Convex Functions

Note Title 3/7/2009

X topological vector space

$F: X \to \mathbb{R}$ is convex iff $\forall x, y \in X$

$$F((1-t)x + ty) \leq (1-t)F(x) + t\,F(y) \quad \forall\, 0 \leq t \leq 1$$

Theorem 1 If $F: X \to \mathbb{R}$ is continuous then the following are equivalent:

1. F is convex
2. $\forall x \in X \ \exists h \in X^*$ such that
$$F(y) \geq F(x) + \langle h, y-x \rangle$$

Proof $1. \Rightarrow 2.$

Fix $x \in X$ arbitrary. Define
$$\text{epi}\, F = \{ (y, \alpha) \in X \times \mathbb{R} \mid \alpha \geq F(y) \}$$

$$C = \text{epi}\, F - (x, F(x)) = \{ (z, \beta) \in X \times \mathbb{R} \mid \beta \geq F(z+x) - F(x) \}$$

F convex $\Rightarrow$ epi F convex $\Rightarrow$ C convex $\Rightarrow$ Int C convex
(Int C = interior of C).

F continuous $\Rightarrow$ Int $C = \{ (z, \beta) \in X \times \mathbb{R} \mid \beta > F(z+x) - F(x) \}$

So Int $C \neq \emptyset$ and $(0,0) \notin$ Int C

All in all we have

$$\left.\begin{array}{l}\text{Int } C \text{ open, convex} \\ \text{Int } C_1 \cap (0,0) = \phi\end{array}\right\} \xrightarrow[\text{geometric form}]{\text{Hahn - Banach}} \exists$$

$l^* \in (X \times \mathbb{R})^*$ such that

$$l^*(z, \beta) > 0 \quad \forall (z, \beta) \in C$$

But $l^* \in (X \times \mathbb{R})^* \Rightarrow \exists\ h^* \in X^*$ and $a \in \mathbb{R}$ such that

$$l^*(z, \beta) = \langle h^*, z \rangle + a\beta$$

Hence $\langle h^*, y - x \rangle + a(\alpha - F(x)) > 0$ for all $(y, \alpha) \in \text{Int}(\text{epi } F)$.

Since $(x, F(x)+1) \in \text{Int}(\text{epi } F) \Rightarrow$

$$\langle h^*, x - x \rangle + a(F(x)+1-F(x)) > 0$$

$\Rightarrow a > 0$ hence:

$$\alpha > F(x) + \langle -\tfrac{1}{a} h^*, y - x \rangle$$

Let $h \in X^*$ $h = -\frac{1}{a} h^*$ we get:

$$\alpha > F(x) + \langle h, y - x \rangle \quad \forall\ \alpha > F(y)$$

By letting $\alpha \to F(y)$ $\alpha > F(y)$ we get

$$F(y) \geq F(x) + \langle h, y - x \rangle$$

$2 \Rightarrow 1$

Fix $z_1, z_2 \in X$ and $t \in (0,1)$ we want to show

$$F((1-t)z_1 + t z_2) \leq (1-t)F(z_1) + t F(z_2)$$

Use 2. with $x = (1-t)z_1 + t z_2$ and $y = z_1, y = z_2$:

$$F(z_1) \geq F(x) + \langle h, z_1 - x \rangle = F(x) + \langle h, t(z_1 - z_2) \rangle$$
$$F(z_2) \geq F(x) + \langle h, z_2 - x \rangle = F(x) + \langle h, (1-t)(z_2 - z_1) \rangle$$

So $(1-t)F(z_1) + t F(z_2) \geq (1-t+t)F(x) + (1-t)t \underbrace{\langle h, z_1 - z_2 + z_2 - z_1 \rangle}_{=0}$

Hence $(1-t)F(z_1) + t F(z_2) \geq F((1-t)z_1 + t z_2)$ q.e.d

<u>Remark 1</u> $2 \Rightarrow 1$ does not use continuity of F.

<u>Lemma 2</u> If X is a normed space and F is convex and $\exists x_0 \in X$, $\exists V \ni x_0$ open $\exists M > 0$ such that

$$F(x) \leq M \quad \forall x \in V$$

then F is continuous on X. (The result generalizes to locally convex spaces)

Proof WLOG we can assume $x_0 = 0$
(Use $G(\cdot) = F(\cdot - x_0)$ convex and bdd above)
Then $\exists \varepsilon > 0$ such that

$$B(0, 2\varepsilon) = \{ x \in X : \|x\| \leq 2\varepsilon \} \subset V$$

Want we have that F is actually bounded on $B(0, 2\varepsilon)$ indeed

$$F\left(\tfrac{1}{2}(-x) + \tfrac{1}{2}x\right) = F(0) \leq \tfrac{1}{2}F(-x) + \tfrac{1}{2}F(x)$$

$\Rightarrow F(x) \geq 2F(0) - F(-x) \geq 2F(0) - M \quad \forall x \in B(0, 2\varepsilon)$
Then $\exists K > 0$ such that

$$|F(x_1) - F(x_2)| \leq K \|x_1 - x_2\| \quad \forall x_1, x_2 \in B(0, \varepsilon)$$

Assume not $\Rightarrow \exists x_1, x_2 \in B(0, \varepsilon)$:

$$\frac{F(x_2) - F(x_1)}{\|x_2 - x_1\|} > \frac{2M}{\varepsilon}$$

Construct $x_3 = x_2 + \alpha(x_2 - x_1)$, $\alpha > 0$ such that

$$\|x_3 - x_2\| = \varepsilon \Rightarrow x_3 \in B(0, 2\varepsilon)$$

Since $\phi : \mathbb{R} \to \mathbb{R}$ $\phi(t) = F(x_2 + t(x_2 - x_1))$ is convex
$\Rightarrow \dfrac{\phi(\cdot) - \phi(0)}{\cdot - 0}$ is increasing. Using $t = \alpha$, $t = -1$ we

get

$$\frac{F(x_3)-F(x_2)}{\|x_3-x_2\|} \geq \frac{F(x_1)-F(x_2)}{-\|x_2-x_1\|} > \frac{2M}{\varepsilon}$$

$\Rightarrow$ $F(x_3)-F(x_2) > 2M$ contradiction with
$|F(x)| \leq M \quad \forall x \in B(0,2\varepsilon)$.

So F is Lipschitz on $B(0,\varepsilon) \Rightarrow F$ continuous on $B(0,\varepsilon)$.

It remains to show that F bounded above on $B(0,2\varepsilon)$ implies $\forall y \in X \ \exists V_y \ni y$ open such that $F|_{V_y}$ is bounded above.

Fix $y \in X$ $y \neq 0$ and $\rho > 1$. Let $z = \rho y$, $\lambda = 1/\rho$.
Then

$$V_y = \{(1-\lambda)x + \lambda z \mid x \in B(0,2\varepsilon)\} \ni y$$

includes $B(y,(1-\lambda)2\varepsilon)$ (actually $B(y,(1-\lambda)2\varepsilon) = V_y$)
and

$$F((1-\lambda)x + \lambda z) \leq (1-\lambda)F(x) + \lambda F(z) \leq (1-\lambda)M + \lambda F(z)$$

$\Rightarrow F|_{V_y}$ bounded above.

The lemma is now completely proven

Corollary 3 If X is finite dimensional and F is convex then F is continuous.

Proof Take the simplex S with vertices $0, e_1, e_2, \dots, e_n$ where $e_1, e_2, \dots, e_n$ is a basis in X. It forms a neighborhood of a point in its interior (which is nonempty) and

$$F\big|_S \leq \max\{F(0), F(e_1), \dots, F(e_n)\} < \infty$$

Apply now previous Lemma.

Theorem 4 If X is a topological vector space and $\forall\, x, y \in X$

$$dF(x)[y] = \lim_{t \to 0} \frac{F(x+ty) - F(x)}{t}$$ exists and is finite

$$d^2F(x)[y,y] = \lim_{t \to 0} \frac{dF(x+ty)[y] - dF(x)[y]}{t}$$ exists

and is finite and $d^2F(x)[y,y] \geq 0$ then

F is convex!

Proof Fix $x_1, x_2 \in X$ we want to show

$$F((1-t)x_1 + t x_2) \leq (1-t)F(x_1) + t F(x_2) \quad \forall\, 0 \leq t \leq 1$$

It suffices to show that F restricted to the line passing through x_1, x_2 is convex i.e:

$$\Phi : \mathbb{R} \to \mathbb{R} \quad \Phi(s) = F(x_1 + s(x_2 - x_1)) \quad t \in \mathbb{R}$$

is convex. But, by hypothesis Φ has a nonnegative second derivative everywhere (use $x = x_1 + s(x_2 - x_1)$, s arbitrary and $y = x_2 - x_1$). Hence Φ has a nondecreasing first derivative hence Φ is convex.

The hypothesis in Theorem 4 are not necessary for F to be convex. They require that F restricted to any line $x_1 + t x_2$, $t \in \mathbb{R}$ be twice classically differentiable with nonnegative second order derivative

Remark 2 For $\Phi : \mathbb{R} \to \mathbb{R}$ convex we have

1° $\forall x$: $\Phi_x(\cdot) = \dfrac{\Phi(\cdot) - \Phi(x)}{\cdot - x}$ is nondecreasing.

This is also sufficient to imply convexity

2° The right derivative:

$$D_+ \Phi(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{\Phi(x+\varepsilon) - \Phi(x)}{\varepsilon} \quad \text{exists}$$

everywhere and is nondecreasing. This, combined with Φ continuous, is also sufficient to show convexity.

3° The left derivative:

$$D_- \Phi(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{\Phi(x) - \Phi(x-\varepsilon)}{\varepsilon} \quad \text{exists}$$

everywhere and is nondecreasing. This, combined with Φ continuous, is also sufficient to show convexity

4° $D_- \Phi(x) = D_+ \Phi(x)$ everywhere except maybe a countable number of points. So the classical derivative exist almost everywhere (actually everywhere except a countable number of poi) and is a nondecreasing function. This is <u>not</u> sufficient to show convexity. But a nondecreasing classical derivative <u>everywhere</u> is sufficient to show convexity.

5° $D_+\Phi$ has nonnegative classical derivative almost everywhere. This can be interpreted as the "second derivative of Φ" exists almost everywhere. However this is not sufficient to show convexity. But, of course, existence of nonnegative second order classical derivative everywhere implies convexity.

6° $D_-\Phi$ has nonnegative classical derivative almost everywhere. In fact the derivative of $D_-\Phi$ exists exactly at the same points where the derivative of $D_+\Phi$ exists and the two are equal.

7° Φ is a continuous (hence locally integrable) function which generates a distribution with second order distributional derivative being a locally finite, nonnegative Borel measure. Reciprocal, any locally finite, nonnegative Borel measure is the second order distributional derivative of a convex function

8° $$\limsup_{\varepsilon \to 0} \frac{\Phi(x+\varepsilon) + \Phi(x-\varepsilon) - 2\,\Phi(x)}{\varepsilon^2} \geq 0$$

for all x. This, and Φ continuous, also suffices to show convexity.