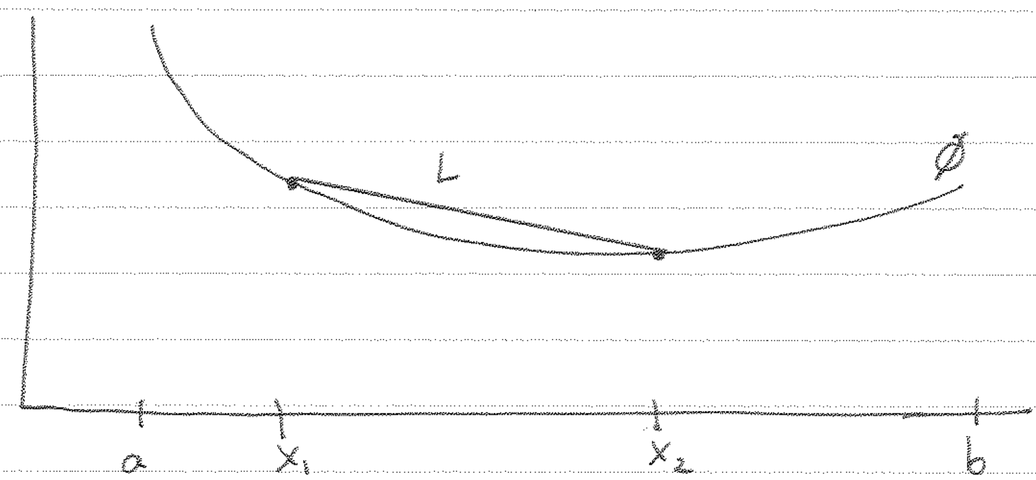
Convex Functions & Jensen's Inequality

Definition

A function $\phi : (a,b) \longrightarrow \mathbb{R}$ is <u>convex</u> if for each $a < x_1 < x_2 < b$, the graph of ϕ on $[x_1, x_2]$ lies on or below the line segment connecting the points $(x_1, \phi(x_1))$ & $(x_2, \phi(x_2))$.

Note: We allow the possibilities $a = -\infty$ or $b = \infty$.



We can parameterize the interval $[x_1, x_2]$ by

$$X(t) = t x_1 + (1-t) x_2 , \quad t \in [0,1].$$

The equation for the line L is the parameterized as

Theorem

If $\phi : (a,b) \to \mathbb{R}$ is differentiable everywhere and ϕ' is monotone increasing, then ϕ is convex.

Proof:

Fix any $a < x_1 < x < x_2 < b$. By the Mean-Value Theorem, $\exists\, \xi_1 \in (x_1, x)$ & $\xi_2 \in (x, x_2)$ such that

$$\frac{\phi(x) - \phi(x_1)}{x - x_1} = \phi'(\xi_1) \qquad \frac{\phi(x_2) - \phi(x)}{x_2 - x} = \phi'(\xi_2).$$

Since ϕ' is monotone increasing, we therefore have

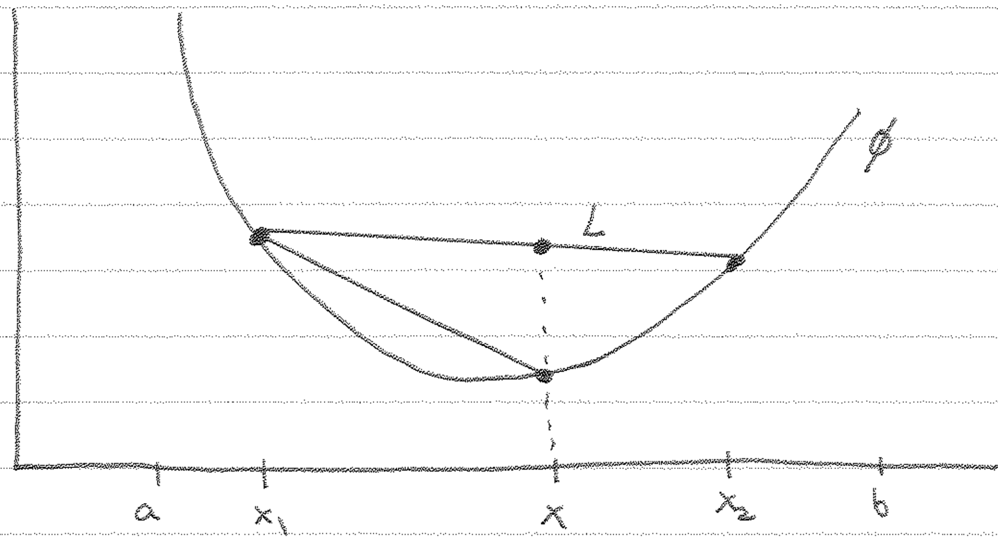
$$\frac{\phi(x) - \phi(x_1)}{x - x_1} = \phi'(\xi_1) = \min\left\{\frac{\phi(x) - \phi(x_1)}{x - x_1}, \frac{\phi(x_2) - \phi(x)}{x_2 - x}\right\}$$

$$\leq \frac{\big(\phi(x) - \phi(x_1)\big) + \big(\phi(x_2) - \phi(x)\big)}{(x - x_1) + (x_2 - x)}$$

$$= \frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1}$$

Hence ϕ is convex by an earlier exercise. $\blacksquare$

Exercise



Show that ϕ is convex if & only if

$$\frac{\phi(x_2) - \phi(x_1)}{x_2 - x_1} \geq \frac{\phi(x) - \phi(x_1)}{x - x_1}$$

for all $a < x_1 < x < x_2 < b$

(Consider the slopes of the two lines in the illustration).

Exercise

If $b_1, b_2 > 0$, then

$$\min\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\} \leq \frac{a_1 + a_2}{b_1 + b_2} \leq \max\left\{\frac{a_1}{b_1}, \frac{a_2}{b_2}\right\}$$

$$L(x(t)) = t\,\phi(x_1) + (1-t)\,\phi(x_2), \quad t \in [0,1].$$

Hence ϕ is convex if & only if

$$\phi(t x_1 + (1-t) x_2) \le t\,\phi(x_1) + (1-t)\,\phi(x_2)$$

for all $a < x_1 < x_2 < b$, $0 \le t \le 1$.

Exercise: Discrete Form of Jensen's Inequality

Assume ϕ is convex in (a,b). Show that if

$x_1, \dots, x_N \in (a,b)$ & $\sum_{j=1}^{N} t_j = 1$ with $t_j > 0$, then

$$\phi\left(\sum_{j=1}^{N} t_j x_j\right) \le \sum_{j=1}^{N} t_j\,\phi(x_j)$$

Exercise

a. ϕ_1, ϕ_2 convex in (a,b) $\Longrightarrow$ $\phi_1 + \phi_2$ is convex.

b. ϕ convex in (a,b) & $c > 0$ $\Longrightarrow$ $c\phi$ is convex.

c. ϕ_k convex in (a,b) & $\phi_k \to \phi$ pointwise in (a,b)

$\Longrightarrow$ ϕ is convex

Examples: x^p is convex on $(0,\infty)$ if $p \geq 1$

e^{ax} is convex on $\mathbb{R}$ for any $a \in \mathbb{R}$

$\log \frac{1}{x} = -\log x$ is convex on $(0,\infty)$

Theorem

If ϕ is convex on (a,b), then:

a. ϕ is continuous,

b. ϕ is differentiable at all but at most Countably many points in (a,b), and

c. ϕ' is monotone increasing (on the set where it is defined).

d. ϕ is Lipschitz on every closed subinterval $[x_1,x_2] \subseteq (a,b)$.

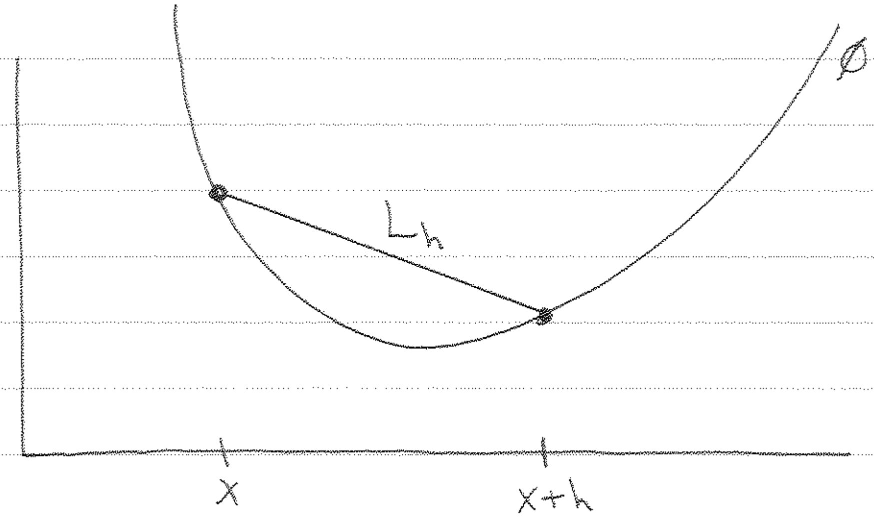
e. $\phi \in AC[x_1,x_2]$ for all $a < x_1 < x_2 < b$.

Proof:

a. Fix any $x \in (a,b)$. Let L_h be the line segment joining $(x, \phi(x))$ to $(x+h, \phi(x+h))$, for all h with $a < x+h < b$. The slope of this line is

$$\frac{\phi(x+h) - \phi(x)}{h} = \text{slope}(L_h).$$

Because ϕ is convex, for $h>0$ we have that slope(L_h) decreases as h decreases.



Since slope(L_h) is decreasing & finite, it has a limit in the range $[-\infty, \infty)$. Thus, the derivative of ϕ from the right exists in an extended real sense:

$$D^+\phi(x) = \lim_{h \to 0^+} \frac{\phi(x+h) - \phi(x)}{h} \in [-\infty, 0)$$

Similarly, the derivative from the left exists, and

$$D^-\phi(x) = \lim_{h \to 0^-} \frac{\phi(x+h) - \phi(x)}{h} \in (-\infty, \infty]$$

Hence for each x we have

$$-\infty < D^-\phi(x) \leq D^+\phi(x) < \infty.$$

Since ϕ has both right & left derivatives at x, it must be continuous at x.

b. Choose any $x \in (a,b)$. Then, as above, $D^+\phi(x)$ & $D^-\phi(x)$ both exist. Since $\text{slope}(L_{-h}) \leq \text{slope}(L_h)$ for $h>0$, we have

$$D^-\phi(x) \leq D^+\phi(x), \quad \text{all } x \in (a,b)$$

Now choose any $a<x<y<b$. Then $y = x+h$ with $h>0$, so again by the decreasing slope remarks from the proof of part a, we have

$$D^+\phi(x) \leq \frac{\phi(x+h)-\phi(x)}{h} = \frac{\phi(y)-\phi(x)}{y-x}.$$

Similarly, since $x = y-h$, we also have

$$\frac{\phi(y)-\phi(x)}{y-x} \leq D^-\phi(y) \leq D^+\phi(y).$$

Combining the above two equations,

$$D^+\phi(x) \leq D^-\phi(y) \leq D^+\phi(y) \quad (*)$$

so $D^+\phi$ is monotone increasing. Similarly, $D^-\phi$

is also monotone increasing. Therefore each of these functions can have at most countably many discontinuities. If y is not one of those points then $D^+\phi$ is continuous at y, so

$$D^+\phi(y) = \lim_{x \to y^-} D^+\phi(x) \leq D^-\phi(y) \quad \text{by } (*)$$

Since we established before that $D^-\phi(y) \leq D^+\phi(y)$, we conclude that $D^-\phi(y) = D^+\phi(y)$, & hence ϕ is differentiable at y.

c. Since $\phi'(x) = D^+\phi(x)$ at any point where ϕ is differentiable, ϕ' is monotone increasing because $D^+\phi$ is.

d. Fix $[x_1, x_2] \subseteq (a,b)$, and choose $x_1 \leq x < y \leq x_2$. Then, as shown before,

$$D^+\phi(x) \le \frac{\phi(y)-\phi(x)}{y-x} \le D^-\phi(y).$$

And since both $D^+\phi$ & $D^-\phi$ are increasing, this implies

$$D^+\phi(x_1) \le \frac{\phi(y)-\phi(x)}{y-x} \le D^-\phi(x_2).$$

Hence, if we set

$$C = \max\{|D^+\phi(x_1)|, \ |D^-\phi(x_2)|\}$$

then we have $\left|\frac{\phi(y)-\phi(x)}{y-x}\right| \le C$, so

$$|\phi(y)-\phi(x)| \le C|y-x|, \quad \text{all } x_1 \le x < y \le x_2.$$

Thus ϕ is Lipschitz on $[x_1, x_2]$.

e. This follows from the fact that Lipschitz functions on $[x_1, x_2]$ are absolutely continuous. ▨

Actually, this result can be improved.

Theorem
Given $\phi : (a,b) \to \mathbb{R}$, TFAE.

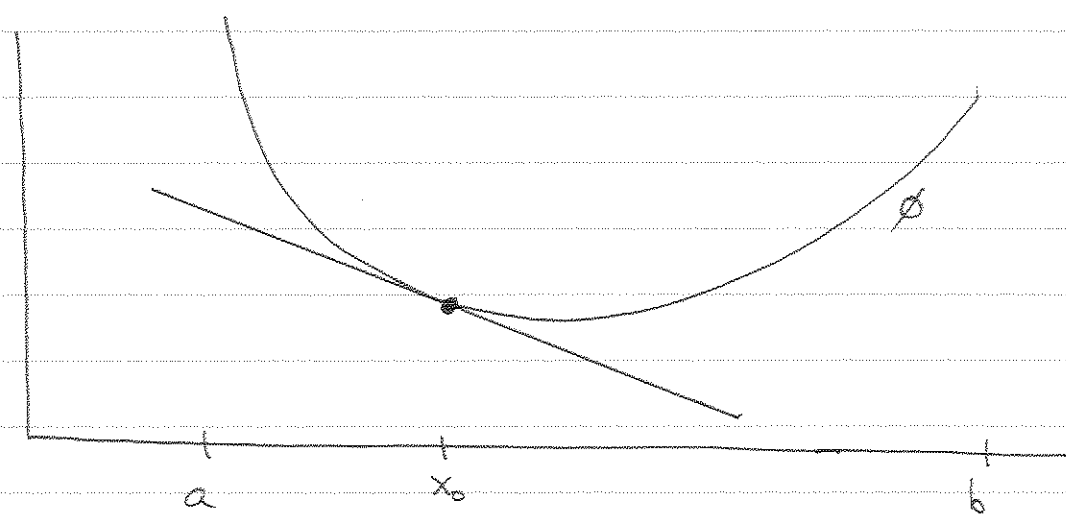
a. ϕ is convex on (a,b).

b. $\phi \in AC[x_1, x_2]$ $\forall\, a < x_1 < x_2 < b$, and
ϕ' is increasing (on the set where it is defined).

Next we will prove an integral version of Jensen's inequality.

Definition

If $\phi : (a,b) \to \mathbb{R}$ is convex & $x_0 \in (a,b)$, then a supporting line at x_0 is any line through $(x_0, \phi(x_0))$ which lies on or below the graph of ϕ.



Exercise

Show that any line through $(x_0, \phi(x_0))$ whose slope m satisfies

$$D^- \phi(x_0) \leq m \leq D^+ \phi(x_0)$$

is a supporting line at x_0.

Theorem: Jensen's Inequality

Let $(X, \mathcal{M}, \mu)$ be any finite measure space with $\mu(X) = 1$ (note μ is a positive measure). If $\phi : (a,b) \to \mathbb{R}$ is convex, then

$$\forall \text{ integrable } g : X \to (a,b), \quad \phi\left(\int g \, d\mu\right) \le \int \phi \circ g \, d\mu.$$

Proof:

By hypothesis, $a < g(x) < b$ for all $x \in X$. Since $\mu(X) = 1$, we therefore have

$$a = \int_X a \, d\mu \le \int_X g \, d\mu \le \int_X b \, d\mu = b.$$

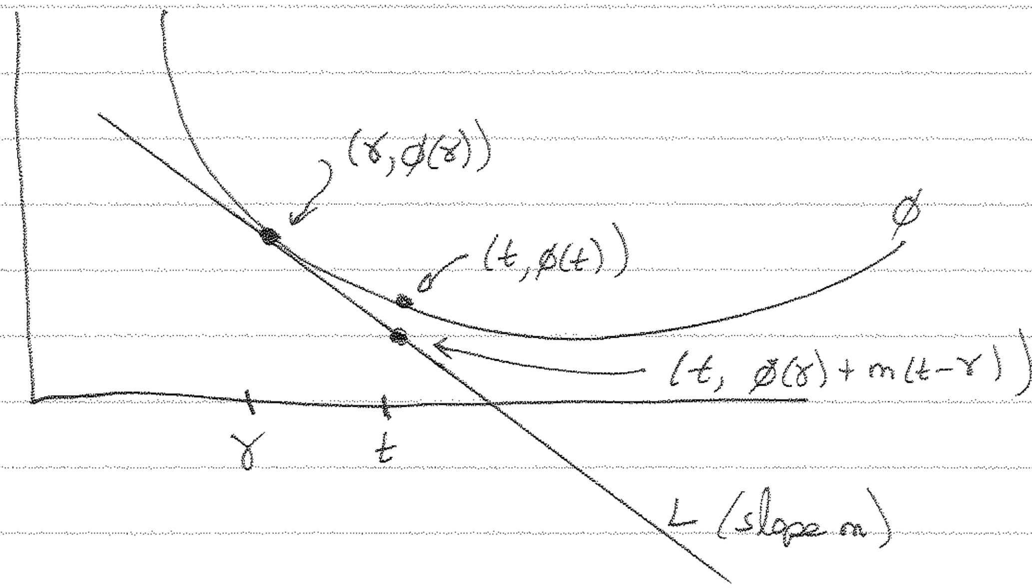
In fact, we must have strict inequalities above since, for example, $\int_X (g - a) \, d\mu = 0$ implies $g(x) = a$ a.e. Hence

$$\gamma = \int g \, d\mu \in (a,b).$$

Let L be any supporting line at γ, & let m be its slope. Then for any $t \in (a,b)$, the point

$\phi(\gamma) + m(t-\gamma)$ lies on L, so we have

$$\phi(\gamma) + m(t-\gamma) \le \phi(t), \qquad t \in (a,b).$$



In particular, if $x \in X$ then $t = g(x) \in (a,b)$, so

$$\phi(\gamma) + m(g(x) - \gamma) \le \phi(g(x)), \qquad x \in X.$$

Integrating over X therefore yields:

$$\begin{aligned} \int \phi(g(x))\, d\mu(x) &\ge \int (\phi(\gamma) + m g(x) - m\gamma)\, d\mu(x) \\ &= \phi(\gamma) + \cancel{m \int g\, d\mu} - \cancel{m\gamma} \\ &= \phi(\gamma) \\ &= \phi\left(\int g\, d\mu\right). \end{aligned}$$

Exercise

Use Jensen's inequality to prove that if $1 < p, q < \infty$ and

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then

$$\forall a, b \geq 0, \qquad ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Remark: This is the inequality that is needed to prove Hölder's Theorem.

Hint: Convexity of e^x (write $a = e^{x/p}$).

Generalize to

$$a_1 \cdots a_N \leq \sum_{j=1}^{N} \frac{a_j^{p_j}}{p_j} \qquad \text{if} \quad \sum_{j=1}^{N} \frac{1}{p_j} = 1$$

Exercise: Show ϕ is convex on (a,b) if & only if ϕ is continuous and

$$\phi\left(\frac{x_1+x_2}{2}\right) \leq \frac{\phi(x_1) + \phi(x_2)}{2} \qquad \forall a < x_1 < x_2 < b.$$

<u>Exercise</u> Use the preceding exercise to show that if f is monotone increasing and ϕ is its antiderivative, i.e.,

$$\phi(x) = \int_a^x f(t)\,dt + \phi(a),$$

then ϕ is convex.