Convex Functions

Andrew D Smith
University College Dublin

25 January 2020

1 Definitions

1.1 Dyadic Convexity

IF we know the value of a function at x and y, what can we say about the value at the midpoint, $(x+y)/2$? We are interested in cases where the value at the midpoint is systematically lower than the values at the ends:

Suppose we have a function $f : \mathbb{R} \mapsto \mathbb{R}$ such that, for all $x, y, \in \mathbb{R}$

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

Examples of such functions include:

- $f(x)$ is linear (equality holds)
- $f(x) = x^2$
- $f(x) = 10^x$
- $f(x) = |x|$ (equality holds if x and y have the same sign)

1.2 Convex Function Definition

A function $f : \mathbb{R} \mapsto \mathbb{R} \cup \{\infty\}$ is *convex* if for all $x, y \in \mathbb{R}$ and $0 \leq \lambda \leq 1$:

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \tag{1}$$

The left hand side is the function evaluated at a point between x and y. The right hand side is the linear interpolation between $f(x)$ and $f(y)$.

A function $f(x)$ is *concave* if $-f(x)$ is convex. Linear functions (and only linear functions) are both concave and convex.

1.3 Adding the Point at Infinity

Sometimes we want to consider a convex function only on a particular range. For example, we might consider $f(x) = 1/x$ on $x > 0$ or $f(x) = -\sqrt{x}$ on $x \geq 0$. These are both convex functions, but over smaller ranges.

In these cases, we define $f(x) = +\infty$ for values of x where $f(x)$ would not otherwise be defined.

1.4 Examples of Convex Functions

1.4.1 Powers

$y = x^n$ on $x > 0$ is:

- Convex if $n \geq 1$ or if $n \leq 0$
- Concave if $0 \leq n \leq 1$.

1.4.2 Exponents

The function $f(x) = 10^x$ is convex.

1.4.3 Logarithms

For $y > 0$, the *logarithm* of y, written $\log_{10}(y)$ is the value of x such that $10^x = y$.

The number of decimal digits in y is $1 + \lfloor \log_{10}(y)$.

The logarithm is not a convex function, but this function is:

$$f(x) = \begin{cases} \infty & x \leq 0 \\ -\log_{10}(x) & x > 0 \end{cases}$$

1.4.4 Piecewise Linear Functions

$$f(x) = \begin{cases} \infty & x < 0 \\ 0 & 0 \leq x \leq 1 \\ x - 1 & x \geq 1 \end{cases}$$

2 Combining Convex Functions

2.1 Sums of Convex Functions are Convex

2.2 Maximum of Convex Functions are Convex

2.3 Minimum of Convex Functions

True or False: the minimum of two convex functions is convex.

False: Consider $(x + 1)^2$ and $(x - 1)^2$.

2.4 Increasing function of a Convex Function

$f(x)$ convex, $g(y)$ increasing.

True or False: $g(f(x))$ is convex?

False: Try $\sqrt{|x|}$ for example.

2.5 Convex Function of a Convex Function

$f(x)$ convex, $g(y)$ convex.

True or False: $g(f(x))$ is convex?

False. Try $\frac{1}{1+|x|}$.

2.6 Increasing Convex Function of a Convex Function

Suppose $f(x)$ is convex and $g(y)$ is both increasing and convex.

Then $f(g(x))$ is convex.

Proof As f is convex,

$$f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x)+\lambda f(y)$$

Now as g is an increasing function, we can apply g to eaxch side of the inequality and the result still holds:

$$g[f((1-\lambda)x+\lambda y)] \leq g[(1-\lambda)f(x)+\lambda f(y)]$$

But now we use the convexity of $g(x)$ to get an upper bound for the right hand side:

$$g[(1-\lambda)f(x)+\lambda f(y)] \leq (1-\lambda)g(f(x))+\lambda g(f(y))$$

Putting the last two inequalities together,

$$g[f((1-\lambda)x+\lambda y)] \leq (1-\lambda)g(f(x))+\lambda g(f(y))$$

This is exactly the definition of the convexity of $g(f(x))$.

Corollary: If $f(x)$ is convex, then so is $10^{f(x)}$.

However, if $f(x)$ is convex it does NOT follow that $\log_{10}(f(x))$ is convex. It might be, or it might not.

3 Jensen's Inequality

3.1 Extending from Pairs to Multiples

Consider the situation where:

- $f : \mathbb{R} \mapsto \mathbb{R}$ is a convex function
- $x_1, x_2, \ldots x_n \in \mathbb{R}$
- $\lambda_1, \lambda_2, \ldots \lambda_n$ satisfy

$$\lambda_j \geq 0$$
$$\sum_{j=1}^{n} \lambda_j = 1$$

Then Jensen's inequality states that

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \leq \sum_{j=1}^{n} \lambda_j f(x_j)$$

3.2 Proof of Jensen's Inequality

Proof by induction on n.

Trivial (equality holds) if $n = 1$.

Suppose we now have the situation as described for some $n \geq 2$ and we have proved the relation for $n-1$ points.

For $j = 1, 2, \ldots n-1$, define:

$$\mu_j = \frac{\lambda_j}{1-\lambda_n}$$

These satisfy the same conditions as the λ's, but there are only $n-1$ of them rather than n.

Of course we have:

$$\sum_{j=1}^{n} \lambda_j x_j = (1-\lambda_n)\left[\sum_{j=1}^{n-1} \mu_j x_j\right] + \lambda_n x_n$$

So applying the convexity of f to the mixture on the right hand side, we have:

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \leq (1-\lambda_n) f\left(\sum_{j=1}^{n-1} \mu_j x_j\right) + \lambda_n f(x_n)$$

Now we can can apply the inductive hypothesis:

$$f\left(\sum_{j=1}^{n-1} \mu_j x_j\right) \leq \sum_{j=1}^{n-1} \mu_j f(x_j)$$

Substituting this back gives:

$$f\left(\sum_{j=1}^{n} \lambda_j x_j\right) \leq (1-\lambda_n)\left[\sum_{j=1}^{n-1} \mu_j f(x_j)\right] + \lambda_n f(x_n)$$
$$= \sum_{j=1}^{n} \lambda_j x_j$$

Thus then proves the result claimed.

4 Convex Functions and Optimisation

4.1 Unique Minimum

A convex function

True or False: If a convex function has a minimum, then that minimum is unique.

False: the minimum might be in a range of constant values.

5 Harmonic - Arithmetic - Geometric Means

5.1 Definitions

Positive numbers $x_1, x_2, \dots x_n$.

$$\text{arithmetic mean} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

$$\text{geometric mean} = \sqrt[n]{\prod_{j=1}^{n} x_j}$$

$$\text{harmonic mean} = \left[\frac{2}{n}\sum_{j=1}^{n}\frac{1}{x_j}\right]^{-1}$$

5.2 Numerical Example

With observations $18, 20, 75$ we have

$$\begin{aligned}
\frac{18+20+75}{3} &= \frac{113}{3} & &= 37\frac{2}{3} \\
\sqrt[3]{18 \times 29 \times 75} &= \sqrt[3]{27000} & &= 30 \\
\frac{3}{\frac{1}{18}+\frac{1}{20}+\frac{1}{75}} &= \frac{2700}{107} & &= 25\frac{25}{107}
\end{aligned}$$

5.3 Ordering

We observe that $HM \leq GM \leq AM$ Trial and error suggests this often holds. Equality applies if all the observations are equal.

Can we prove this?

5.4 Jensen Implies $GM \leq AM$

Proof: given n positive observations $x_1, x_2, \dots n$, let

$$y_j = \log_{10} x_j$$

Then applying Jensen's inequality to the convex function 10^y we have:

$$10^{\left(\frac{1}{n}\sum_{j=1}^{n} y_j\right)} \leq \frac{1}{n}\sum_{j=1}^{n} 10^{y_j}$$

This is equivalent to:

$$\sqrt[n]{\prod_{j=1}^{n} x_j} \leq \frac{1}{n}\sum_{j=1}^{n} x_j$$

This is the famous arithmetic-geometric mean inequality.

5.5 General Weighted Results

We can use the same methodology to show the $GM \leq AM$ inequality when the weights are λ_j (with $\sum \lambda_j = 1$) rather than necessarily being equal to $1/n$. IN that case, the $GM \leq AM$ inequality takes the form:

$$\prod_{j=1}^{n} x_j^{\lambda_j} \leq \sum_{j=1}^{n} \lambda_j x_j$$

5.6 $GM \leq AM$ implies $GM \geq HM$

To prove this, apply $GM \leq AM$ to the reciprocals of the observations.

6 Convex Conjugates

6.1 Definition

Suppose we have an arbitrary function $f : \mathbb{R} \mapsto \mathbb{R}$. The *convex conjugate* function, written $f^*(y)$ is defined as:

$$f^*(y) = \sup_{x \in \mathbb{R}} \{xy - f(x)\}$$

This immediately implies *Fenchel's Inequality*

$$xy \leq f(x) + f^*(y)$$

Remark: The Convex Conjugate $f^*(y)$ is always a convex function, regardless of whether $f(x)$ is convex. This follows because $f^*(y)$ is the maximum of a set of convex (indeed, linear) functions of y.

6.2 Examples

6.2.1 Quadratic

Suppose $f(x) = x^2$.

To find $f^*(y)$ we need to maximise (over x)

$$xy - x^2 = \frac{y^2}{4} - \left(\frac{y}{2} - x\right)^2$$

Obviously maximised when $y = 2x$ and the maximum is

$$f^*(y) = \frac{y^2}{4}$$

6.2.2 Square Root

We know that $f(x) = -\sqrt{x}$ is convex on $x \geq 0$.

The convex conjugate is:

$$f^*(y) = \sup_{x \geq 0} \{xy + \sqrt{x}\}$$

This is obviously unbounded if $y \geq 0$. So we look at the less obvious case $y < 0$. Here, we complete the square:

$$xy + \sqrt{x} = -\left((-y)x - \sqrt{x} + \frac{1}{-4y}\right) + \frac{1}{-4y}$$
$$= -\left(\sqrt{x}\sqrt{-y} - \frac{1}{2\sqrt{-y}}\right)^2 + \frac{1}{-4y}$$

Once again, we use the fact that a square is at least zero to complete the maximisation, which finally gives:

$$f^*(y) = \begin{cases} -\frac{1}{4y} & y < 0 \\ +\infty & y \geq 0 \end{cases}$$

6.2.3 Reciprocal

The function $f(x) = 1/x$ is convex on $x > 0$. Then:

$$f^*(y) = \sup_{x>0} \left\{ xy - \frac{1}{x} \right\}$$

It is clear that for $y \geq 0$ we have $f^*(y) = \infty$, while if $f^*(0) = 0$.

So we focus on the non-trivial case when $y < 0$. Once again, we look for a square; this time we note:

$$-(-y)x - \frac{1}{x} = -\left(\sqrt{-y}\sqrt{x} - \frac{1}{\sqrt{x}}\right)^2 - 2\sqrt{-y}$$

Putting the pieces together, we have:

$$f^*(y) = \begin{cases} -2\sqrt{-y} & y \leq 0 \\ \infty & y > 0 \end{cases}$$

6.3 Guessing Duality

The seems to be some kind of duality here: the convex conjugate of a square root on the positive axis is a reciprocal on the negative axis, while the convex conjugate of a reciprocal on the positive axis is a square root on the negative axis.

We also saw that the convex conjugate of a quadratic function is another quadratic function. Indeed, if $f(x) = \frac{x^2}{2}$ then $f^*(y) = y^2/2$ so f is self-conjugate.

True or False: The only self-conjugate function is $f(x) = x^2/2$.

Hint for solution: try Fenchel's inequality with $x = y$.

6.4 Piecewise Linear Example

Let us define $f(x)$ by:

$$f(x) = \begin{cases} \infty & x < 0 \\ 0 & 0 \leq x \leq 1 \\ x - 1 & x \geq 1 \end{cases}$$

To find the convex conjugate $f^*(y)$ we consider separately three cases in the maximisation of $xy - f(x)$.

- If $y < 0$ then the objective function (the thing you're trying to optimise) is decreasing for $x \geq 0$ so the maximum is at $x = 0$, whence $f^*(y) = 0$.
- If $0 < y < 1$ then the objective function in increasing on $0 \leq x \leq 1$ and decreasing on $x \geq 1$ so the maximum is attained at $x = 1$ and so $f^*(y) = y$
- If $y > 1$ then the objective function is increasing for all y and so $f^*(y) = \infty$.
- If $y = 0$ or $y = 1$ then there is a range of x for which the objective function is flat.

Putting these together, we have:

$$f^*(y) = \begin{cases} 0 & y \leq 0 \\ y & 0 \leq y \leq 1 \\ \infty & y > 1 \end{cases}$$

7 Convex Bi-Conjugate

7.1 Definition

Define $f^{**}(x)$ to be the convex conjugate of the convex conjugate of f.

7.2 Bi-Conjugate Does not Exceed f

Fenchel's inequality implies:

$$xy - f^*(y) \leq f(x)$$

Taking the supremum of the left hand side over values of y:

$$f^{**}(x) \leq f(x)$$

7.3 The Case of a Tangent

Suppose $f(x)$ has a supporting line with gradient y_0 at a single point x_0, so that:

$$f(x) \geq f(x_0) + y_0(x - x_0); x \in \mathbb{R}$$

This works for continuous convex functions - take this on trust for now as the proof is tricky. Obviously if such a tangent exists then equality holds at $x = x_0$.

Then we have:

$$\begin{aligned} f^*(y_0) &= \sup_{x \in \mathbb{R}} \{xy_0 - f(x)\} \\ &= x_0 y_0 - f(x_0) + \sup_{x \in \mathbb{R}} \{f(x_0) + y_0(x - x_0) - f(x)\} \\ &= x_0 y_0 - f(x_0) \end{aligned}$$

Taking a specific value $y = y_0$ in the sup defining $f^{**}(y)$, we have:

$$\begin{aligned} f^{**}(x_0) &= \sup_{y \in \mathbb{R}} \{x_0 y - f^*(y)\} \\ &\geq x_0 y_0 - f^*(y_0) \\ &= f(x_0) \end{aligned}$$

As we already know $f^{**}(x) \leq f(x)$, we deduce that in the case where a tangent exists at x_0 we have equality and $f^{**}(x_0) = f(x_0)$.

8 Non-Integer Factorials

8.1 Definition

The factorial $n!$ for integers $n \geq 1$ is defined by the product:

$$n! = 1 \times 2 \times 3 \times \ldots \times (n-1) \times n$$

In more formal terms we define $n!$ inductively:

$$n! = \begin{cases} 1 & n = 0 \\ n \times (n-1)! & n \geq 1 \end{cases}$$

8.2 Log Factorials are Convex

I claim the log factorial is convex, at least when measured on non-negative integers.

Proof: Let $0 \leq a < b < c$ be non-negative integers. Then:

$$\begin{aligned} \frac{b!}{a!} &= (a+1)(a+2)\cdots(b-1)b \\ &\leq b^{b-a} \\ \frac{c!}{b!} &= (b+1)(b+2)\cdots(c-1)c \\ &> b^{c-b} \end{aligned}$$

Then

$$\left(\frac{b!}{a!}\right)^{\frac{c-b}{c-a}} \leq b^{\frac{(b-a)(c-b)}{c-a}} < \left(\frac{c!}{b!}\right)^{\frac{b-a}{c-a}}$$

Write

$$\begin{aligned} \lambda &= \frac{b-a}{c-a} \\ 1-\lambda &= \frac{c-b}{c-a} \end{aligned}$$

Then the inequality we proved is:

$$\left(\frac{b!}{a!}\right)^{1-\lambda} < \left(\frac{c!}{b!}\right)^{\lambda}$$

Cross-multiplying, and taking logs, this is equivalent to

$$\begin{gathered} b = (1-\lambda)a + \lambda c \\ \log_{10} b! \leq (1-\lambda)\log_{10} a! + \lambda \log_{10} c! \end{gathered}$$

We have shown that $log_{10}(n!)$ is convex.

8.3 Extending Factorials to Non-Integers

The factorial function is defined for non-negative integers n. Can we extend it in a logical way for non-integers? We want to preserve:

- The recurrence relation $n! = n \times (n-1)!$
- The log convex property.

Let us try, for example, to devise a logical value of $(-\frac{1}{2})!$.
The recurrence relation implies:

$$\left(\frac{1}{2}\right)! = \frac{1}{2}\left(-\frac{1}{2}\right)!$$

Expressing 0 as the average of $-1/2$ and $1/2$, the log convexity of factorials would imply:

$$1 = (0!)^2 \leq \left(-\frac{1}{2}\right)! \times \left(\frac{1}{2}\right)! = \frac{1}{2}\left(-\frac{1}{2}\right)!^2$$

But then also as $1/2$ is the average of 0 and 1, logarithmic convexity gives:

$$\frac{1}{4}\left(-\frac{1}{2}\right)!^2 = \left(\frac{1}{2}\right)!^2 \leq 0! \times 1! = 1$$

Combining these two inequalities, we have:

$$2 \leq \left(-\frac{1}{2}\right)!^2 \leq 4$$

We can strengthen the inequality by applying the same techniques for larger n.

Expressing n as the average of $n-\frac{1}{2}$ and $n+\frac{1}{2}$, logarithmic convexity implies:

$$(n!)^2 \leq \left(n-\frac{1}{2}\right)!\left(n+\frac{1}{2}\right)!$$

Substituting the recurrence relation for factorials, we have:

$$(n!)^2 \leq \left(-\frac{1}{2}\right)!^2 \prod_{j=1}^{n}\left(j-\frac{1}{2}\right)\prod_{j=1}^{n+1}\left(j-\frac{1}{2}\right)$$

Likewise, expressing $n+\frac{1}{2}$ as the average of n and $n+1$, logarithmic convexity implies:

$$\left(n+\frac{1}{2}\right)^2 \leq (n)!(n+1)!$$

We can expand the left side by the recurrence relation, to give:

$$\left(-\frac{1}{2}\right)!^2 \prod_{j=0}^{n}\left(j+\frac{1}{2}\right)^2 \leq (n)!(n+1)!$$

Combining those equations, we have upper and lower bounds:

$$\frac{(n!)^2}{\prod_{j=1}^{n}\left(j-\frac{1}{2}\right)\prod_{j=1}^{n+1}\left(j-\frac{1}{2}\right)} \leq \left(-\frac{1}{2}\right)!^2 \leq \frac{n!(n+1)!}{\prod_{j=0}^{n}\left(j+\frac{1}{2}\right)^2}$$

Let us evaluate both sizes of this inequality:

n	Lower	Upper
0	2.0000	4.0000
1	2.6667	3.5556
2	2.8444	3.4133
3	2.9257	3.3437
4	2.9722	3.3024
5	3.0022	3.2751
10	3.0677	3.2138
20	3.1035	3.1792
50	3.1261	3.1570
100	3.1338	3.1494
200	3.1377	3.1455
500	3.1400	3.1432
1000	3.1408	3.1424

We guess the limit is π (this is hard to prove, but true). That implies:

$$\left(-\frac{1}{2}\right)! = \sqrt{\pi}$$

We haven't proved that $n!$ admits a logarithmically convex extension to reals (at least those exceeding -1), although that also turns out to be true.