Konvexität der Exponentialfunktion

Jendrik Stelzner

9. Dezember 2014

Definition 1. Eine Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ heißt konvex, falls für alle $x, y \in \mathbb{R}$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \text{für alle } \lambda \in [0,1].$$

Übung 1.
Es seien $f, g\colon \mathbb{R} \to \mathbb{R}$ konvex. Zeigen Sie: Ist g monoton steigend, so ist auch $g \circ f$ konvex.

Wir wollen nun zeigen, dass die Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}$ konvex ist. Wir erinnern daran, dass

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{für alle } x \in \mathbb{R},$$

und dass

$$\exp(x+y) = \exp(x)\exp(y) \quad \text{für alle } x, y \in \mathbb{R}. \tag{1}$$

Außerdem ist $\exp(x) > 0$ für alle $x \in \mathbb{R}$.

Übung 2.
Zeigen Sie, dass exp auf $[0, \infty)$ konvex ist, d.h. dass für alle $x, y \in [0, \infty)$ ist

$$\exp(\lambda x + (1-\lambda)y) \leq \lambda \exp(x) + (1-\lambda)\exp(y) \quad \text{für alle } \lambda \in [0,1].$$

(*Hinweis*: Nutzen Sie, dass für alle $k \in \mathbb{N}$ die Abbildung $x \mapsto x^k$ auf $[0, \infty)$ konvex ist.)

Übung 3.
Folgern Sie, dass exp konvex ist. (*Hinweis:* Verschieben Sie die Exponentialfunktion und nutzen Sie (1), um diese Verschiebung in eine Skalierung umzuwandeln.)

Lösung 1.
Es seien $x, y \in \mathbb{R}$ und $\lambda \in [0,1]$. Da f konvex ist, ist

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

Wegen der Monotonie von g ergibt sich damit, dass

$$g(f(\lambda x + (1-\lambda)y)) \leq g(\lambda f(x) + (1-\lambda) f(y)).$$

Durch die Konvexität von g erhalten wir auch, dass

$$g(\lambda f(x) + (1-\lambda) f(y)) \leq \lambda g(f(x)) + (1-\lambda) g(f(y)).$$

Insgesamt erhalten wir damit, dass

$$\begin{aligned}
&(g \circ f)(\lambda x + (1-\lambda)y) \\
&= g(f(\lambda x + (1-\lambda)y)) \\
&\leq g(\lambda f(x) + (1-\lambda) f(y)) \\
&\leq \lambda g(f(x)) + (1-\lambda) g(f(y)) \\
&= \lambda (g \circ f)(x) + (1-\lambda)(g \circ f)(y).
\end{aligned}$$

Lösung 2.
Für alle $k \in \mathbb{N}$ ist die Abbildung

$$[0,\infty) \to \mathbb{R}, x \mapsto x^k$$

konvex, d.h. für alle $x, y \in [0,\infty)$ ist

$$(\lambda x + (1-\lambda)y)^k \leq \lambda x^k + (1-\lambda) y^k \quad \text{für alle } \lambda \in [0,1].$$

Damit ergibt sich, dass für alle $x, y \in [0,\infty)$ und $\lambda \in [0,1]$

$$\begin{aligned}
\exp(\lambda x + (1-\lambda)y) &= \sum_{k=0}^{\infty} \frac{(\lambda x + (1-\lambda)y)^k}{k!} \\
&\leq \sum_{k=0}^{\infty} \frac{\lambda x^k + (1-\lambda)y^k}{k!} \\
&= \lambda \sum_{k=0}^{\infty} \frac{x^k}{k!} + (1-\lambda) \sum_{k=0}^{\infty} \frac{y^k}{k!} \\
&= \lambda \exp(x) + (1-\lambda)\exp(y).
\end{aligned}$$

Lösung 3.
Es seien $x, y \in \mathbb{R}$ beliebig aber fest. Es sei $z \in \mathbb{R}$, so dass $x+z, y+z \geq 0$, etwa $z := \max\{|x|, |y|\}$. Für alle $\lambda \in [0,1]$ ist dann

$$\begin{aligned}
&\exp(\lambda x + (1-\lambda)y) \\
&= \exp(-z)\exp(z)\exp(\lambda x + (1-\lambda)y) \\
&= \exp(-z)\exp(z + \lambda x + (1-\lambda)y) \\
&= \exp(-z)\exp(\lambda(x+z) + (1-\lambda)(y+z)) \\
&\leq \exp(-z)\left(\lambda \exp(x+z) + (1-\lambda)\exp(y+z)\right) \\
&= \lambda \exp(-z)\exp(x+z) + (1-\lambda)\exp(-z)\exp(y+z) \\
&= \lambda \exp(x) + (1-\lambda)\exp(y).
\end{aligned}$$

Dabei haben wir genutzt, dass $\exp(-z) > 0$ und $\exp(-z) = 1/\exp(z)$. Wegen der Beliebigkeit von $x, y \in \mathbb{R}$ folgt, dass $\exp\colon \mathbb{R} \to \mathbb{R}$ konvex ist.