine Summe von Termen der Form

$$\sigma(X, Y)g_1(X)^{e_1} \dots g_s(X)^{e_s} (Yf(X) - 1)^{e_{s+1}} (-Yf(X) + 1)^{e_{s+2}} ,$$

wobei $e_i \in \{0, 1\}$ für $i = 1, \dots, s + 2$ und $\sigma(X, Y) \in \sum \mathbb{R}[X, Y]^2$. Setze $\sigma(X, Y) = \sum h_j(X, Y)^2$. Nun ersetzen wir wieder $Y = \frac{1}{f(X)}$. Dadurch verschwinden die Terme mit $e_{s+1} = 1$ oder $e_{s+2} = 1$. Anschließend multiplizieren wir wieder mit $f(X)^{2m}$, wobei $m \geq$ der höchsten Potenz von Y in $h_j(X, Y)$ ist. Das heißt

$$h_j(X, Y) = \sum_{i=0}^v h_{ij}(X)Y^i, \quad v \leq m .$$

Dann gilt

$$f(X)^m h_j \left(X, \frac{1}{f(X)} \right) = \sum_{i=0}^v h_{ij}(X) f(X)^{m-i} \in \mathbb{R}[X] .$$

Durch quadrieren und bilden der Summation folgt

$$f(X)^{2m} \sigma \left(X, \frac{1}{f(X)} \right) = \sum_j \left(f(X)^m h_j \left(X, \frac{1}{f(X)} \right) \right)^2 \in \sum \mathbb{R}[X]^2 .$$

Das Argument für q ist analog. \square

3.2 Beweis (2) \Rightarrow (3):

Annahme: $f = 0$ in K_S .

Dann liefert (2) angewendet auf f und $-f$

$$\left. \begin{array}{l} p_1 f = f^{2m_1} + q_1 \\ -p_2 f = f^{2m_2} + q_2 \end{array} \right\} p_i, q_i \in T_S, \quad m_i \in \mathbb{N}_0, \quad \text{für } i = 1, 2 .$$

Multiplikation der beiden Gleichungen liefert

$$-p_1 p_2 f^2 = f^{2(m_1+m_2)} + f^{2m_1} q_2 + f^{2m_2} q_1 + q_1 q_2 .$$

Also $-f^{2m} = p$, wobei $m = m_1 + m_2$ und

$$p = p_1 p_2 f^2 + f^{2m_1} q_2 + f^{2m_2} q_1 + q_1 q_2 .$$

Da T_S eine Präordnung ist, ist T_S abgeschlossen unter Addition und Multiplikation und enthält alle Quadratsummen. Somit folgt $p \in T_S$. \square

3.3 Beweis (3) \Rightarrow (1):

Sei $S' = S \cup \{-f\}$. Dann gilt

$$f > 0 \text{ in } K_S \Leftrightarrow K_{S'} = \emptyset \stackrel{\textcircled{*}}{\Leftrightarrow} 1 = 0 \text{ in } K_{S'} \stackrel{\textcircled{(3)}}{\Leftrightarrow} -1 \in T_{S'} .$$

$\textcircled{*}$ folgt mit $f = 1$ und aus der Tatsache, das man für eine leere Menge jede Aussage treffen kann. Zusätzlich folgt, da $S' = S \cup \{-f\}$, daß $T_{S'} = T_S - fT_S$. Also $-1 = q - pf$, daß heißt $pf = 1 + q$ für $p, q \in T_S$. \square

Für den Beweis von (1) benötigen wir noch zwei Lemma.

3.4 Lemma

Voraussetzung: R sei ein kommutativer Ring mit 1 und P eine Präordnung in R . Weiter sei P maximal unter der Auflage $-1 \notin P$.

Behauptung: $P \cup -P = R$ und $P \cap -P$ ist ein Primideal von R .

Beweis:

(a) Behauptung: $P \cup -P = R$.

Annahme: $g \in R$ und $g \notin P \cup -P$, das heißt $g \notin P$ und $g \notin -P$.

Dann gilt $-1 \in P + gP$, $-1 \in P - gP$, weil P sonst nicht maximal wäre. Also

$$\left. \begin{array}{l} -1 = s_1 + gt_1 \\ -1 = s_2 - gt_2 \end{array} \right\} s_i, t_i \in P \text{ für } i = 1, 2 .$$

Dann gilt $-gt_1 = 1 + s_1$ und $gt_2 = 1 + s_2$. Multiplizieren der beiden Gleichungen liefert

$$-g^2 t_1 t_2 = (1 + s_1)(1 + s_2) = 1 + s_1 + s_2 + s_1 s_2 .$$

Also $-1 = s_1 + s_2 + s_1s_2 + g^2t_1t_2 \in P$, was einen Widerspruch zur Voraussetzung ist. Somit gilt $P \cup -P = R$.

(b) Behauptung: $\mathcal{P} := P \cap -P$ ist ein Ideal in R .

Es gilt:

- $0 \in \mathcal{P}$, da $0 = 0^2 \in P$ und $0 = -0 \in -P$.
- $\mathcal{P} + \mathcal{P} \subseteq \mathcal{P}$, da P als Präordnungen additiv abgeschlossen ist und $-P - P = -(P + P) \subseteq -P$.
- $-\mathcal{P} = \mathcal{P}$, da der Schnitt zweier Mengen kommutativ ist.
- $P\mathcal{P} \subseteq \mathcal{P}$, da $P\mathcal{P} = P(P \cap -P) = PP \cap -PP \subseteq P \cap -P = \mathcal{P}$.

Dann

$$R\mathcal{P} = (P \cup -P)\mathcal{P} \subseteq P\mathcal{P} \cup -P\mathcal{P} \subseteq \mathcal{P} .$$

Dieses liefert, das \mathcal{P} ein Ideal ist.

(c) Behauptung: $\mathcal{P} := P \cap -P$ ist prim.

Annahme: $gh \in \mathcal{P}$, $g, h \notin \mathcal{P}$.

Ersetze, falls nötig, g durch $-g$ bzw. h durch $-h$, und wir können $g, h \notin P$ voraussetzen. Somit folgt analog zu oben, $-1 \in P + gP$ und $-1 \in P + hP$, also

$$\left. \begin{array}{l} -1 = s_1 + gt_1 \\ -1 = s_2 + ht_2 \end{array} \right\} s_i, t_i \in P \text{ für } i = 1, 2 .$$

Dann gilt $-gt_1 = 1 + s_1$ und $-ht_2 = 1 + s_2$. Multiplizieren der beiden Gleichungen liefert

$$ght_1t_2 = (1 + s_1)(1 + s_2) = 1 + s_1 + s_2 + s_1s_2 .$$

Also $-1 = s_1 + s_2 + s_1s_2 - ght_1t_2$. Da $gh \in \mathcal{P}$ und \mathcal{P} ein Ideal ist, ist $-1 \in P$, was einen Widerspruch zur Annahme darstellt. Also ist \mathcal{P} prim. \square

3.5 Lemma

Voraussetzung: R sei ein kommutativer Ring mit 1 und P eine Präordnung in R , so daß $R = P \cup -P$ und $\mathcal{P} = P \cap -P$ ein Primideal.

Behauptung: P induziert durch

$$\frac{\bar{g}}{\bar{h}} \geq 0 \Leftrightarrow gh \in P$$

eine eindeutige Ordnung im Quotientenkörper $Q = Q(R/\mathcal{P})$. Hier ist $\bar{g} = g + \mathcal{P}$ und $\bar{h} = h + \mathcal{P}$ mit $\bar{h} \neq \bar{0}$, das heißt $h \notin \mathcal{P}$.

Beweis:

Zu zeigen ist, dass durch $\frac{\bar{g}}{\bar{h}} \geq 0 \Leftrightarrow gh \in P$ eine Ordnung auf Q gegeben ist. Dazu müssen wir Reflexivität ($x \leq x$), Antisymmetrie ($x \leq y, y \leq x \Rightarrow x = y$) und Transitivität ($x \leq y, y \leq z \Rightarrow x \leq z$) zeigen.

(a) Reflexivität:

$$\frac{\bar{g}}{h} \geq \frac{\bar{g}}{h} \Leftrightarrow \frac{\bar{g}}{h} - \frac{\bar{g}}{h} \geq 0 \Leftrightarrow \frac{\bar{g}h - \bar{g}h}{h^2} \geq 0 \Leftrightarrow h^2(gh - gh) \in P \Leftrightarrow 0 \in P$$

(b) Antisymmetrie:

$$\frac{\bar{g}_1}{h_2} \geq \frac{\bar{g}_2}{h_2} \Leftrightarrow \frac{\bar{g}_1}{h_1} - \frac{\bar{g}_2}{h_2} \geq 0 \Leftrightarrow \frac{\bar{g}_1\bar{h}_2 - \bar{g}_2\bar{h}_1}{h_1\bar{h}_2} \geq 0 \Leftrightarrow h_1h_2(g_1h_2 - g_2h_1) \in P$$

$$\frac{\bar{g}_2}{h_2} \geq \frac{\bar{g}_1}{h_1} \Leftrightarrow \frac{\bar{g}_2}{h_2} - \frac{\bar{g}_1}{h_1} \geq 0 \Leftrightarrow \frac{\bar{g}_2\bar{h}_1 - \bar{g}_1\bar{h}_2}{h_2\bar{h}_1} \geq 0 \Leftrightarrow h_1h_2(g_2h_1 - g_1h_2) \in P$$

Damit folgt aber $h_1h_2(g_1h_2 - g_2h_1) \in -P$ und $h_1h_2(g_2h_1 - g_1h_2) \in -P$ (, weil das erste Element das negative des zweiten Elementes ist). Also gilt $h_1h_2(g_1h_2 - g_2h_1)$, $h_1h_2(g_2h_1 - g_1h_2) \in \mathcal{P}$. Das bedeutet aber $g_1 = g_2$ und $h_1 = h_2$.

(c) Transitivität:

Hierfür zeigen wir, dass unsere Ordnungsdefinition $\frac{\bar{g}}{h} \geq 0 \Leftrightarrow gh \in P$ additionsverträglich ist. Es gilt

$$\frac{\bar{g}_1}{h_1} \geq 0 \Leftrightarrow g_1h_1 \in P \text{ und } \frac{\bar{g}_2}{h_2} \geq 0 \Leftrightarrow g_2h_2 \in P .$$

Nach der Addition der beiden Brüche und Anwendung der Ordnungsdefinition folgt

$$\frac{\bar{g}_1}{h_1} + \frac{\bar{g}_2}{h_2} \geq 0 \Leftrightarrow \frac{\bar{g}_1\bar{h}_2 + \bar{g}_2\bar{h}_1}{h_1\bar{h}_2} \geq 0 \Leftrightarrow h_1h_2(g_1h_2 + g_2h_1) \in P .$$

Die letzte Aussage gilt, da wir wissen das die Präordnung P additiv und multiplikativ abgeschlossen ist, sowie alle Quadratsummen und die Elemente g_2h_1 , g_1h_2 enthält.

Damit folgt die Transitivität aus

$$\left. \begin{array}{l} \frac{\bar{g}_1}{h_1} \geq \frac{\bar{g}_2}{h_2} \Leftrightarrow \frac{\bar{g}_1}{h_1} - \frac{\bar{g}_2}{h_2} \geq 0 \\ \frac{\bar{g}_2}{h_2} \geq \frac{\bar{g}_1}{h_1} \Leftrightarrow \frac{\bar{g}_2}{h_2} - \frac{\bar{g}_1}{h_1} \geq 0 \end{array} \right\} + \Rightarrow \frac{\bar{g}_1}{h_1} - \frac{\bar{g}_3}{h_3} \geq 0 \Leftrightarrow \frac{\bar{g}_1}{h_1} \geq \frac{\bar{g}_3}{h_3} . \quad \square$$

3.6 Beweis von (1):

Voraussetzung: $f \in \mathbb{R}[X]$, $f > 0$ in K_S .

Zu zeigen: $\exists p, q \in T_S : pf = 1 + q$.

Annahme: Es existieren keine $p, q \in T_S$, so das $pf = 1 + q \Leftrightarrow -1 = q - pf$.

Dann gilt

$$-1 \notin T_S - fT_S .$$

Nach dem Zornschen Lemma existiert eine Präordnung P in $\mathbb{R}[X]$, welche, unter der Auflage $-1 \notin P$, maximal ist und für die $T_S - fT_S \subseteq P$ gilt.

(Die Voraussetzungen des Zornschen Lemmas sind erfüllt, da $T_S - fT_S$ selbst eine

Präordnung ist.) Aus dem ersten Lemma folgt: $\mathbb{R}[X] = P \cup -P$ und $\mathcal{P} = P \cap -P$ prim. Mit dem zweiten Lemma erhalten wir, daß P eine eindeutige Ordnung \geq im Quotientenkörper $Q := Q(\mathbb{R}[X]/\mathcal{P})$ durch

$$\frac{\bar{g}}{\bar{h}} \geq 0 \Leftrightarrow gh \in P$$

induziert.

Q ist eine Erweiterung von \mathbb{R} mit dem Zusammenhang

$$\mathbb{R} \hookrightarrow \mathbb{R}[X] \rightarrow \mathbb{R}[X]/\mathcal{P} \hookrightarrow Q(\mathbb{R}[X]/\mathcal{P}) = Q .$$

Die Ordnung von Q eingeschränkt auf \mathbb{R} liefert die eindeutige Ordnung \leq auf \mathbb{R} .

Behauptung: Es existiert ein Element $x = (x_1, \dots, x_n) \in Q^n$, so daß $g_i(x) \geq 0$ für $i = 1, \dots, s$ und $f(x) \leq 0$.

Setze $x_j = \bar{X}_j = X_j + \mathcal{P}$ für $j = 1, \dots, n$. Für alle $g \in \mathbb{R}[X]$, wenn $g = \sum aX_1^{k_1} \dots X_n^{k_n}$, gilt

$$\bar{g} = \sum \bar{a}\bar{X}_1^{k_1} \dots \bar{X}_n^{k_n} = \sum ax_1^{k_1} \dots x_n^{k_n} = g(x) .$$

Also ist nur zu testen, daß $\bar{g}_i \geq 0$ für $i = 1, \dots, s$ und $\bar{f} \leq 0$.

$\bar{g}_i \geq 0$ weil $g_i \in S \subseteq T_S \subseteq T_S - fT_S \subset P$. $\bar{f} \leq -f \in T_S - fT_S \subseteq P$. Daraus folgt die Behauptung.

Mit der Behauptung und Tarski's Transfer Prinzip (siehe 1. Vortrag) folgt, daß $x \in \mathbb{R}^n$ existiert, so daß $g_i(x) \geq 0$ für $i = 1, \dots, s$ (, das heißt $x \in K_S$) und $f(x) \leq 0$.

Dies ist ein Widerspruch zu der Annahme, daß f in K_S echt positiv ist. \square