Convex Functions

\( X \) topological vector space

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

\[ F((1-t)x + ty) \leq (1-t)F(x) + tF(y) \quad \forall 0 \leq t \leq 1 \]

**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 \( G \) convex (Int \( C = \) interior of \( C \)).

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

So Int \( C \neq \emptyset \) and \( (0, 0) \in \text{Int } C \).
All in all we have

\[ \text{Int } C \text{ open, convex} \]
\[ \text{Int } C \cap \mathbb{Q} \times \{0\} = \emptyset \]
\[ \text{geometrizes to} \]
\[ \exists \rho \in (X \times \mathbb{R})^* \text{ such that} \]
\[ (\rho, (z, \beta)) > 0 \quad \forall (z, \beta) \in C \]

But \( (\rho, (x, \beta)) \in (X \times \mathbb{R})^* \Rightarrow \exists x^* \in X^* \text{ and } a \in \mathbb{R} \)

such that

\[ (\rho, (x, \beta)) = (x^*, (z, \beta)) + a \beta \]

Hence \( (x^*, (y-x)) + a (\mathbb{1} - F(x)) > 0 \) for all \((y, z) \in \text{Int } (\text{epi } F)\).

Since \((x, F(x)+1) \in \text{Int } (\text{epi } F) \Rightarrow \)
\[ (x^*, (x-x)) + a (\mathbb{1} - F(x)) > 0 \]
\[ \Rightarrow a > 0 \text{ hence:} \]
\[ a > F(x) + \langle \frac{-1}{a} x^*, y-x \rangle \]

Let \( h \in X^* \quad h = -\frac{1}{a} x^* \quad \text{we get:} \]
\[ a > F(x) + \langle h, y-x \rangle \quad \forall \quad a \geq F(y) \]

By letting \( x \rightarrow F(y) \)
\[ a \geq F(y) \]

we get
\[ F(y) > F(x) + \langle h, y-x \rangle \]
\[ 2 = 3 \rightarrow 1 \]

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 \( y = z_1, y = z_2 \):

\[ F(z_1) \geq F(x) + \langle u, z_1 - x \rangle = F(x) + \langle u, t (z_1 - z_2) \rangle \]
\[ F(z_2) \geq F(x) + \langle u, z_2 - x \rangle = F(x) + \langle u, (1-t) (z_2 - z_1) \rangle \]

So

\[ (1-t) F(z_1) + t F(z_2) \geq (1-t+t) F(x) + (1-t) t \langle u, z_1 - z_2 \rangle \]

Hence

\[ (1-t) F(z_1) + t F(z_2) \geq F((1-t)z_1 + tz_2) \] \[ \geq 2 \quad \text{a.d.} \]

Remark 1: \( 2 = 3 \) does not use continuity of \( F \).

Lemma 2: If \( X \) is a normed space and \( F \) is convex and \( \exists x_0 \in X, \exists V \exists x_0 \text{ open } \exists M > 0 \text{ such that } -F(x) < M \forall x \in V \)

then \( F \) is continuous on \( X \). (The result generalizes to locally convex spaces.)
Proof: WirLOG we can assume \( x_0 = 0 \) (Use \( G(\cdot) = F(\cdot - x_0) \) convex and hold above.) Then \( \exists \varepsilon > 0 \) such that
\[
\beta (0, 2\varepsilon) = \{ x \in X : \| x \| \leq 2\varepsilon \} \subset V
\]
Want we prove that \( F \) is absolutely bounded on \( \beta (0, 2\varepsilon) \) indeed,
\[
F \left( \frac{1}{2} (-x) + \frac{1}{2} x \right) = F(0) \leq \frac{1}{2} F(-x) + \frac{1}{2} F(x)
\]
\[
\Rightarrow F(x) \geq 2F(0) - F(-x) \geq 2F(0) - M \forall x \in \beta (0, 2\varepsilon)
\]
Then \( \exists \| \cdot \| > 0 \) such that
\[
\| F(x_1) - F(x_2) \| \leq \varepsilon \| x_1 - x_2 \| \forall x_1, x_2 \in \beta (0, 2\varepsilon)
\]
Assume not \( \Rightarrow \exists x_1, x_2 \in \beta (0, 2\varepsilon) : 
\[
\frac{F(x_2) - F(x_1)}{\| x_2 - x_1 \|} < \frac{2M}{\varepsilon}
\]
Construct \( x_3 = x_2 + d(x_2 - x_1) \), \( d > 0 \) such that
\[
\| x_3 - x_2 \| = \varepsilon \Rightarrow x_3 \in \beta (0, 2\varepsilon)
\]
Since \( \Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \Phi(t) = F(x_2 + \frac{t}{d}(x_2 - x_1)) \) is convex
\[
\Rightarrow \frac{\Phi(t) - \Phi(0)}{t} \text{ is increasing. Using } t = x_3, t = 0 \text{ we get}
\]
\[
\frac{F(x_2) - F(x_1)}{||x_2 - x_1||} \geq \frac{F(x_3) - F(x_2)}{||x_3 - x_2||} \geq \frac{2M}{\varepsilon}
\]

\[
\Rightarrow F(x_2) - F(x_1) \geq 2M \quad \text{contradiction with } |F(x)| \leq M \quad \forall x \in \mathcal{B}(0, 2\varepsilon).
\]

So \( F \) is upper bounded on \( \mathcal{B}(0, \varepsilon) \); \( F \) continuous on \( \mathcal{B}(0, \varepsilon) \).

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

Fix \( y \in X \) \( y \neq 0 \) and \( \varepsilon > 1 \). Let \( z = \varepsilon y \), \( \lambda = 1/\varepsilon \).

Then
\[
V_y = \{(1-\lambda)x + \lambda z \mid x \in \mathcal{B}(0, 2\varepsilon)\}, \quad \forall y
\]
includes \( \mathcal{B}(y, (1-\lambda)2\varepsilon) \) (actually \( \mathcal{B}(y, (1-\lambda)2\varepsilon) = V_y \)) and

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

\[
\Rightarrow F|_{V_y} \text{ bounded above.}
\]

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 $0, e_1, e_2, \ldots, e_n$ where $e_1, e_2, \ldots, e_n$ is a basis in $X$. It forms a neighborhood of a point in $\operatorname{int}(S)$ (which is nonempty) and

$$F \mid_S \leq \max \{ F(0), F(e_1), \ldots, F(e_n) \} < \infty$$

Apply your previous lemma.

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

$$d_F(x,y) = \lim_{t \to 0} \frac{F(x+ty) - F(x)}{t}$$

exists and is finite.

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

exists and is finite. If $d^2 F(x)[y,y] > 0$ then

$F$ is convex!
Proof. Fix $x_1, x_2 \in X$ we want to show

$$F((1-t)x_1 + tx_2) \leq (1-t)F(x_1) + tF(x_2) \quad \forall 0 \leq t \leq 1$$

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

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

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

The hypothesis in Theorem 4 are not necessary for $F$ to be convex. They require that $F$ restricted to any line $x_1 + tx_2$, $t \in \mathbb{R}$ be twice continuously differentiable with nonnegative second order derivative.

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

$$\forall x \in X : \Phi_x(\cdot) = \frac{\Phi(\cdot) - \Phi(x)}{\cdot - x}$$

is nondecreasing.

This is also sufficient to imply convexity.
2° The right derivative:

\[ D^+ \Phi(x) = \lim_{\varepsilon \to 0^+} \frac{\Phi(x+\varepsilon) - \Phi(x)}{\varepsilon} \]

everywhere and is nondecreasing. This, combined with \( \Phi \) continuous, is also sufficient to show convexity.

3° The left derivative:

\[ D^- \Phi(x) = \lim_{\varepsilon \to 0^-} \frac{\Phi(x) - \Phi(x-\varepsilon)}{\varepsilon} \]

everywhere and is nondecreasing. This, combined with \( \Phi \) 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 exists almost everywhere (actually everywhere except a countable number of points) and is a nondecreasing function. This is not sufficient to show convexity. But a nondecreasing classical derivative everywhere is sufficient to show convexity.
5° $D_+ \phi$ has nonnegative classical derivative almost everywhere. This can be interpreted as the "second derivative of $\phi" exist 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° $\phi$ is a continuous (hence locally integrable) function which generates a distribution with second order distributional derivative, being a locally finite, nonnegative Borel measure. Reciprocally, any locally finite, nonnegative Borel measure is the second order distributional derivative of a convex function.
for all \( x \). Thus, and \( \phi \) continuous, also suffices to show convexity.