40.2 Using Second Derivatives to Find Extrema

For the sake of brevity, we consider only the case of local minima; analogous results are obtained for local maxima (replace J by $-J$, since $\max_u J(u) = -\min_u -J(u)$). We begin with a necessary condition for an unconstrained local minimum.

Proposition 40.5. *Let E be a normed vector space and let $J\colon \Omega \to \mathbb{R}$ be a function, with Ω some open subset of E. If the function J is differentiable in Ω, if J has a second derivative $\mathrm{D}^2 J(u)$ at some point $u \in \Omega$, and if J has a local minimum at u, then*

$$\mathrm{D}^2 J(u)(w, w) \geq 0 \quad \textit{for all } w \in E.$$

Proof. Pick any nonzero vector $w \in E$. Since Ω is open, for t small enough, $u + tw \in \Omega$ and $J(u + tw) \geq J(u)$, so there is some open interval $I \subseteq \mathbb{R}$ such that

$$u + tw \in \Omega \quad \text{and} \quad J(u + tw) \geq J(u)$$

for all $t \in I$. Using the Taylor–Young formula and the fact that we must have $dJ(u) = 0$ since J has a local minimum at u, we get

$$0 \leq J(u + tw) - J(u) = \frac{t^2}{2}\mathrm{D}^2 J(u)(w, w) + t^2 \left\|w\right\|^2 \epsilon(tw),$$

with $\lim_{t \mapsto 0} \epsilon(tw) = 0$, which implies that

$$\mathrm{D}^2 J(u)(w, w) \geq 0.$$

Since the argument holds for all $w \in E$ (trivially if $w = 0$), the proposition is proven. ∎

One should be cautioned that there is no converse to the previous proposition. For example, the function $f\colon x \mapsto x^3$ has no local minimum at 0, yet $df(0) = 0$ and $\mathrm{D}^2 f(0)(u, v) = 0$. Similarly, the reader should check that the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = x^2 - 3y^3$$

has no local minimum at $(0, 0)$; yet $df(0, 0) = 0$ since $df(x, y) = (2x, -9y^2)$, and for $u = (u_1, u_2)$, $\mathrm{D}^2 f(0, 0)(u, u) = 2u_1^2 \geq 0$ since

$$D^2 f(x, y)(u, u) = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -18y \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

See Figure 40.3.

When $E = \mathbb{R}^n$, Proposition 40.5 says that a necessary condition for having a local minimum is that the Hessian $\nabla^2 J(u)$ be positive semidefinite (it is always symmetric).

We now give sufficient conditions for the existence of a local minimum.