Die Jensensche Ungleichung

Hans-Gert Gräbe, Univ. Leipzig

3. Februar 1998

1 Konvexe und konkave Funktionen

Wir betrachten eine stetige Funktion $y = f(x)$, die auf einem offenen Intervall $]a, b[$ definiert sein möge. Eine solche Funktion können wir in einem x-y-Koordinatensystem zeichnen, wobei man die Menge

$$\Gamma_f := \{(x, f(x)) \ : \ x \in]a, b[\}$$

als den *Graphen* von f bezeichnet.

Wir nennen eine solche Funktion f *konvex* oder *nach unten gekrümmt*, wenn jede Sekante, die zwei Funktionswerte auf dem Graphen von f verbindet, vollständig oberhalb Γ_f verläuft.

Wir nennen eine solche Funktion f *konkav* oder *nach oben gekrümmt*, wenn jede Sekante, die zwei Funktionswerte auf dem Graphen von f verbindet, vollständig unterhalb Γ_f verläuft.

Beispiele für derartige Funktionen, deren Graph aus dem Schulunterricht gut bekannt ist.

1. $f(x) = \log(x)$. Die Logarithmusfunktion (zu einer Basis $b > 1$) ist in ihrem gesamten Definitionsbereich $x > 0$ nach oben gekrümmt, also *konkav*.

2. $f(x) = x^k$. Die Potenzfunktion (mit $k > 1$) ist für $x \geq 0$ nach unten gekrümmt, also *konvex*.

3. $f(x) = \frac{1}{x}$. Diese Funktion ist für $x > 0$ nach unten gekrümmt, also *konvex*.

4. $f(x) = \sin(x)$. Die Sinusfunktion ist im Bereich $[0, \pi]$ nach oben gekrümmt, also *konkav*, im Bereich $[\pi, 2\pi]$ dagegen nach unten gekrümmt, also *konvex*.

5. $f(x) = \tan(x)$. Die Tangensfunktion ist im Bereich $]-\frac{\pi}{2}, 0]$ konkav und im Bereich $[0, \frac{\pi}{2}[$ konvex.

Für genügend gute Funktionen, deren Graph uns nicht so geläufig ist wie die Graphen der bisher aufgeführten Funktionen, kann man das Krümmungsverhalten auch mit Mitteln der Differentialrechnung bestimmen. Offensichtlich trennen die Wendepunkte der Funktion f die Bereiche unterschiedlichen Krümmungsverhaltens voneinander. Ist eine Funktion f (glatt und) nach oben gekrümmt und legen wir die Tangente an verschiedene Punkte des Graphen von f, so sehen wir, dass der Anstieg dieser Tangenten, also der Wert von $f'(x)$, für wachsendes x immer kleiner wird. f' muss also eine monoton fallende Funktion sein. Eine monoton fallende Funktion erkennen wir aber daran, dass deren Ableitung, also hier $f''(x)$, kleiner als 0 ist. Ähnlich verhält es sich mit konvexen Funktionen.

Satz 1 *Eine (genügend oft stetig differenzierbare) Funktion f ist genau dann in einem Intervall $]a, b[$ konvex (konkav), wenn dort ihre erste Ableitung f' monoton wächst (fällt). Das gilt genau dann, wenn ihre zweite Ableitung $f''(x)$ im gesamten Intervall ≥ 0 (≤ 0) ist.*

Beispiel: Betrachten wir die Funktion $f(x) = \sin(x) + \tan(x)$ im Intervall $[0, \frac{\pi}{2}[$. Als Summe einer nach oben gekrümmten und einer nach unten gekrümmten Funktion ist das Krümmungsverhalten nicht auf den ersten Blick zu erkennen. Allerdings gilt

$$f''(x) = -\sin(x) + 2\tan(x)^3 + 2\tan(x) \geq 0,$$

da bekanntlich $\sin(x) \leq \tan(x)$ ist. Die Funktion f ist also konvex.

2 Die Jensensche Ungleichung

Sei $y = f(x)$ eine im Intervall $]a, b[$ (streng) konvexe Funktion, $x_1, x_2 \in]a, b[$ und $y_1 = f(x_1)$, $y_2 = f(x_2)$ die zugehörigen Funktionswerte. Dann verläuft die Sekante von $P_1 = (x_1, y_1)$ nach $P_2 = (x_2, y_2)$ vollkommen oberhalb des Graphen von f. Insbesondere liegt der Mittelpunkt $S = (\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$ der Sekante P_1P_2 *oberhalb* des Funktionswerts $f(\frac{x_1+x_2}{2})$ an der Stelle $\frac{x_1+x_2}{2}$. Mithin gilt *für eine konvexe Funktion* stets

$$f(\frac{x_1 + x_2}{2}) \leq \frac{f(x_1) + f(x_2)}{2}$$

und analog *für eine konkave Funktion*

$$f(\frac{x_1 + x_2}{2}) \geq \frac{f(x_1) + f(x_2)}{2}.$$

Mehr noch, Gleichheit gilt in beiden Fällen nur dann, wenn die Sekante entartet, also für $x_1 = x_2$.

Diese recht einfachen Überlegungen führten zu zwei Ungleichungen, die in vielen Anwendungen ein wichtiges Hilfsmittel sein können. Bevor wir im nächsten Punkt derartige Anwendungen betrachten, wollen wir diese Überlegungen noch verallgemeinern.

Betrachten wir dazu n verschiedene Argumente $x_1, x_2, \ldots, x_n$ aus dem gegebenen Intervall, deren Funktionswerte $y_i = f(x_i)$, $i = 1, \ldots, n$ und die zugehörigen Punkte $P_i = (x_i, y_i)$ auf dem Graphen der Funktion f. Diese Punkte spannen ein n-Eck auf, dessen Schwerpunkt S bekanntlich die Koordinaten

$$S = \left(\frac{x_1 + x_2 + \ldots + x_n}{n}, \frac{y_1 + y_2 + \ldots + y_n}{n} \right)$$

hat. Ist f wiederum eine konvexe Funktion, dann liegt zusammen mit jeder der Sekanten, die zwei der Punkte verbinden, auch das gesamte n-Eck *oberhalb* des Graphen von f. Insbesondere liegt der Schwerpunkt S *oberhalb* des Punktes auf dem Graphen mit derselben Abszisse.

Es gilt also

Satz 2 (Jensensche Ungleichung) *Ist f eine im Intervall $]a,b[$ konvexe Funktion und $x_1,\ldots,x_n$ aus diesem Intervall, so gilt die Ungleichung*

$$f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right) \le \frac{f(x_1)+f(x_2)+\ldots+f(x_n)}{n}.$$

Analog gilt für eine konkave Funktion f die Ungleichung

$$f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right) \ge \frac{f(x_1)+f(x_2)+\ldots+f(x_n)}{n}.$$

Gleichheit gilt (für streng konvexes bzw. konkaves f) wiederum nur dann, wenn alle Argumente x_i übereinstimmen.

3 Anwendungen

Als erste Anwendung wollen wir zeigen, dass in jedem Dreieck für die Innenwinkel

$$\sin(\alpha)+\sin(\beta)+\sin(\gamma) \le \frac{3\sqrt{3}}{2}$$

gilt.
Diese Ungleichung ergibt sich als eine einfache Folgerung aus der Jensenschen Ungleichung, denn $\sin(x)$ ist im Intervall $[0,\pi]$ konkav. Also gilt

$$\frac{\sin(\alpha)+\sin(\beta)+\sin(\gamma)}{3} \le \sin\left(\frac{\alpha+\beta+\gamma}{3}\right) = \sin(60^0) = \frac{\sqrt{3}}{2}.$$

Gleichheit gilt bei dieser Ungleichung nur für das gleichseitige Dreieck.

Aufgabe 1 Zeigen Sie, dass in einem Dreieck für die Innenwinkel stets

$$\tan^2(\frac{\alpha}{2})+\tan^2(\frac{\beta}{2})+\tan^2(\frac{\gamma}{2}) \ge 1$$

gilt.

In einem Dreieck gilt für die Innenwinkel stets auch

$$\sin(\frac{\alpha}{2})\cdot\sin(\frac{\beta}{2})\cdot\sin(\frac{\gamma}{2}) \le \frac{1}{8}$$

Hier betrachten wir die Funktion $f(x) = \log(\sin(x))$ im Intervall $]0,\frac{\pi}{2}]$. Aus der zweiten Ableitung $f''(x) = -\frac{1}{\sin(x)^2}$ (oder durch genauere Analyse des Graphen) erkennen wir, dass diese Funktion im ganzen Definitionsbereich konkav ist, d.h. dass

$$\log(\sin(\frac{\alpha}{2}))+\log(\sin(\frac{\beta}{2}))+\log(\sin(\frac{\gamma}{2})) \le 3\,\log(\sin(30^0)) = \log(2^{-3})$$

gilt, woraus die Behauptung unmittelbar folgt (denn $\log(x)$ ist ja eine monotone wachsende Funktion).

Aus der Jensenschen Ungleichung lassen sich auch viele andere wohlbekannte Ungleichungen herleiten. Betrachen wir z.B. die Funktion $f(x) = \log(x)$, die ja im ganzen Definitionsbereich $x > 0$ konkav ist, so ergibt sich nach den Logarithmengesetzen

$$\log\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right) \geq \frac{\log(x_1) + \log(x_2) + \ldots + \log(x_n)}{n} = \log(\sqrt[n]{x_1\, x_2\, \ldots\, x_n}),$$

also die **Ungleichung vom arithmetisch-geometrischen Mittel**.

Aufgabe 2 Leiten Sie unter Verwendung der Potenzfunktion als f die Ungleichung vom quadratischen Mittel und vom harmonischen Mittel her.

Auch kompliziertere Ungleichungen lassen sich auf die Jensensche zurückführen. Die Aufgabe 061243 der 6. Mathematikolympiade lautete etwa

> $a_1, \ldots, a_n$ seien positive reelle Zahlen und $s = \sum_{i=1}^{n} a_i$ deren Summe. Beweisen Sie die Ungleichung
>
> $$\Sigma := \sum_{i=1}^{n} \frac{a_i}{s - a_i} \geq \frac{n}{n-1}$$

Aufgabe 3 Zeigen Sie die Ungleichung für $n = 3$, d.h.

$$\frac{a_1}{a_2 + a_3} + \frac{a_2}{a_2 + a_3} + \frac{a_3}{a_1 + a_2} \geq \frac{3}{2},$$

ohne die Jensensche Ungleichung zu verwenden.

Für den Beweis der Ungleichung für eine beliebige natürliche Zahl n betrachten wir die Funktion $f(x) = \frac{x}{s-x}$ im Intervall $0 < x < s$. Da deren zweite Ableitung $f''(x) = \frac{2\,s}{(s-x)^3}$ überall positiv, f also konvex ist, erhalten wir

$$f\left(\frac{a_1 + a_2 + \ldots + a_n}{n}\right) = f\left(\frac{s}{n}\right) \leq \frac{f(x_1) + f(x_2) + \ldots + f(x_n)}{n} = \frac{\Sigma}{n}.$$

Daraus ergibt sich die Behauptung unmittelbar.

4 Eine weitere Verallgemeinerung

Betrachten wir noch einmal das von den Punkten $P_i = (x_i, y_i)$ aufgespannte n-Eck. Zum Beweis der Jensenschen Ungleichung haben wir einzig die Tatsache verwendet, dass der Schwerpunkt S für eine konvexe Funktion f *oberhalb* des Graphen von f liegt. Das ist aber für *jeden* Punkt aus dem Inneren des n-Ecks richtig. Die x-y-Koordinaten eines solchen Punktes kann man aus seinen *baryzentrischen Koordinaten* bzgl. $P_1, \ldots, P_n$ bestimmen. Dazu versehen wir die Eckpunkte mit nichtnegativen Gewichten $w_1, \ldots, w_n \in \mathbb{R}$, die in der Summe gerade 1 ergeben. Der Punkt P mit den Koordinaten

$$P = w_1 P_1 + \ldots + w_n P_n = (w_1 x_1 + \ldots + w_n x_n,\ w_1 y_1 + \ldots + w_n y_n)$$

ist dann ein Punkt im Inneren des n-Ecks und liegt somit oberhalb des Punkts auf dem Graphen von f mit derselben Abszisse. Wir erhalten als Verallgemeinerung der oben bewiesenen Ungleichung den folgenden

Satz 3 (Gewichtete Jensensche Ungleichung) *Ist f eine im Intervall $]a,b[$ konvexe Funktion, $x_1,\ldots,x_n$ aus diesem Intervall und $w_1,\ldots,w_n$ nichtnegative Gewichte mit $w_1+\ldots+w_n=1$, so gilt die Ungleichung*

$$f\left(w_1x_1+w_2x_2+\ldots+w_nx_n\right)\leq w_1f(x_1)+w_2f(x_2)+\ldots+w_nf(x_n).$$

Analog gilt für eine konkave Funktion f die Ungleichung

$$f\left(w_1x_1+w_2x_2+\ldots+w_nx_n\right)\geq w_1f(x_1)+w_2f(x_2)+\ldots+w_nf(x_n).$$

Gleichheit gilt (für streng konvexes bzw. konkaves f) wiederum nur dann, wenn alle Argumente x_i, deren Gewicht w_i positiv ist, übereinstimmen.