We now establish a convexity criterion using the second derivative of f. This criterion is often easier to check than the previous one.

Proposition 40.12. *(Convexity and second derivative) Let $f\colon \Omega \to \mathbb{R}$ be a function twice differentiable on some open subset Ω of a normed vector space E and let $U \subseteq \Omega$ be a nonempty convex subset.*

(1) The function f is convex on U iff

$$\mathrm{D}^2 f(u)(v-u, v-u) \geq 0 \quad \text{for all } u, v \in U.$$

(2) If

$$\mathrm{D}^2 f(u)(v-u, v-u) > 0 \quad \text{for all } u, v \in U \text{ with } u \neq v,$$

then f is strictly convex.

Proof. First assume that the inequality in Condition (1) is satisfied. For any two distinct points $u, v \in U$, the formula of Taylor–Maclaurin yields

$$\begin{aligned} f(v) - f(u) - df(u)(v-u) &= \frac{1}{2}\mathrm{D}^2 f(w)(v-u, v-u) \\ &= \frac{\rho^2}{2}\mathrm{D}^2 f(w)(v-w, v-w), \end{aligned}$$

for some $w = (1-\lambda)u + \lambda v = u + \lambda(v-u)$ with $0 < \lambda < 1$, and with $\rho = 1/(1-\lambda) > 0$, so that $v - u = \rho(v-w)$. Since $\mathrm{D}^2 f(w)(v-w, v-w) \geq 0$ for all $u, w \in U$, we conclude by applying Proposition 40.11(1).

Similarly, if (2) holds, the above reasoning and Proposition 40.11(2) imply that f is strictly convex.

To prove the necessary condition in (1), define $g\colon \Omega \to \mathbb{R}$ by

$$g(v) = f(v) - df(u)(v),$$

where $u \in U$ is any point considered fixed. If f is convex, since

$$g(v) - g(u) = f(v) - f(u) - df(u)(v-u),$$

Proposition 40.11 implies that $f(v) - f(u) - df(u)(v-u) \geq 0$, which implies that g has a local minimum at u with respect to all $v \in U$. Therefore, we have $dg(u) = 0$. Observe that g is twice differentiable in Ω and $\mathrm{D}^2 g(u) = \mathrm{D}^2 f(u)$, so the formula of Taylor–Young yields for every $v = u + w \in U$ and all t with $0 \leq t \leq 1$,

$$\begin{aligned} 0 \leq g(u+tw) - g(u) &= \frac{t^2}{2}\mathrm{D}^2 f(u)(tw, tw) + \|tw\|^2 \epsilon(tw) \\ &= \frac{t^2}{2}(\mathrm{D}^2 f(u)(w, w) + 2\|w\|^2 \epsilon(wt)), \end{aligned}$$

with $\lim_{t \mapsto 0} \epsilon(wt) = 0$, and for t small enough, we must have $\mathrm{D}^2 f(u)(w, w) \geq 0$, as claimed. ∎