Chapter 5

Convex minimization; some sufficient conditions

This is some material from chapters 2 and 3 of Troutman. The point of view concerning the first and second variations is somewhat different in that rather than giving specific conditions under which these variations exist and explicit domains of perturbations on which they are defined, Troutman prefers to allow perturbation in any direction in which the formal definition of the variation leads to a well-defined real value. We will start off with admissible classes in the nominal universal set $C^1[a,b]$, though Troutman suggests the consideration of various other vector spaces as universal sets. In particular, the larger space $C^0[a,b] \cap C^1(a,b)$.

5.1 The basic results

Definition 5. *Given* $F \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$ *and* $\mathcal{F}[u] = \int_a^b F(x,u,u')\,dx$*, we say* $\mathcal{F}$ *is* **convex on** $\mathcal{A} \subset C^1[a,b]$ *if*

$$\mathcal{F}[u+v] - \mathcal{F}[u] \geq \delta\mathcal{F}_u[v] \quad \textit{whenever } u, u+v \in \mathcal{A}. \tag{5.1}$$

Notice that the first variation will exist under these assumptions even if $u + \epsilon v \notin \mathcal{A}$ *for all* ϵ.

A convex integral functional is said to be **strictly convex** *if equality in (5.1) occurs only for* $v \equiv 0$.

Theorem 13. *If $\mathcal{F}$ is convex on $\mathcal{A}$, as in the definition above, and for every $v \in C^1[a,b]$ such that $u_0 + v \in \mathcal{A}$, we have*

$$\delta \mathcal{F}_{u_0}[v] = 0,$$

then $\mathcal{F}[u_0] \leq \mathcal{F}[u]$ for all $u \in \mathcal{A}$.

If $\mathcal{F}$ is strictly convex, then u_0 is the unique minimizer.

Proof: If $u \in \mathcal{A}$, then $v = u - u_0 \in C^1[a,b]$ and $u + v \in \mathcal{A}$, so

$$\mathcal{F}[u] - \mathcal{F}[u_0] \geq \delta \mathcal{F}_{u_0}[v] = 0.$$

Clearly, if $\mathcal{F}$ is strictly convex, then the inequality is strict for all $u \neq u_0$. □

Theorem 14. *If $F \in C^2([a,b] \times \mathbb{R} \times \mathbb{R})$ is* **strictly second order convex** *in z and p, i.e., for each fixed x, the matrix*

$$\begin{pmatrix} F_{zz} & F_{zp} \\ F_{zp} & F_{pp} \end{pmatrix} \qquad \text{is positive definite,} \tag{5.2}$$

then $\mathcal{F}[u] = \int_a^b F(x,u,u')\,dx$ is strictly convex.

Proof: For fixed x we have a Taylor expansion formula

$$\begin{aligned} F(x, u+v, u'+v') - F(x,u,u') &= F_z(x,u,u')v + F_z(x,u,u')v' \\ &\quad + \frac{1}{2}(v,v') \begin{pmatrix} F_{zz}(x,z^*,p^*) & F_{zp}(x,z^*,p^*) \\ F_{zp(x,z^*,p^*)} & F_{pp}(x,z^*,p^*) \end{pmatrix} \begin{pmatrix} v \\ v' \end{pmatrix} \end{aligned}$$

where (z^*, p^*) is a point on the line segment connecting (u,u') to $(u+v, u'+v')$ in $\mathbb{R}^2$. Since

$$\begin{aligned} &(v,v') \begin{pmatrix} F_{zz}(x,z^*,p^*) & F_{zp}(x,z^*,p^*) \\ F_{zp(x,z^*,p^*)} & F_{pp}(x,z^*,p^*) \end{pmatrix} \begin{pmatrix} v \\ v' \end{pmatrix} \\ &\qquad = \left\langle \begin{pmatrix} v \\ v' \end{pmatrix}, \begin{pmatrix} F_{zz}(x,z^*,p^*) & F_{zp}(x,z^*,p^*) \\ F_{zp}(x,z^*,p^*) & F_{pp}(x,z^*,p^*) \end{pmatrix} \begin{pmatrix} v \\ v' \end{pmatrix} \right\rangle \geq 0 \end{aligned}$$

with equality only if $(v,v') = (0,0)$. we have

$$F(x, u+v, u'+v') - F(x,u,u') \geq F_z(x,u,u')v + F_z(x,u,u')v' \tag{5.3}$$

with equality only if $(v, v') = (0, 0)$. Integrating, we find

$$\mathcal{F}[u+v] - \mathcal{F}[u] \geq \int_a^b [F_z(x,u,u')v + F_z(x,u,u')v']\,dx = \delta\mathcal{F}_u[v]. \qquad (5.4)$$

This means $\mathcal{F}$ is convex on $\mathcal{A}$. Initially, the equality condition of (5.3) gives only that $(v(x), v'(x)) = (0, 0)$ at the point under consideration. If, however, we have equality in (5.4) in the presence of the inequality of (5.3), then by continuity, we must have equality in (5.3) at *every* point. Thus, we get $v \equiv 0$ in the case of equality, and $\mathcal{F}$ is strictly convex. □

5.2 Applications to the Brachistochrone problem

It is pointed out by Troutman that if we consider points $(1, -d)$ with $d > 2/\pi$, then the resulting cycloid curve is the graph of a function $x = v(y)$ with $v \in C^1[0, -d]$, and the transit time may be expressed as

$$\begin{aligned} T[v] &= \int_{-d}^0 \frac{1}{\frac{dy}{dt}}\,dy \\ &= \int_{-d}^0 \frac{1}{\frac{dy}{ds}\frac{ds}{dt}}\,dy \\ &= \int_{-d}^0 \sqrt{\frac{1+v'^2}{-2gy}}\,dy \end{aligned}$$

since

$$s = \int_{-d}^0 \sqrt{1+[v'(\eta)]^2}\,d\eta \qquad \text{implies} \qquad \frac{dy}{ds} = \frac{1}{\sqrt{1+v'^2}}.$$

The Lagrangian

$$F(y,p) = \sqrt{\frac{1+p^2}{-2gy}}$$

while non-autonomous is strictly second order convex in p since

$$\frac{d^2}{dp^2}\sqrt{1+p^2} = \frac{d}{dp}\frac{p}{\sqrt{1+p^2}} = \frac{1}{(1+p^2)^{3/2}} > 0.$$

This implies these particular cycloids are unique minimizers among curves given as graphs $\{(v(y), y) : y \in [-d, 0] : v(-d) = 1,\ v(0) = 0\}$ determined by functions $v \in C^1[-d, 0]$.

Exercise 36. *Technically, Theorem 14 does not apply to the functional considered by Troutman because the Hessian matrix with respect to* v *and* v' *is degenerate in the* v *direction. Prove a version of the theorem which applies to this case.*

It is pointed out in BGH that under certain assumptions, the transit time may be written as

$$T[v] = \frac{1}{g}\int_0^1 \sqrt{\frac{g^2}{v^2} + v'^2}\, dx.$$

To be precise, consider

$$\mathcal{A}_1 = \left\{ u \in C^0[0,1] \cap C^1(0,1] : u \le 0,\ u(0) = 0,\ u(1) = -d,\ \int_0^1 \frac{1}{\sqrt{-u}}\, dx < \infty \right\}.$$

Setting

$$\mathcal{A}_2 = \{ v \in C^0[0,1] \cap C^1(0,1] : v = \sqrt{-2gu} \text{ for some } u \in \mathcal{A}_1 \}$$

we have the integrability condition

$$\int_0^1 \frac{1}{v}\, dx < \infty,$$

$2gu = -v^2$ so $gu' = -2vv'$, and

$$T[v] = \int_0^1 \sqrt{\frac{1 + u'^2}{-2gu}}\, dx = \frac{1}{g}\int_0^1 \sqrt{\frac{g^2 + v^2 v'^2}{v^2}}\, dx$$

as given above. The Lagrangian

$$F(z, p) = \sqrt{\frac{g^2}{z^2} + p^2}$$

is still autonomous and has

$$F_z = -\frac{g^2/z^3}{\sqrt{\frac{g^2}{z^2} + p^2}} \quad \text{and} \quad F_p = \frac{p}{\sqrt{\frac{g^2}{z^2} + p^2}};$$

$$F_{zz} = \frac{3g^2/z^4}{\sqrt{\frac{g^2}{z^2}+p^2}} - \frac{g^4/z^6}{\left(\frac{g^2}{z^2}+p^2\right)^{3/2}} = \frac{2g^4/z^6 + 3g^2p^2/z^4}{\left(\frac{g^2}{z^2}+p^2\right)^{3/2}} > 0;$$

$$F_{zp} = -\frac{g^2p/z^3}{\left(\frac{g^2}{z^2}+p^2\right)^{3/2}} \quad \text{and} \quad F_{pp} = \frac{g^2/z^2}{\left(\frac{g^2}{z^2}+p^2\right)^{3/2}} > 0.$$

Finally,

$$F_{zz}F_{pp} - F_{zp}^2 = \frac{1}{\left(\frac{g^2}{z^2}+p^2\right)^3}\left[\frac{2g^6}{z^8} + 3\frac{g^4p^2}{z^6} - \frac{g^4p^2}{z^6}\right] > 0.$$

This means the Hessian matrix is positive definite and the functional T is strictly convex on $\mathcal{A}_2$.

Exercise 37. *Theorem 13 and 14 (and technically even the definition of convexity) given above do not apply in this case. Modify the construction above to treat the admissible set $\mathcal{A}_2$ in the universal vector space $C^0[0,1] \cap C^1(0,1]$ to show the cycloids provide global minimizers in $\mathcal{A}_2$.*