5.3 Konvexe Funktionen

Def.: Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $f: I \to \mathbb{R}$ heißt konvex, wenn für alle $x, y \in I$ und $\lambda \in (0,1)$ gilt

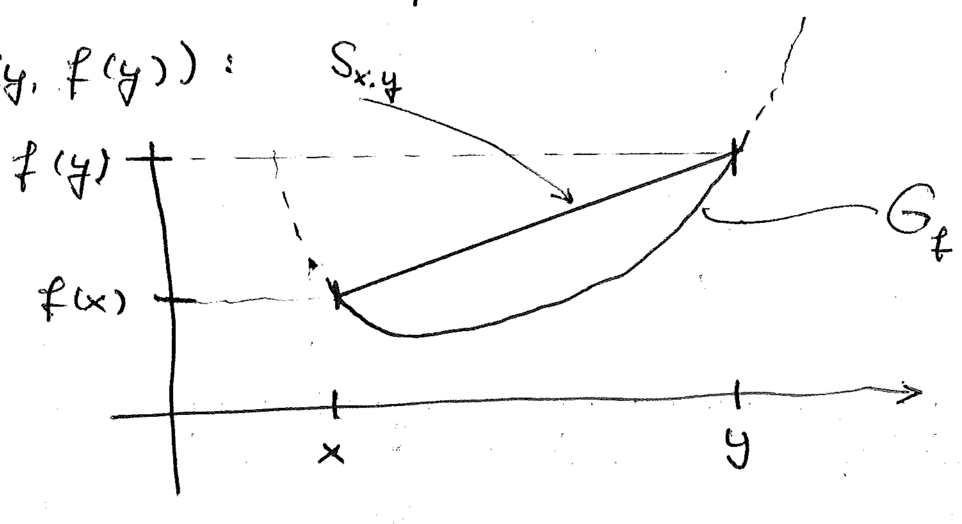
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \qquad (K)$$

f heißt streng konvex, wenn (K) mit strikter Ungleichung "<" gilt; f heißt (streng) konkav, wenn $-f$ (streng) konvex ist.

Geometrische Interpretation: Die Begriffe 'konvex' und 'konkav' charakterisieren das Krümmungsverhalten des Graphen G_f. Man sagt, G_f sei nach links (bzw. rechts) gekrümmt, wenn f konvex (bzw. konkav) ist. Bei dieser Sprechweise wird angenommen, dass G_f in Richtung wachsenden Arguments x durchlaufen wird. Die Menge

$$S_{x,y} := \{ \lambda(x, f(x)) + (1-\lambda)(y, f(y)) \in \mathbb{R}^2 : 0 \leq \lambda \leq 1 \}$$

ist die Sekante zu G_f durch die Punkte $(x, f(x))$ und $(y, f(y))$:



Die Ungleichung (K) in der Def. von "konvex" bedeutet geometrisch also gerade die Forderung, dass alle Sekanten durch G_f oberhalb dieses Graphen verlaufen. Entsprechend liegen alle Sekanten durch den Graphen einer konkaven Funktion f unterhalb von G_f.

Um den Bezug zur Differenzierbarkeit herzustellen, schreiben wir die Bedingung (K) etwas um:

Lemma 1: $f: I \to \mathbb{R}$ ist genau dann konvex, wenn für alle $x, y, z \in I$ mit $x < z < y$ die Ungleichung

$$\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(z)}{y-z} \qquad (1)$$

gilt. In diesem Fall ist

$$\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(z)}{y-z}. \qquad (2)$$

Bem.: $f: I \to \mathbb{R}$ ist genau dann streng konvex, wenn die Ungleichungen (1) und (2) mit strikter Ungleichheit "<" gelten.

Bew.: (K) $\Rightarrow$ (2). Sind $x < z < y$ in I gegeben, so existiert genau ein $\lambda \in (0,1)$, so dass $z = \lambda x + (1-\lambda) y$. Hierfür ist dann

$$z = x + (1-\lambda)(y-x) = \lambda(x-y) + y,$$

so dass

$$(1-\lambda) = \frac{z-x}{y-x} \qquad \text{und} \qquad \lambda = \frac{y-z}{y-x}.$$

Nun folgt aus (K)

$$f(z) \leq \lambda f(x) + (1-\lambda) f(y) = f(x) + (1-\lambda)(f(y)-f(x)) \quad (1.1)$$

$$= \lambda (f(x)-f(y)) + f(y) \quad (1.2)$$

Umstellung von (1.1) unter Beachtung von $(1-\lambda) = \frac{z-x}{y-x}$ ergibt

$$\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(x)}{y-x}$$

und aus (1.2) erhält man mit $\lambda = \frac{y-z}{y-x}$

$$\frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(z)}{y-z}.$$

Zusammen ergibt sich die Ungleichungskette (2), die (1) enthält.

(1) $\Rightarrow$ (K): Sind $x, y \in I$ mit $x < y$ und $\lambda \in (0,1)$ gegeben, wählt man $z = \lambda x + (1-\lambda) y$. Wegen $1-\lambda = \frac{z-x}{y-x}$ und $\lambda = \frac{y-z}{y-x}$ (s.o.) folgt dann aus (1) nach Division durch $y-x$

$$\lambda (f(z)-f(x)) \leq (1-\lambda)(f(y)-f(z)) \quad \Big| \begin{array}{l} + \lambda f(x) \\ + (1-\lambda) f(z) \end{array}$$

$$\Rightarrow \quad f(z) \leq (1-\lambda) f(y) + \lambda \cdot f(x).$$

□

Wie regulär sind konvexe Funktionen? Zwei Beispiele:

(1) $f: [-1,1] \to \mathbb{R},\ x \mapsto f(x) := \begin{cases} x^2 & : |x| < 1 \\ 2 & : |x| = 1 \end{cases}$

ist in $(-1,1)$ stetig, aber in den Randpunkten $x = \pm 1$ unstetig. f ist konvex, sogar streng konvex.

(2) $f: \mathbb{R} \to \mathbb{R},\ x \mapsto \underbrace{f(x)}_{:= |x|}$ ist konvex, aber in $x_0 = 0$ nicht differenzierbar. Dort existieren aber immerhin noch die einseitigen Ableitungen

$$D_+ f(0) = \lim_{\substack{h \to 0 \\ h > 0}} \frac{1}{h}(f(h) - f(0)) = 1 \quad \text{und}$$

$$D_- f(0) = \lim_{\substack{h \to 0 \\ h < 0}} \frac{1}{h}(f(h) - f(0)) = -1$$

Damit ist der Rahmen des möglichen bereits einigermaßen abgesteckt.

<u>Satz 1</u> (über die Regularität konvexer Funktionen):

Es seien $-\infty \le a < b \le \infty$ und $f: (a,b) \to \mathbb{R}$ konvex.

Dann gelten

(1) f ist stetig.

(2) In jedem $x \in (a,b)$ existieren die einseitigen Ableitungen

$$D_- f(x) := \lim_{\substack{z \to x \\ z < x}} \frac{f(z) - f(x)}{z - x} \quad \text{und}$$

$$D_+ f(x) := \lim_{\substack{z\to x\\ z>x}} \frac{f(z)-f(x)}{z-x}.$$

(3) Für alle $x,y \in (a,b)$ mit $x<y$ ist

$$D_- f(x) \le D_+ f(x) \le D_- f(y) \le D_+ f(y).$$

Falls f streng konvex ist, gilt $D_+ f(x) < D_- f(y)$.

Bew.: Die Stetigkeit folgt aus der Differenzierbarkeitsaussage in (2), vgl. Bem. (6) zur Definition der Ableitung. Um (2) in einem beliebigen $x \in (a,b)$ zu zeigen, wählen wir Punkte r,t,z,y mit $a<r<t<x<z<y<b$. Dann gibt Lemma 1 die folgenden Ketten von Ungleichungen:

für $x<z<y \longrightarrow$ $$\frac{f(z)-f(x)}{z-x} \le \frac{f(y)-f(x)}{y-x} \le \frac{f(y)-f(z)}{y-z}$$

$$\frac{f(x)-f(t)}{x-t} \le \frac{f(z)-f(t)}{z-t} \le \frac{f(z)-f(x)}{z-x}$$ $\longleftarrow$ für $t<x<z$

$$\frac{f(t)-f(r)}{t-r} \le \frac{f(x)-f(r)}{x-r} \le \frac{f(x)-f(t)}{x-t}$$ $\longleftarrow$ für $r<t<x$

Hiervon benötigen wir ~~[illegible]~~ nur die letzte und diejenigen, in denen x auftritt, zusammengefasst also

$$\frac{f(r)-f(x)}{r-x} \underset{1}{\le} \frac{f(t)-f(x)}{t-x} \underset{2}{\le} \frac{f(z)-f(x)}{z-x} \underset{3}{\le} \frac{f(y)-f(x)}{y-x} \underset{4}{\le} \frac{f(y)-f(z)}{y-z}.$$

1 und 2 sagen aus, dass die Abbildung

$$t \longmapsto \frac{f(t)-f(x)}{t-x}$$

auf (a,x) monoton steigt und für jedes $z \in (x,b)$

durch $\frac{f(z)-f(x)}{z-x}$ beschränkt ist. Daher existiert der (K6)

Grenzwert $$\lim_{\substack{t\to x\\ t<x}} \frac{f(t)-f(x)}{t-x} = D_- f(x)$$

und für jedes $z\in(x,b)$ gilt $D_- f(x) \leq \frac{f(z)-f(x)}{z-x}$. (3)

liefert nun, dass die Abbildung

$$z \longmapsto \frac{f(z)-f(x)}{z-x}$$

auf (x,b) monoton steigt, sodass für $z\searrow x$ die Ausdrücke $\frac{f(z)-f(x)}{z-x}$ monoton fallen und nach unten durch $D_- f(x)$ beschränkt sind. Also existiert auch

$$\lim_{\substack{z\to x\\ z>x}} \frac{f(z)-f(x)}{z-x} = D_+ f(x) \geq D_- f(x),$$

womit bereits die erste Teilaussage von (3) gezeigt ist. $z\nearrow y$ in (4) und $z\searrow x$ in (3) liefern

$$D_+ f(x) \leq \frac{f(y)-f(x)}{y-x} \leq D_- f(y),$$

wobei (wegen der Monotonie des Limites) die Ungleichungen strikt sind, wenn f streng konvex ist. $\square$

Folgerungen: (1) Ist $f:(a,b)\to\mathbb{R}$ differenzierbar und (streng) konvex, so ist $f':(a,b)\to\mathbb{R}$ (streng) monoton steigend.

(2) Ist $f:(a,b)\to\mathbb{R}$ konkav, so ist f stetig und

in jedem $x \in (a,b)$ existieren die einseitigen Ableitungen $D_- f(x)$ und $D_+ f(x)$. Hierfür gilt $D_- f(x) \geq D_+ f(x)$. (K7)

(3) Ist $f: (a,b) \to \mathbb{R}$ differenzierbar und (streng) konkav, so ist $f': (a,b) \to \mathbb{R}$ (streng) monoton fallend.

Lemma 2: Sei $-\infty \leq a < b \leq \infty$ und $f: (a,b) \to \mathbb{R}$ differenzierbar. Dann gelten:

(1) f ist genau dann (streng) konvex, wenn $f': (a,b) \to \mathbb{R}$ (streng) monoton steigt.

(2) f ist genau dann (streng) konkav, wenn $f': (a,b) \to \mathbb{R}$ (streng) monoton fällt.

Bew.: (2) folgt aus (1) durch Übergang von f zu $-f$. Die Implikation "$\Rightarrow$" in (1) wurde bereits als Folgerung (1) aus Satz 1 festgestellt. Seien also $f': (a,b) \to \mathbb{R}$ (streng) monoton steigend und $x < z < y$, alle in (a,b). Dann gibt es nach dem MWS $\xi \in (x,z)$ und $\eta \in (z,y)$, so dass

$$\frac{f(z)-f(x)}{z-x} = f'(\xi) \underset{(<)}{\leq} f'(\eta) = \frac{f(y)-f(z)}{y-z}.$$

Nach Lemma 1 ist also f (streng) konvex. $\square$

Setzen wir schließlich f als zweimal differenzierbar voraus und verbinden das Bisherige mit dem Monotoniesatz (KS) (anzuwenden auf f'), so ergibt sich das folgende gut handhabbare Kriterium für Konvexität:

Satz 2 (Konvexitätssatz). Es seien $-\infty \leq a < b \leq \infty$ und $f:(a,b)\to\mathbb{R}$ zweimal differenzierbar. Dann gelten:

(1) f ist konvex genau dann, wenn $f''(x)\geq 0$ für alle $x\in(a,b)$ ist.

(2) f ist streng konvex, wenn $f''(x)>0$ ist für alle $x\in(a,b)$.

Gegenbeispiel für "$\not\Rightarrow$" in (2): $f:\mathbb{R}\to\mathbb{R}$, $x\mapsto x^4$, ist streng konvex, obwohl $f''(0)=0$ gilt.

Anwendung: Youngsche Ungleichung

Die reelle Exponentialfunktion $\exp:\mathbb{R}\to\mathbb{R}$ ist zweimal differenzierbar mit $\exp''(x)=\exp(x)>0$ für alle $x\in\mathbb{R}$ und somit streng konvex auf $\mathbb{R}$. Daraus ergibt sich ein kurzer Beweis der häufig benutzten Youngschen Ungleichung:

Lemma 3: Es seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ und $a, b \geq 0$.

Dann gilt

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q.$$

Bew.: Das ist klar, wenn $ab = 0$, d.h. wenn $a = 0$ oder $b = 0$.

Andernfalls setzen wir

$$x = \ln(a) \qquad \text{und} \qquad y = \ln(b).$$

Dann ist

$$ab = e^x \cdot e^y = e^{x+y} = \exp\Big(\underbrace{\frac{1}{p}}_{\lambda} \cdot px + \underbrace{\frac{1}{q}}_{1-\lambda} qx\Big)$$

$$\underset{\text{Konvexität}}{\leq} \frac{1}{p} \cdot \exp(px) + \frac{1}{q} \exp(qy) = \frac{1}{p} a^p + \frac{1}{q} b^q.$$