Real Analysis Exchange Vol. 7 (1981-82)

John L. Troutman, Department of Mathematics,
Syracuse University, Syracuse, N.Y. 13210

Partially Convex Functions in the Variational Calculus

Functions which exhibit an elementary partial convexity provide direct access to the solution of minimization problems from the classical variational calculus. However such convexity is not necessarily preserved under relevant coordinate transformations and a better understanding of how this occurs would be valuable. This article offers motivation for an inquiry.

Various extensions of the following results can be made, but we shall consider here only the standard problem of minimizing

$$F(y) = \int_a^b f[y(x)]dx \overset{\text{def}}{=} \int_a^b f(x,y(x),y'(x))dx$$

on a subset of $\mathcal{Y} = C^1[a,b]$.

Let D be a domain in $\mathbb{R}^2$.

<u>Definition</u>: $f(\underline{x},y,z)$ <u>is said to be</u> [strongly] <u>convex on</u> $S=[a,b]\times D$ <u>if</u> $f = f(x,y,z)$ <u>and its partial derivatives</u> f_y <u>and</u> f_z <u>are continuous on</u> S <u>and satisfy the inequality</u>

$$f(x,y+v,z+w) - f(x,y,z) \geq f_y(x,y,z)v + f_z(x,y,z)w,$$

$$\forall (x,y,z),\ (x,y+v,z+w) \in S,$$

[with equality at (x,y,z) only if $v = 0$ or $w = 0$].

Observe that the set S need <u>not</u> be convex.

As examples, $f(\underline{x},y,z) = y^2$, $z + \sqrt{1+y^2}$, $-\sqrt{1-z^2}$, $\underline{x}^2y^2+(1+\sin\underline{x})z^2$, $\sqrt{1+y^2+z^2}$, are strongly convex on the sets of continuity of f_z, while a function linear in y and z is of course only convex. Also, the sum of a [strongly] convex function with one or more convex functions of this type is again [strongly] convex.

Let $\mathcal{D} = \{y \in \mathcal{Y}\colon (y(x), y'(x)) \in D,\ a \le x \le b\}$.

Proposition. If $f(\underline{x}, y, z)$ is [strongly] convex on S then each solution $y_0 \in \mathcal{D}$ of the differential equation $\frac{d}{dx} f_z[y(x)] = f_y[y(x)]$ on (a,b) minimizes F on

(i) $\mathcal{D}_0 = \{y \in \mathcal{D}: y(a) = y_0(a); y(b) = y_0(b)\}$ [uniquely];

(ii) $\mathcal{D}^b = \{y \in \mathcal{D}: y(a) = y_0(a)\}$, if $f_z[y_0(b)] = 0$ [uniquely];

(iii) $\mathcal{D}$, if $f_z[y_0(a)] = f_z[y_0(b)] = 0$ [uniquely within an additive constant].

[Indeed, if y_0 is as hypothesized and $y_0 + v \in \mathcal{D}$ then the convexity inequality implies

$$F(y_0+v) - F(y_0) \ge \int_a^b (f_y[y_0(x)]v(x) + f_z[y_0(x)]v'(x))dx$$

$$= f_z[y_0(x)]v(x)\Big|_a^b ,$$

and this will vanish when y_0+v is in $\mathcal{D}_0$, or, under the additional conditions imposed, in $\mathcal{D}^b$ or $\mathcal{D}$. With strong convexity, equality is possible only if $v(x)v'(x) \equiv 0$ so that $v(x) = \text{const.}$ which must also vanish when $y_0+v \in \mathcal{D}_0$ or $\mathcal{D}^b$.]

The differential equation is of course that of Euler-Lagrange and the additional conditions are the natural boundary conditions known to be necessary for a minimum. However, this sufficiency proposition is more elementary, and in each instance where a solution y_0 of the stated problem can be found by-passes the far more difficult questions of existence. Some applications of the principles embodied in this proposition were given in [1]. Observe that additional smoothness for y_0 may be assumed as needed.

It is straigntforward to extend these methods to minimization problems involving constraints either of the isoperimetric or Lagrangian form. Again we shall examine here only the case of a single constraint determined by a function $g = g(x,y,z)$ defined on S. Set $\tilde{f} = f + \lambda g$ where $\lambda \in C[a,b]$ is unspecified.

Corollary. If $f(\underline{x},y,z)$ is [strongly] convex on S, then each solution $y_0 \in \mathcal{D}_0$ of the differential equation $\frac{d}{dx}\tilde{f}_z[y(x)] = \tilde{f}_y[y(x)]$ on (a,b) minimizes F on $\mathcal{D}_0$ [uniquely] when restricted to the set $\{y \in \mathcal{D}_0 : \int_a^b \lambda(x)g[y(x)]dx \le \int_a^b \lambda(x)g[y_0(x)]dx\}$.
[For, by the proposition, y_0 minimizes
$\tilde{F}(y) \stackrel{\text{def}}{=} F(y) + \int_a^b \lambda(x)g[y(x)]dx$ on $\mathcal{D}_0$ [uniquely]].

No information is given in this result as to how λ should be selected. However when λ = const. then λ may possibly be chosen so that the λ-dependent solution y_0 meets the isoperimetric constraint $G(y_0) = \int_a^b g[y_0(x)]dx = \ell$. (Although a Lagrangian constraint of the form $g[y_0(x)] \equiv 0$ on (a,b) cannot be fulfilled in general by a solution y_0 of the differential equation, such automatic exclusion disappears in the analogous case when y is a vector valued function, and the approach does permit solutions to Lagrangian problems.)

The corollary furnishes a solution to the catenary or hanging cable problem of Euler, in a non-standard formulation using the arc length s as the independent variable. Then the problem becomes that of minimizing $F(y) = \mu\int_0^L y(s)ds$ (for $\mu>0$), on $\mathcal{D} = \{y \in C^1[0,L] : y(0) = a_1, \ y(L) = b_1\}$ subject to the isoperimetric constraint $G(y) = \int_0^L \sqrt{1-y'(s)^2}\, ds = H$ (where H is the horizontal distance between the supports, so that $H < L$). The modified function $\tilde{f}(\underline{s},y,z) = \mu y - \lambda\sqrt{1-z^2}$ <u>is</u> strongly convex for $\lambda > 0$, and the (unique) solution to the problem may be found by integration. This application motivates the search for integrand functions which transform into those which will be strongly convex when expressed in a new coordinate system - perhaps with accompanying isoperimetric constraints - since the problem as usually formulated does not exhibit the requisite convexity.

In conclusion, we also note that the [strong] convexity of $f(\underline{x},\underline{y},z)$ is precisely the statement that the Weierstrass excess function $\mathcal{E}$ is ≥ 0 [> 0] and this condition is well known to be intimately related to the problem if minimizing F on $\mathcal{D}_0$. Finally, far more sophisticated uses of convexity have been made in attacking general variational problems. [2]. However the elementary approach presented here is remarkably useful in providing at least partial solutions to a wide class of problems.

[1] W. Hrusa and J. L. Troutman. *Elementary characterization of classical minima.* AMS Monthly, May 1981, 321-327.

[2] V. Barbu & Th. Precupanu. *Convexity and Optimization in Banach Spaces.* Bucuresti: Sijthoff & Nordhoff (1978).