Here is a necessary condition for a function to have a local minimum with respect to a convex subset U.

Theorem 40.9. *(Necessary condition for a local minimum on a convex subset) Let $J\colon \Omega \to \mathbb{R}$ be a function defined on some open subset Ω of a normed vector space E and let $U \subseteq \Omega$ be a nonempty convex subset. Given any $u \in U$, if $dJ(u)$ exists and if J has a local minimum in u with respect to U, then*

$$dJ(u)(v-u) \geq 0 \quad \textit{for all } v \in U.$$

Proof. Let $v = u + w$ be an arbitrary point in U. Since U is convex, we have $u + tw \in U$ for all t such that $0 \leq t \leq 1$. Since $dJ(u)$ exists, we can write

$$J(u+tw) - J(u) = dJ(u)(tw) + \|tw\|\, \epsilon(tw)$$

with $\lim_{t\mapsto 0} \epsilon(tw) = 0$. However, because $0 \leq t$,

$$J(u+tw) - J(u) = t(dJ(u)(w) + \|w\|\, \epsilon(tw))$$

and since u is a local minimum with respect to U, we have $J(u+tw) - J(u) \geq 0$, so we get

$$t(dJ(u)(w) + \|w\|\, \epsilon(tw)) \geq 0.$$

The above implies that $dJ(u)(w) \geq 0$, because otherwise we could pick $t > 0$ small enough so that

$$dJ(u)(w) + \|w\|\, \epsilon(tw) < 0,$$

a contradiction. Since the argument holds for all $v = u + w \in U$, the theorem is proven. ☐

Observe that the convexity of U is a substitute for the use of Lagrange multipliers, but we now have to deal with an *inequality* instead of an equality.

In the special case where U is a subspace of E we have the following result.

Corollary 40.10. *With the same assumptions as in Theorem 40.9, if U is a subspace of E, if $dJ(u)$ exists and if J has a local minimum in u with respect to U, then*

$$dJ(u)(w) = 0 \quad \textit{for all } w \in U.$$

Proof. In this case since $u \in U$ we have $2u \in U$, and for any $u + w \in U$, we must have $2u - (u+w) = u - w \in U$. The previous theorem implies that $dJ(u)(w) \geq 0$ and $dJ(u)(-w) \geq 0$, that is, $dJ(u)(w) \leq 0$, so $dJ(u) = 0$. Since the argument holds for $w \in U$ (because U is a subspace, if $u, w \in U$, then $u + w \in U$), we conclude that

$$dJ(u)(w) = 0 \quad \text{for all } w \in U.$$ ☐

We will now characterize convex functions when they have a first derivative or a second derivative.