Convex Sets
and Jensen's Inequality

ANDREW D SMITH

School of Mathematics and Statistics

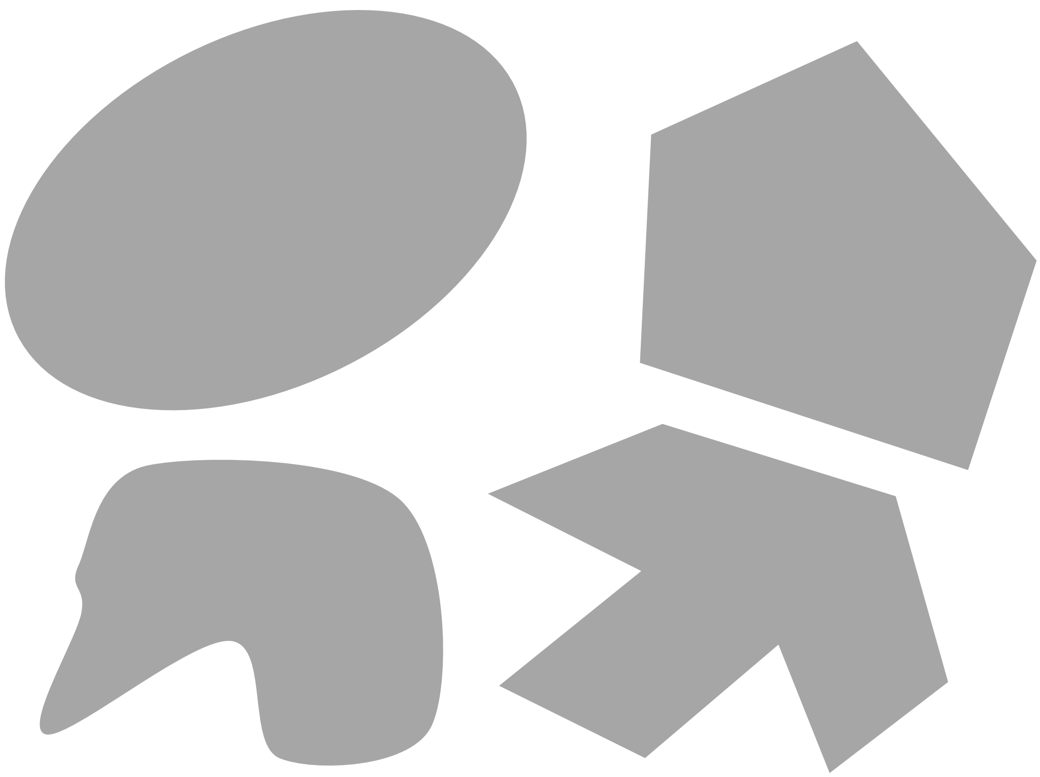
University College Dublin

Definition of Convex Sets: A set $A \subset \mathbb{R}^n$ is *convex* if:

- For any vectors $a, b \in A$
- For any $\lambda \in [0, 1]$
- The point $\lambda a + (1 - \lambda) b \in A$.

This says that if two points, a and b lie in the set, then so does the straight line segment connecting a to b.

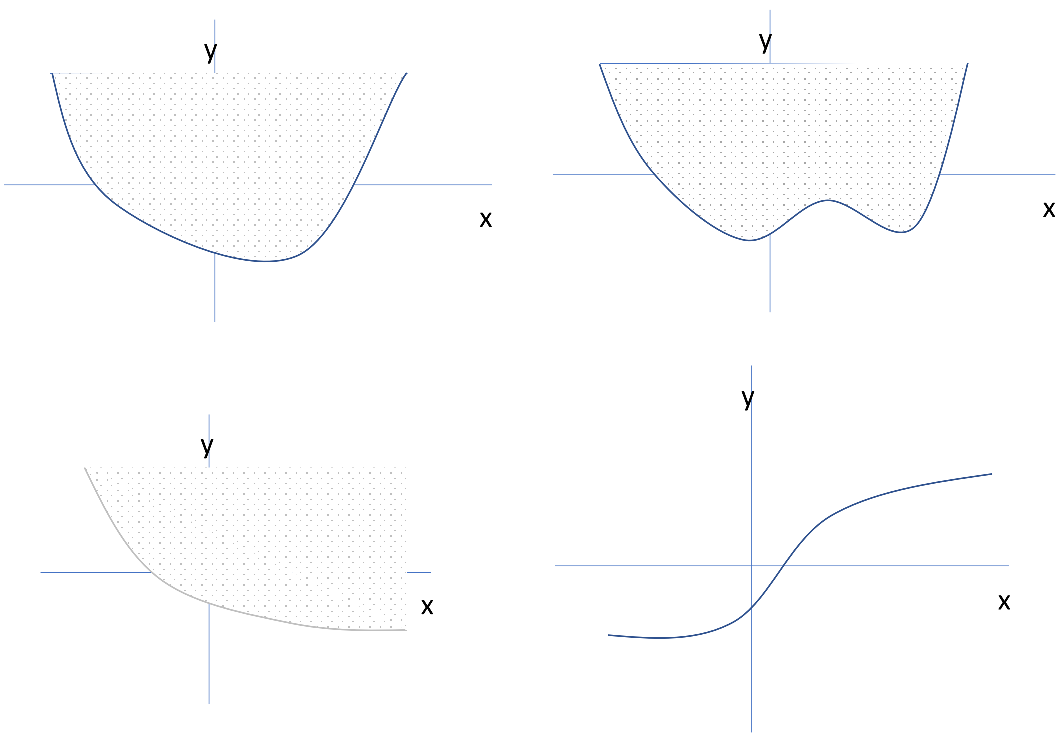
Which of these sets are convex?



Definition of Convex Functions: A function $f : \mathbb{R} \to \mathbb{R}$ is *convex* if the *epigraph* of $f(x)$ is a convex set.

The epigraph is the set of points lying on or above the graph of $f(x)$:

$$\text{epi} f = \{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$$



A function f is *concave* if $-f$ is convex, or equivalently, if the *subgraph* is a convex set.

Similar definitions apply if f is defined on a sub-interval or $\mathbb{R}$, of if f is defined on (a convex subset of) $\mathbb{R}^n$.

Proposition: A function $f : \mathbb{R} \to \mathbb{R}$ is *convex* if and only if, for all x, y and $0 \leq \lambda \leq 1$:

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

Proof We prove this in two stages. Firstly, we show that this definition implies that the epigraph is convex (the 'if' part), and then that a convex epigram implies this inequality (the 'only if' part).

If. Suppose that the inequality holds. We need to show that the epigraph is convex.

Suppose then that the vectors (x, a) and (y, b) are in the epigraph, which is equivalent to:

$$a \geq f(x)$$

$$b \geq f(y)$$

We then consider an intermediate point $(\lambda x+(1-\lambda)y, \lambda a+(1-\lambda)b)$. We then have:

$$\lambda a + (1-\lambda)b \geq \lambda f(x) + (1-\lambda)f(y)$$

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

Thus, the intermediate point lies in the epigraph of $f(x)$. This proves that the epigraph is convex.

Only if. Suppose that the epigraph of $f(x)$ is convex. Then we need to prove that $f(x)$ satisfies the inequality.

Let us then pick x, y in the domain of f. Then the vectors $(x, f(x))$ and $(y, f(y))$ lie in the epigraph of f.

By hypothesis, the epigraph is convex and so, for $0 \leq \lambda \leq 1$ the following point is in the epigraph:

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

By definition of the epigraph, this implies that:

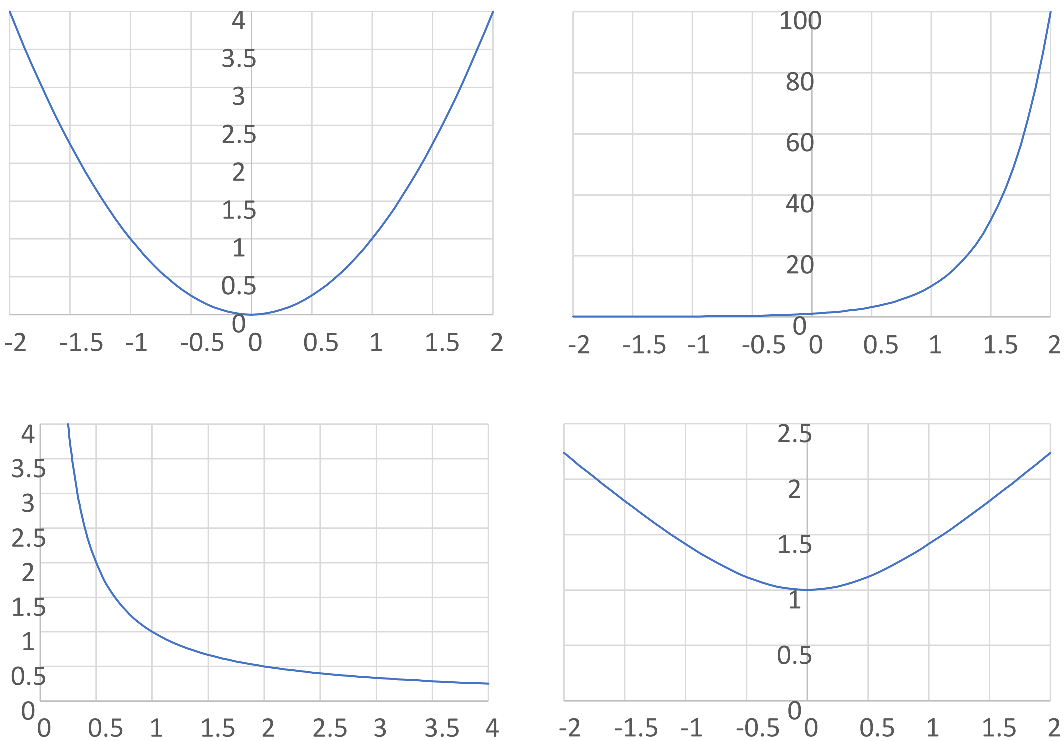
$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y)$$

Therefore, any convex function satisfies the inequality claimed.

We have therefore proved that the inequality holds for convex functions, and only for convex functions.

Examples of Convex Functions

- $y = x^2$
- $y = 10^x$
- $y = 1/x$ for $x > 0$.
- $y = \sqrt{1 + x^2}$



Jensen's Inequality:

Let $f(x)$ be a convex function and let $w_1, w_2, \ldots w_n$ be weights with

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

Then, for arbitrary $x_1, x_2, \ldots x_n$ Jensen's inequality states:

$$f\left(w_1 x_1 + w_2 x_2 + \ldots w_n x_n\right) \leq w_1 f(x_1) + w_2 f(x_2) + \ldots + w_n f(x_n)$$

Proof We proceed by induction on n, the number of weights.

If $n = 1$ then equality holds and the inequality is trivially true.

Let us suppose, inductively, that Jensen's inequality holds for $n = k-1$. We seek to prove the inequality when $n = k$.

Let us then suppose that $w_1, w_2, \ldots w_k$ be weights with

- $w_j \geq 0$
- $\sum_{j=1}^{k} w_j = 1$

If $w_k = 1$ then the inequality reduces to $f(x_k) \geq f(x_k)$ which is trivially true, so we concentrate on the case $w_k < 1$. Then, applying the inductive hypothesis to the $k-1$ points $x_1, x_2, \ldots x_k$:

$$f\left(\frac{w_1}{1-w_k} x_1 + \frac{w_2}{1-w_k} x_2 + \ldots + \frac{w_{k-1}}{1-w_k} x_{k-1}\right)$$
$$\leq \frac{w_1 f(x_1) + w_2 f(x_2) + \ldots w_{k-1} f(x_{k-1})}{1-w_k}$$

Trivially we also have:

$$f(x_k) \leq f(x_k)$$

Taking a weighted average of the last two formulas with weights $1 - w_k$ and w_k respectively, we have:

$$(1-w_k)f\left(\frac{w_1}{1-w_k}x_1 + \frac{w_2}{1-w_k}x_2 + \ldots + \frac{w_{k-1}}{1-w_k}x_{k-1}\right) + w_k f(x_k)$$

$$\leq w_1 f(x_1) + w_2 f(x_2) + \ldots w_{k-1} f(x_{k-1}) + w_k f(x_k)$$

But by the convexity of f we can compare the left hand side:

$$f(w_1x_1 + w_2x_2 + \ldots w_kx_k) \leq$$

$$(1-w_k)f\left(\frac{w_1}{1-w_k}x_1 + \frac{w_2}{1-w_k}x_2 + \ldots + \frac{w_{k-1}}{1-w_k}x_{k-1}\right) + w_k f(x_k)$$

Combining these last two inequalities, we finally have proved the inductive hypothesis when $n = k$:

$$f(w_1x_1 + w_2x_2 + \ldots w_kx_k)$$

$$\leq w_1 f(x_1) + w_2 f(x_2) + \ldots w_{k-1} f(x_{k-1}) + w_k f(x_k)$$

By induction we have proved Jensen's inequality for arbitrary positive integers n.

When does Equality Hold?

Equality holds in Jensen's inequality if:

- All the x_j are equal
- All but one of the w_j are zero.

If the function $f(x)$ is *strictly* convex then these are the only cases of equality.

Example Problem

Show that:

$$\sqrt{1^2+1}+\sqrt{2^2+1}+\ldots+\sqrt{n^2+1} \geq \frac{n}{2}\sqrt{n^2+2n+5}$$

Solution: Apply Jensen's inequality to the convex function $f(x) = \sqrt{1+x^2}$ at the points $x_n = n$ with weight $1/n$. Then

$$\begin{aligned}
&\frac{\sqrt{1^2+1}+\sqrt{2^2+1}+\ldots+\sqrt{n^2+1}}{n} \\
&\geq \sqrt{1+\left(\frac{1+2+\ldots+n}{n}\right)^2} \\
&= \sqrt{1+\frac{(n+1)^2}{4}} \\
&= \frac{1}{2}\sqrt{(n+1)^2+4}
\end{aligned}$$

Multiplying by n, we obtain the result we set out to prove.

Example: Suppose a, b and c are positive real numbers with:

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = a + b + c$$

Find the minimal value of this expression.

Solution By Jensen's inequality applied to the convex function $f(x) = 1/x$, for arbitrary $a, b, c > 0$:

$$\left[\frac{a+b+c}{3}\right]^{-1} \leq \frac{1}{3}\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right)$$

We can therefore conclude that:

$$\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right)(a + b + c) \geq 9$$

In this example, we are told the two factors are equal and positive; therefore:

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = a + b + c \geq 3$$

Equality holds when $a = b = c = 1$.

Example: Suppose $\{x_i : 1 \leq i \leq n\}$ are non-negative real numbers with

$$\sum_{i=1}^{n} x_i = 1$$

What is the lowest possible value of:

(1) $\sum_{i=1}^{n} x_i^2$

(2) $\sum_{i=1}^{n} \sqrt{x_i}$?

For the first problem, we note by the convexity of x^2 that

$$\left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right)^2 \leq \frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}$$

Therefore,

$$\sum_{i=1}^{n} x_i^2 \geq \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = \frac{1}{n}$$

Equality holds when all the x_i are equal to $1/n$.

For the second problem, we cannot apply Jensen's inequality because $\sqrt{x}$ is concave, not convex.

However, we note that for $0 \leq x \leq 1$:

$$\sqrt{x} \geq x$$

Therefore:

$$\sum_{i=1}^{n} \sqrt{x}_i \geq \sum_{i=1}^{n} x_i = 1$$

Equality holds when one of the $x_i = 1$ and all the others are zero.

Intersections of Convex Sets:

The intersection of a collection of convex sets is still convex:

Corollary The maximum of convex functions is convex (because the epigraph of the maximum is the intersection of the epigraphs).

Example: Arithmetic-Geometric Mean Inequality

Let $a_1, a_2, \ldots a_n \geq 0$. Then:

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

In this expression, the left hand side is the *geometric mean* and the right hand side is the *arithmetic mean*.

Proof If any of the a_j are zero then the result holds trivially setting the left hand side to zero.

So let us suppose all the a_j are strictly positive. Then we can write $a_j = 10^{x_j}$ for some (positive or negative) x_j.

Then applying Jensen's inequality to the convex function 10^x, with weights equal to $1/n$, we have:

$$10^{(x_1+x_2+\ldots x_n)/n} \leq \frac{1}{n} \sum_{j=1}^{n} 10^{x_j}$$

This is the inequality we set out to prove.

Note: Equality holds when all the a_j are equal.

Applications of AM $\geq$ GM.

Problem AMGM #1

If $\{b_1, b_2, \ldots b_n\}$ is a permutation of the sequence $\{a_1, a_2, \ldots a_n\}$ of positive real numbers, then show that:

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \ldots + \frac{a_n}{b_n} \geq n$$

Problem AMGM #2 Let a, b, c be positive real numbers. Show that:

$$a^3 + b^3 + c^3 \geq a^2b + b^2c + c^2a$$

Hint: Start by showing:

$$\frac{a^3 + a^3 + b^3}{3} \geq a^2b$$

Problem AMGM #3 Let a, b, c be positive real numbers. Show that:

$$\frac{c}{a} + \frac{a}{b+c} + \frac{b}{c} \geq 2$$

Hint Add 1 to each side and apply AM $\geq$ GM.

Example Problem Let $x_1, x_2, \ldots x_n$ and $y_1, y_2, \ldots y_n$ be real sequences, satisfying:

$$\sum_{i=1}^{n} |x_i|^p = 1$$

$$\sum_{i=1}^{n} |y_i|^q = 1$$

Here, the exponents $p, q > 1$ satisfy:

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

Prove that

$$-1 \leq \sum_{i=1}^{n} x_i y_i \leq 1$$

Solution: Without loss of generality, we may assume that $x_i > 0$ and $y_i > 0$, otherwise we could increase the absolute value of the left hand side by replacing each x and y by their absolute value.

We apply Jensen's inequality to the convex function $f(z) = z^p$, writing:

$$w_i = y_i^q$$
$$z_i = \frac{x_i}{y_i^{q-1}}$$

Jensen's inequality then implies:

$$\left[\sum_{i=1}^{n} x_i y_i\right]^p \leq \sum_{i=1}^{p} y_i^q \frac{x_i^p}{y_i^{p(q-1)}} = \sum_{i=1}^{p} x_i^p = 1$$

In the middle step, the $y's$ cancel because the exponent is zero:

$$q - p(q-1) = pq\left(\frac{1}{p} - 1 + \frac{1}{q}\right) = 0$$

Taking the p^{th} root of the previous inequality gives the result we set out to prove.

Remark. This result is more often stated in the equivalent form:

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$$

It is known as Hölder's Inequality.

Unit Balls and Duality

Hölder's inequality is a special case of a profound result in the theory of convex sets.

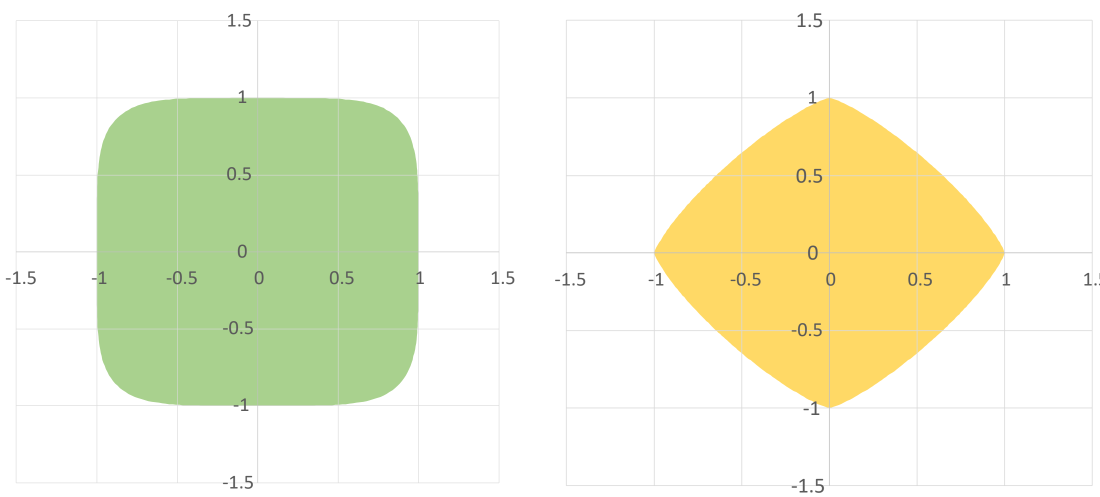
Let a *unit ball* $\mathcal{B} \subset \mathbb{R}^n$ be a closed, bounded, convex set containing a neighborhood of the origin. An example of such a unit ball is:

$$\mathcal{B} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^{n} |x_i|^p \leq 1 \right\}$$

Then the *dual ball*, $\mathcal{B}'$, is defined by:

$$\mathcal{B}' = \{ y \in \mathbb{R}^n : x.y \leq 1, \forall x \in \mathcal{B} \}$$

Hölder's inequality identifies the dual ball in our example. These are shown in $\mathbb{R}^2$ for $p = 5$ and $q = 1.25$.



Generalising Factorials to Non-Integers

The factorial function $n!$ is defined for non-negative integers n by:

$$0! = 1$$

$$1! = 1$$

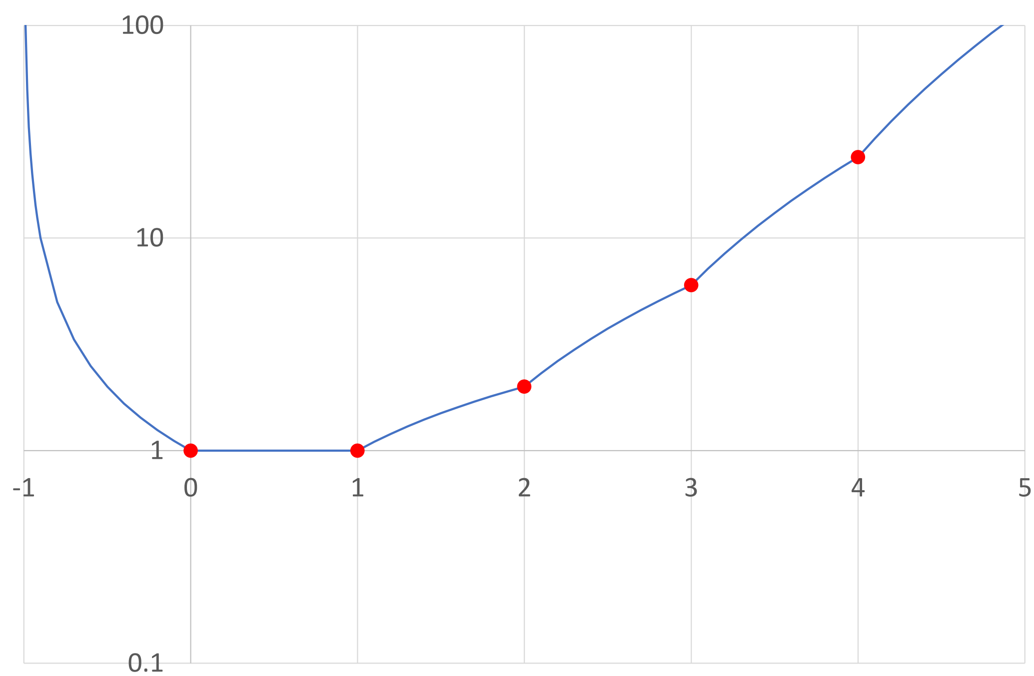
$$2! = 2$$

$$n! = n \times (n-1)!$$

Question: Is there a natural generalisation of $n!$ to non-integer n?

We could generalise $n!$ by choosing an arbitrary function for $0 < n < 1$, for example 1, and then applying the recurrence relation for larger n.

Plotting this on a logarithmic scale, we get the following function:



Convex Generalised Factorial

Let us suppose we want the generalised factorial to be convex when plotted on a geometric scale (unlike the plot above).

In other words, we want to define $x!$ for $x > -1$ such that

- $x! = x \times (x-1)!$
- $x! = 10^{f(x)}$ where $f(x)$ is a convex function.

In particular, taking $0 < x < 1$, the definition of convexity implies:

$$f(n+x) \leq (1-x)f(n) + xf(n+1)$$
$$f(n) \leq xf(n-1+x) + (1-x)f(n+x)$$

Raising 10 to the power of each side, we have:

$$(n+x)! \leq (n!)^{1-x}[(n+1)!]^x$$
$$n! \leq [(n-1+x)!]^x[(n+x)!]^{1-x}$$

Now using the recurrence relation for the factorial, we have:

$$(n+x)! \leq (n+1)^x \times n!$$
$$n! \leq (n+x)^{-x} \times (n+x)!$$

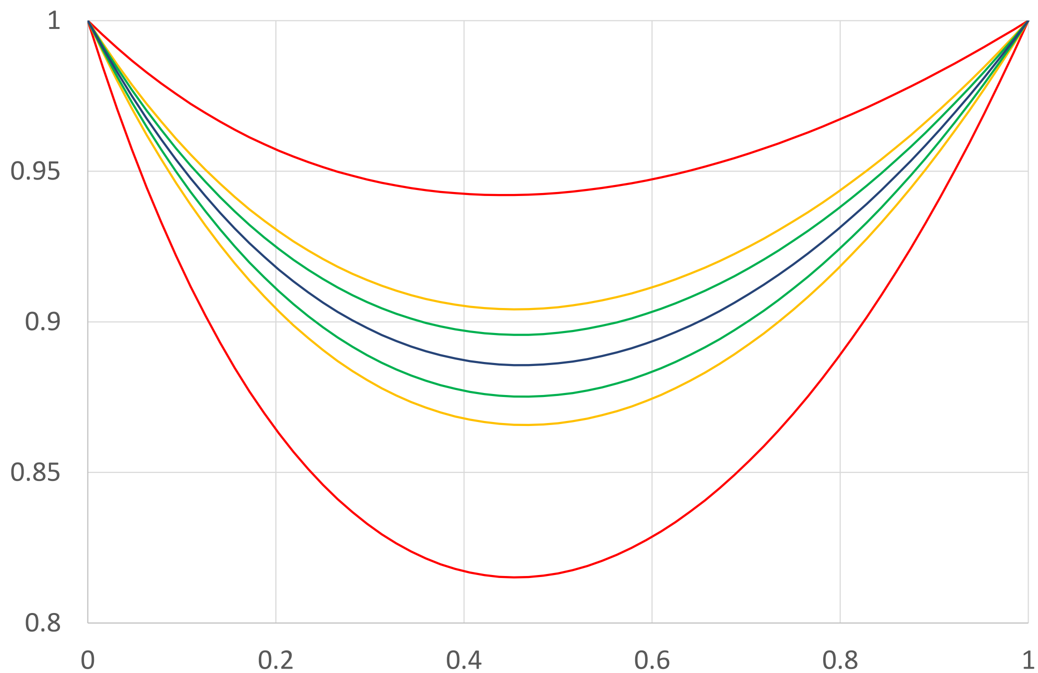
Putting these together, we have upper and lower bounds for $(n+x)!$:

$$(n+x)^x n! \leq (n+x)! \leq (n+1)^x \times n!$$

Dividing through by $(1+x)(2+x)\dots(n+x)$ we have:

$$\frac{(n+x)^x n!}{\prod_{j=1}^{n}(j+x)} \le x! \le \frac{(n+1)^x \times n!}{\prod_{j=1}^{n}(j+x)}$$

This gives upper and lower bounds for x!. The chart below shows the upper and lower bounds for $0 \le x \le 1$ and $n = 1$ (red), $n = 5$ (orange), $n = 10$ (green) and the limit of large n (black):



There is much more that can be said about this function, including that $\frac{1}{2}! = \frac{\sqrt{\pi}}{2}$, but this will have to wait for another time!