§ 23 Konvexe Funktionen und Ungleichungen

Die ersten systematischen Untersuchungen der konvexen Funktionen hat der dänische Mathematiker und Ingenieur Jensen (1859–1925) durchgeführt.

23.1 Konvexe Funktionen

Sei I ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt *konvex* (*konkav*), wenn gilt:

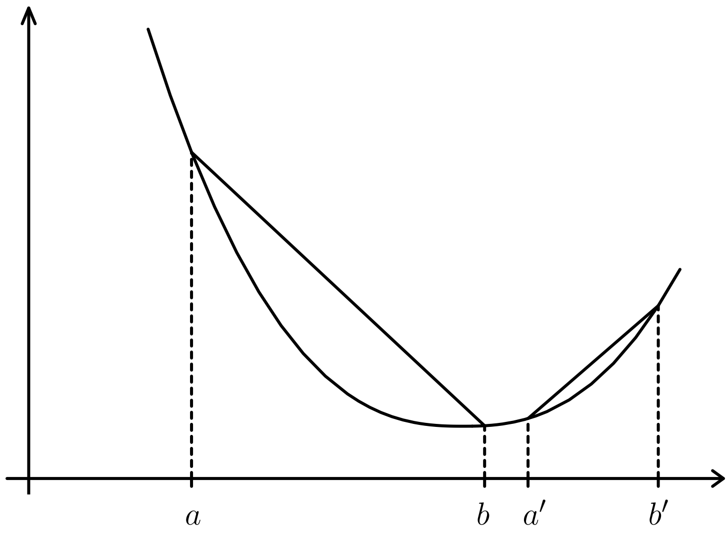
$$(\forall a, b \in I)(a < t < b \Rightarrow f(t) \underset{(\geq)}{\leq} f(a) + \tfrac{f(b)-f(a)}{b-a}(t-a)).$$

Offensichtlich ist f konkav genau dann, wenn $-f$ konvex ist. Es reicht daher im folgenden, die konvexen Funktionen zu untersuchen.

Die geometrische Bedeutung der Konvexität ergibt sich aus folgendem: Sind $a, b \in I$ und $a < b$, so ist der Graph von

$$[a, b] \ni t \to f(a) + \tfrac{f(b)-f(a)}{b-a}(t-a)$$

das Geradenstück, welches die Punkte $(a, f(a))$ und $(b, f(b))$ verbindet, also die sogenannte Sehne. Die Konvexität besagt nun, daß der Graph von $f|[a, b]$ (im unstrengen Sinne) unterhalb der Sehne verlaufen muß. Dies muß für jede Wahl von $a, b \in I$ mit $a < b$ der Fall sein.



23.2 Kriterien für Konvexität

Sei I ein Intervall und $f : I \to \mathbb{R}$ stetig und in allen Punkten von I° differenzierbar. Dann sind (i) bis (iii) äquivalent:

(i) f ist konvex.

(ii) Für alle $t_0 \in I^\circ$ gilt:

$$f(t) \geq f(t_0) + f'(t_0)(t - t_0) \text{ für alle } t \in I,$$

d.h. für alle Punkte $t_0 \in I^\circ$ gilt: Der Graph von f verläuft auf I oberhalb seiner Tangente im Punkte t_0.

(iii) f' ist monoton wachsend auf I°.

Ist f sogar zweimal differenzierbar in allen Punkten von I°, dann ist jede der Aussagen (i) bis (iii) äquivalent zu:

(iv) $f''(t) \geq 0$ für alle $t \in I^\circ$.

Beweis. (i) $\Rightarrow$(iii) Sei $a < b$ mit $a, b \in I^\circ$. Z.z. ist:

(1) $$f'(a) \leq f'(b).$$

Aus der Definition der Konvexität folgt für $t \in \,]a, b[$

$$\tfrac{f(t)-f(a)}{t-a} \leq \tfrac{f(b)-f(a)}{b-a}.$$

Mit $t \downarrow a$ erhält man aus dieser Ungleichung

(2) $$f'(a) \leq \tfrac{f(b)-f(a)}{b-a}.$$

Nun ist

(3) $$f(a) + \tfrac{f(b)-f(a)}{b-a}(t - a) = f(b) + \tfrac{f(b)-f(a)}{b-a}(t - b).$$

Mit (3) folgt daher aus der Konvexität von f

$$f(t) - f(b) \leq \tfrac{f(b)-f(a)}{b-a}(t-b),$$

und hieraus folgt wegen $t - b < 0$

(4) $$\tfrac{f(t)-f(b)}{t-b} \geq \tfrac{f(b)-f(a)}{b-a}.$$

Mit $t \uparrow b$ erhält man aus (4)

(5) $$f'(b) \geq \tfrac{f(b)-f(a)}{b-a}.$$

Aus (2) und (5) folgt dann (1).

(iii) $\Rightarrow$***(ii)*** : Sei $t_0 \in I^\circ$ und $t \in I$ mit o.B.d.A. $t \neq t_0$. Ist $t < t_0 (t > t_0)$, dann ist $f'(\zeta) \leq f'(t_0)$ für alle $\zeta \in]t, t_0[$ ($f'(\zeta) \geq f'(t_0)$ für alle $\zeta \in]t_0, t[$).

Nach dem Mittelwertsatz folgt (beachte f ist stetig auf I):

$$\tfrac{f(t)-f(t_0)}{t-t_0} = f'(\zeta) \underset{(\geq)}{\leq} f'(t_0) \text{ für ein } \zeta \text{ zwischen } t \text{ und } t_0.$$

Hieraus folgt durch Multiplikation mit $t - t_0$ dann (ii).

(ii) $\Rightarrow$***(i)*** Seien $a, b \in I$ und $a < t < b$. Wir wenden (ii) auf $t_0 := t$ und $t := a$ bzw. $t := b$ an. Wegen $t \in I^\circ$ gilt dann

(6) $$f(a) \geq f(t) + f'(t)(a-t),$$

(7) $$f(b) \geq f(t) + f'(t)(b-t).$$

Somit erhalten wir:

(8) $$\tfrac{f(t)-f(a)}{t-a} \underset{(6)}{\leq} f'(t) \underset{(7)}{\leq} \tfrac{f(b)-f(t)}{b-t}.$$

Aus (8) folgt durch Multiplikation mit $(b-t)(t-a)$

(9) $$(b-t)(f(t)-f(a)) \leq (f(b)-f(t))(t-a).$$

Addiert man auf beiden Seiten der Ungleichung in (9) nun $f(t)(t-a) - f(a)(t-a)$, so erhält man:

$$(b-a)f(t) - f(a)(b-a) \leq (f(b)-f(a))(t-a).$$

Hieraus folgt die Konvexitätsbedingung nach Division durch $(b-a)$.

Die Äquivalenz von (iii) und (iv) folgt, weil nach 19.4(ii) die differenzierbare Funktion $f'|I^\circ$ genau dann monoton wachsend ist, wenn $[f'' =](f')'$ über I° nicht-negativ ist. □

Fordert man, daß für $a < b$ aus I die Funktion $f|]a,b[$ *streng* unterhalb der Sekante liegen soll, so gelangt man zum Begriff der *strengen* Konvexität.

23.3 Streng konvexe Funktionen

Sei I ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt *streng konvex* (*streng konkav*), wenn

$$(\forall a, b, t \in I)(a < t < b \Rightarrow f(t) \underset{(>)}{<} f(a) + \tfrac{f(b)-f(a)}{b-a}(t-a)).$$

Die folgenden Bedingungen für die strenge Konvexität von Funktionen sind hinreichend, aber nicht notwendig.

23.4 Kriterien für strenge Konvexität

Sei I ein Intervall und $f : I \to \mathbb{R}$ stetig und in allen Punkten von I° differenzierbar.

(i) Ist dann f' streng monoton wachsend auf I°, so ist f streng konvex.

(ii) Ist f auf I° sogar zweimal differenzierbar und ist f'' auf I° positiv, so ist f streng konvex.

(i) Nach 23.2(iii)$\Rightarrow$(i) ist f konvex. Wäre nun f nicht streng konvex, so existieren $a, b, t \in I$ mit $a < t < b$ und

(1) $$f(t) = f(a) + \tfrac{f(b)-f(a)}{b-a}(t-a).$$

Nun ist

(2) $$f(a) + \tfrac{f(b)-f(a)}{b-a}(t-a) = f(b) + \tfrac{f(b)-f(a)}{b-a}(t-b),$$

und daher folgt:

(3) $$\tfrac{f(t)-f(a)}{t-a} \underset{(1)}{=} \tfrac{f(b)-f(a)}{b-a} \underset{(1),(2)}{=} \tfrac{f(t)-f(b)}{t-b}.$$

Nach dem Mittelwertsatz gibt es dann ein $t_1 \in]a, t[$ und ein $t_2 \in]t, b]$ mit

$$f'(t_1) = \tfrac{f(t)-f(a)}{t-a} \underset{(3)}{=} \tfrac{f(b)-f(t)}{b-t} = f'(t_2).$$

Dies widerspricht wegen $t_1 < t_2$ der strengen Monotonie von f'.

(ii) folgt aus (i) mit 19.4(iv). □

Wir erhalten nun mit Satz 19.4 und 23.4 viele Beispiele für monotone und konvexe bzw. konkave Funktionen, indem wir mit Hilfe der Ableitungsregeln die ersten und zweiten Ableitungen dieser Funktionen berechnen.

(i) Für $n \in \mathbb{N}$ gilt: $(x^n)' \underset{18.4}{=} nx^{n-1}, (x^n)'' \underset{18.4}{=} n(n-1)x^{n-2}$;

(ii) Für $b \in \mathbb{R}$ gilt $(x^b)' \underset{18.11(ii)}{=} bx^{b-1}, (x^b)'' \underset{18.11(ii)}{=} b(b-1)x^{b-2}$;

(iii) $(ln(x))' \underset{18.11(i)}{=} \frac{1}{x}|\mathbb{R}_+, (ln(x))'' \underset{18.11(ii)}{=} -\frac{1}{x^2}|\mathbb{R}_+$;

(iv) $(e^x)' \underset{18.5}{=} e^x, (e^x)'' \underset{18.5}{=} e^x$;

(v) $(e^{-x})' \underset{18.5,18.6}{=} -e^{-x}, (e^{-x})'' \underset{18.5,18.6}{=} e^{-x}$.

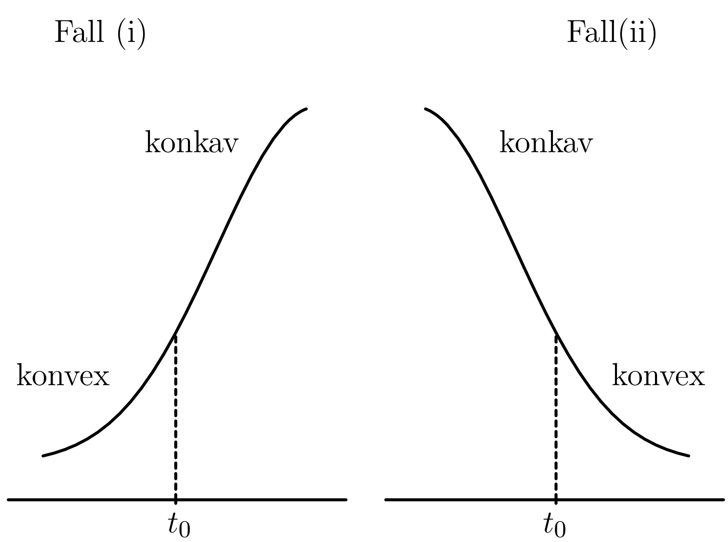
Funktion	*Monotonie*	*Konvexität/Konkavität*
$x^2,\ x^4,\ x^6,\ldots$	streng monoton fallend auf $]-\infty,0]$, streng monoton wachsend auf $[0,\infty[$	streng konvex auf $\mathbb{R}$
$x^3,\ x^5,\ x^7,\ldots$	streng monoton wachsend auf $\mathbb{R}$	streng konkav auf $]-\infty,0]$, streng konvex auf $[0,\infty[$
$x^b,\ 0<b<1$	streng monoton wachsend auf $[0,\infty[$	streng konkav auf $[0,\infty[$
$x^b,\ b>1$	streng monoton wachsend auf $[0,\infty[$	streng konvex auf $[0,\infty[$
$ln(x)$	streng monoton wachsend auf $]0,\infty[$	streng konkav auf $]0,\infty[$
e^x	streng monoton wachsend auf $\mathbb{R}$	streng konvex auf $\mathbb{R}$
e^{-x}	streng monoton fallend auf $\mathbb{R}$	streng konvex auf $\mathbb{R}$

23.5 Wendepunkt

Sei I ein Intervall, und sei $t_0 \in I^\circ$. Ferner sei $f: I \to \mathbb{R}$ stetig.
Dann heißt t_0 *ein Wendepunkt von f*, wenn es ein $\varepsilon \in \mathbb{R}_+$ gibt, so daß $]t_0-\varepsilon, t_0+\varepsilon[\subset I$ und (i) oder (ii) gilt:

(i) $f|]t_0-\varepsilon,t_0[$ ist konvex und $f|]t_0,t_0+\varepsilon[$ ist konkav;

(ii) $f|]t_0-\varepsilon,t_0[$ ist konkav und $f|]t_0,t_0+\varepsilon[$ ist konvex.



23.6 Notwendige bzw. hinreichende Bedingungen für einen Wendepunkt

Sei I ein Intervall und $f : I \to \mathbb{R}$ stetig. Es sei f in allen Punkten von I° zweimal differenzierbar.

(i) Besitzt f in $t_0 \in I^\circ$ einen Wendepunkt, so ist $f''(t_0) = 0$.

(ii) Ist $f''(t_0) = 0$ und f in t_0 dreimal differenzierbar mit $f'''(t_0) \neq 0$, so besitzt f in t_0 einen Wendepunkt.

Beweis. ***(i)*** Liegt die Situation von 23.5(i) vor, so ist $f''(t) \geq 0$ für $t \in]t_0 - \varepsilon, t_0[$ und $f''(t) \leq 0$ für $t \in]t_0, t_0 + \varepsilon[$ nach 23.2. Nun ist $f'|]t_0 - \varepsilon, t_0]$ (bzw. $f'|[t_0, t_0 + \varepsilon[$) stetig und in den Punkten von $]t_0 - \varepsilon, t_0]^\circ =]t_0 - \varepsilon, t_0[$ (bzw. $[t_0, t_0+\varepsilon[^\circ$) differenzierbar mit nicht-negativer (nicht-positiver) Ableitung. Daher ist $f'|]t_0-\varepsilon, t_0]$ monoton wachsend ($f'|[t_0, t_0+\varepsilon[$ monoton fallend). Somit besitzt f' in t_0 ein lokales Maximum. Daher ist $f''(t_0) = 0$ (siehe 18.15).

Liegt die Situation von 23.5(ii) vor, so liegt für $-f$ die Situation von 23.5(i) vor, d.h. es gilt $(-f)''(t_0) = 0$, also $f''(t_0) = 0$.

(ii) Ist $f''(t_0) = 0$ und $(f'')'(t_0) > 0$, so gilt $f''(t) < 0$ für $t \in]t_0 - \varepsilon, t_0[$ und $f''(t) > 0$ für $t \in]t_0, t_0 + \varepsilon[$ mit einem geeigneten $\varepsilon \in \mathbb{R}_+$. (Wende 18.13(i) auf $f''|I^\circ$ an). Also liegt die Situation von 23.5(ii) vor (benutze 23.2(iv) oder 23.4(ii)). Der Fall $f'''(t_0) < 0$ folgt entsprechend mit 18.13(ii). □

Wichtige Ungleichungen der Analysis folgen aus der

23.7 Ungleichung von Jensen

Sei I ein Intervall und $f : I \to \mathbb{R}$ konvex. Sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ mit $\lambda_1 + \ldots + \lambda_n = 1$ sowie $n \geq 2$, so gilt für beliebige $t_1, \ldots, t_n \in I$:

(i) $\lambda_1 t_1 + \ldots + \lambda_n t_n \in I$.

(ii) $f(\lambda_1 t_1 + \ldots + \lambda_n t_n) \leq \lambda_1 f(t_1) + \ldots + \lambda_n f(t_n)$.

(iii) Ist f streng konvex, so gilt die Gleichheit nur für $t_1 = \ldots = t_n$.

Beweis. Wir führen den Beweis von (i)–(iii) simultan, und zwar mit vollständiger Induktion.

(A) Sei $n = 2$. Ist dann $t_1 = t_2$, so sind wegen $\lambda_1 + \lambda_2 = 1$ die Aussagen (i)–(iii) erfüllt. Sei nun $t_1 \neq t_2$. Dann nehmen wir $t_1 < t_2$ an; im Falle $t_1 > t_2$ numeriere man um.

Dann ist

$$t := \lambda_1 t_1 + (1 - \lambda_1) t_2 \in]t_1, t_2[\subset I \tag{1}$$

nach 2.5. Wegen $\lambda_2 = 1 - \lambda_1$ folgt also (i).

Aus (1) folgt

(2) $$t - t_1 \underset{(1)}{=} (1-\lambda_1)(t_2 - t_1) = \lambda_2(t_2 - t_1).$$

Da f konvex ist gilt (mit $<$ bei strenger Konvexität):

$$f(\lambda_1 t_1 + \lambda_2 t_2) \underset{(1)}{=} f(t) \underset{(<)}{\leq} f(t_1) + \tfrac{f(t_2)-f(t_1)}{t_2-t_1}(t - t_1) \underset{(2)}{=} f(t_1) + \lambda_2(f(t_2) - f(t_1))$$
$$= \lambda_1 f(t_1) + \lambda_2 f(t_2).$$

Also gelten (ii) und (iii) für $n = 2$.

(S) Seien nun $n \geq 2$ und für den Schluß von n auf $n+1$ nicht alle t_i gleich. Setze

(3) $$\lambda := \lambda_1 + \ldots + \lambda_n \text{ und } t_1' := \tfrac{\lambda_1}{\lambda} t_1 + \ldots + \tfrac{\lambda_n}{\lambda} t_n.$$

Wegen $\frac{\lambda_i}{\lambda} \in \mathbb{R}_+$ und $\sum_{i=1}^n \frac{\lambda_i}{\lambda} = 1$ ist nach Induktionsvoraussetzung:

(4) $$t_1' \in I;$$

(5) $$f(t_1') \leq \textstyle\sum_{i=1}^n \frac{\lambda_i}{\lambda} f(t_i)$$

mit $<$, falls f streng konvex und nicht alle $t_1, \ldots, t_n$ gleich sind.

Nun ist wegen (4) und $t_{n+1} \in I$ nach Induktionsanfang

(6) $$\lambda_1 t_1 + \ldots + \lambda_n t_n + \lambda_{n+1} t_{n+1} \underset{(3)}{=} (\lambda_1 + \ldots + \lambda_n) t_1' + \lambda_{n+1} t_{n+1} \in I.$$

Also gilt (i) für $n+1$.

Wiederum nach Induktionsanfang (für (ii)) gilt:

$$f(\lambda_1 t_1 + \ldots + \lambda_{n+1} t_{n+1}) \underset{(6)}{=} f((\lambda_1 + \ldots + \lambda_n) t_1' + \lambda_{n+1} t_{n+1})$$
$$\leq (\lambda_1 + \ldots + \lambda_n) f(t_1') + \lambda_{n+1} f(t_{n+1}) \underset{(5)}{\leq} \textstyle\sum_{i=1}^{n+1} \lambda_i f(t_i).$$

Daher ist (ii) für $n+1$ bewiesen. Sei nun f streng konvex. Sind nicht alle t_1 bis t_n gleich, so steht in der letzten Ungleichung wegen (5) ein $<$. Also erhalten wir in diesem Falle (iii). Ist f streng konvex und sind $t_1, \ldots, t_n$ alle gleich, aber $t_{n+1} \neq t_1$, so ist $t_{n+1} \neq t_1' (= t_1)$, und daher steht nach Induktionsanfang (für (iii)) in der vorletzten Gleichung ein $<$. Also erhalten wir auch in diesem Fall (iii). □

23.8 Ungleichung zwischen dem arithmetischen und geometrischen Mittel

Seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ mit $\lambda_1 + \ldots + \lambda_n = 1$. Dann gilt für beliebige $t_1, \ldots, t_n \in \mathbb{R}_+$

$$t_1^{\lambda_1} \cdot \ldots \cdot t_n^{\lambda_n} \leq \lambda_1 t_1 + \ldots + \lambda_n t_n.$$

Also insbesondere

$$\sqrt[n]{t_1 \cdot \ldots \cdot t_n} \leq \tfrac{t_1 + \ldots + t_n}{n}.$$

Das Gleichheitszeichen gilt jedoch nur für $t_1 = \ldots = t_n$.

Beweis. Es ist $-ln(x)$ streng konvex auf $\mathbb{R}_+$. Also gilt nach der Ungleichung von Jensen

$$-ln(\lambda_1 t_1 + \ldots + \lambda_n t_n) \leq -\lambda_1 ln(t_1) - \ldots - \lambda_n ln(t_n)$$

mit $<$, falls nicht alle $t_1, \ldots, t_n$ gleich sind.

Somit ist

$$ln(t_1^{\lambda_1} \cdot \ldots \cdot t_n^{\lambda_n}) = \sum_{i=1}^{n} \lambda_i ln(t_i) \leq ln(\sum_{i=1}^{n} \lambda_i t_i)$$

mit $<$, falls nicht alle $t_1, \ldots, t_n$ gleich sind.

Anwendung der streng monotonen Exponentialfunktion auf diese Ungleichung ergibt die Behauptung für den allgemeinen Fall. Der Spezialfall ergibt sich mit $\lambda_1 = \ldots = \lambda_n = 1/n$. □

23.9 Die p-Normen im $\mathbb{R}^n$

Sei $p \geq 1$. Setze dann für $u := (u_1, \ldots, u_n) \in \mathbb{R}^n$

$$\| u \|_p := \sqrt[p]{\textstyle\sum_{i=1}^{n} |u_i|^p}.$$

Dann gilt:

(i) $\| u \|_p \geq 0$ und $\| u \|_p = 0 \Longleftrightarrow u = 0$;

(ii) $\| \alpha u \|_p = |\alpha| \; \| u \|_p$ für jedes $\alpha \in \mathbb{R}$.

$\sqrt[p]{\sum_{i=1}^{n} |u_i|^p}$ ist hierbei definiert durch $(\sum_{i=1}^{n} |u_i|^p)^{1/p}$.

23.10 Höldersche Ungleichung

Sei $p > 1$ und q durch $1/p + 1/q = 1$ (d.h. $q = \frac{p}{p-1}$) gegeben. Dann gilt für beliebige $u := (u_1, \ldots, u_n), v := (v_1, \ldots, v_n) \in \mathbb{R}^n$

$$\sum_{i=1}^{n} |u_i v_i| \; \leq \| u \|_p \; \| v \|_q \,.$$

Beweis. Es genügt, den Fall $u \neq 0$ und $v \neq 0$ zu behandeln. Nach 23.9(i) sind dann $\| u \|_p \neq 0$ und $\| v \|_p \neq 0$. Nach Ungleichung 23.8 folgt mit $\lambda_1 := 1/p$, $\lambda_2 := 1/q$ und $t_1 := \frac{|u_i|^p}{\|u\|_p^p}, t_2 := \frac{|v_i|^q}{\|v\|_q^q}$ (der Fall $t_1 = 0$ oder $t_2 = 0$ ist trivial):

$$\text{(1)} \qquad \frac{|u_i \cdot v_i|}{\|u\|_p \|v\|_q} \leq \frac{1}{p} \frac{|u_i|^p}{\|u\|_p^p} + \frac{1}{q} \frac{|v_i|^q}{\|v\|_q^q}.$$

Durch Summation ergibt sich dann

$$\frac{1}{\|u\|_p \|v\|_q} \sum_{i=1}^{n} |u_i v_i| \underset{(1)}{\leq} \frac{1}{p} \frac{1}{\|u\|_p^p} \sum_{i=1}^{n} |u_i|^p + \frac{1}{q} \cdot \frac{1}{\|v\|_q^q} \sum_{i=1}^{n} |v_i|^q = \frac{1}{p} + \frac{1}{q} = 1,$$

also durch Multiplikation mit $\| u \|_p \cdot \| v \|_q$ die behauptete Ungleichung. □

23.11 Die Minkowskische Ungleichung

Sei $p \geq 1$, dann gilt für $u, v \in \mathbb{R}^n$

$$\| u+v \|_p \leq \| u \|_p + \| v \|_p \,.$$

Beweis. Für $p = 1$ folgt die Behauptung unmittelbar aus der Dreiecksungleichung für reelle Zahlen. Seien also

(1) $$p > 1 \text{ und } \| u+v \|_p > 0.$$

Mit $w_i := |u_i + v_i|^{p-1},\ w := (w_1, \ldots, w_n)$ und

(2) $$q := p/(p-1), \text{ also } \quad 1/p + 1/q = 1$$

gilt

(3) $$\begin{cases} \| u+v \|_p^p & = \sum_{i=1}^n |u_i + v_i|^p = \sum_{i=1}^n |(u_i+v_i)w_i| \\ & \leq \sum_{i=1}^n |u_i w_i| + \sum_{i=1}^n |v_i w_i| \underset{23.10}{\leq} \| u \|_p \| w \|_q + \| v \|_p \| w \|_q \,. \end{cases}$$

Da $w_i^q = |u_i + v_i|^{(p-1)q} = |u_i + v_i|^p$ ist, gilt

(4) $$\begin{aligned} \| w \|_q &= (\sum_{i=1}^n |w_i|^q)^{1/q} = (\sum_{i=1}^n |u_i + v_i|^p)^{1/q} \\ &= (\| u+v \|_p^p)^{1/q} = \| u+v \|_p^{p-1} \,. \end{aligned}$$

Also ist

$$\| u+v \|_p^p \underset{(3)}{\leq} (\| u \|_p + \| v \|_p) \| w \|_q \underset{(4)}{=} (\| u \|_p + \| v \|_p)(\| u+v \|_p^{p-1}).$$

Division durch $0 \underset{(1)}{<} \| u+v \|_p^{p-1}$ liefert die Behauptung. □