Convex Functions with applications to inequalities.

Proposition 1. *Let $x < y$ be real numbers and $\alpha, \beta > 0$ with $\alpha + \beta = 1$. Then the linear combination $\alpha x + \beta y$ is between x and y. That is $x < \alpha x + \beta y < y$.*

Proof. Write $\alpha x + \beta y = x - x + \alpha x + \beta y = x - (1-\alpha)x + \beta y = x - \beta x + \beta y = x + \beta(y - x)$. But $x < y$ so $(y - x) > 0$ and $0 < \beta < 1$ and thus $0 < \beta(y - x) < (y - x)$. There

$$x < \alpha x + \beta y = x + \beta(y - x) < x + (y - x) = y$$

as required. □

Remark 2. If we do not make the assumption that $x < y$ we can just say that $\alpha x + \beta y$ is between x and y. That is, when $x \neq y$, we have $\min\{x, y\} < \alpha x + \beta y < \max\{x, y\}$. □

Definition 3. Let x, y be real numbers. Then a ***convex combination*** of x and y is a linear combination of the form $\alpha x + \beta y$ where $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$.

Thus Proposition 1 tells us that the convex combination of two real numbers x and y is between x and y. We can make a more general definition

Definition 4. Let $x_1, \ldots, x_n$ be real numbers. Then a ***convex combination*** of these numbers is a linear combination of the form

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \sum_{k=1}^{n} \alpha_k x_k$$

where

$$\alpha_1, \ldots, \alpha_n > 0 \quad \text{and} \quad \alpha_1 + \cdots + \alpha_n = \sum_{k=1}^{n} \alpha_k = 1.$$

The following is useful in the induction step of a couple of the proofs below.

Lemma 5. *Let $\alpha_1, \ldots, \alpha_{n+1} > 0$ with $\alpha_1 + \cdots + \alpha_{n+1} = 1$. Then for any real numbers $x_1, \ldots, x_{n+1}$ we have*

$$\sum_{k=1}^{n+1} \alpha_k x_k = (1 - \alpha_{n+1}) \sum_{k=1}^{n} \left(\frac{\alpha_k}{1 - \alpha_{n+1}} \right) x_k + \alpha_{n+1} x_{n+1}.$$

and

$$\sum_{k=1}^{n} \left(\frac{\alpha_k}{1 - \alpha_{n+1}} \right) = 1.$$

Problem 1. Prove this. □

Remark 6. One way to think about the last lemma is that if x is a convex combination of $x_1, \dots, x_{n+1}$, then x can be written as

$$x = \alpha y + \beta x_{n+1}$$

where $\alpha = 1 - \alpha_{n+1} > 0$, $\beta = \alpha_{n+1} > 0$ (so that $\alpha + \beta = 1$) and y is a convex combination of $x_1, \dots, x_n$. This is exactly the set up needed for induction proofs. □

Proposition 7. *Let x be a convex combination of $x_1, \dots, x_n$. Then*

$$\min\{x_1, \dots, x_n\} \leq x \leq \max\{x_1, \dots, x_n\}.$$

(The reason that we have "$\leq$" rather than "$<$" is to cover the case when $x_1 = x_2 = \cdots = x_n$. In all other cases the inequalities are strict.)

Problem 2. Prove this. *Hint:* See 2 (for the base case) and Remark 6 (for the induction stop).

Definition 8. A function f defined on an interval I is ***convex*** iff for all $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and all $x, y \in I$ the inequality

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y) \tag{1}$$

holds. □

Definition 9. A function f defined on an interval I is ***strictly convex*** iff for all $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and all $x, y \in I$ with $x \neq y$ the inequality

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$$

holds. □

Remark 10. Anther way to say that f is strictly convex is that equality holds in the inequality (1) if and only if $x = y$. □

In the terminology of many calculus books this is the same as being concave up. In terms of the graph of f, the condition that f is convex is that f is bellow any of its secant segments (see Figure 1).

Problem 3. Show that $f(x) = x$ and $g(x) = |x|$ are convex on $\mathbb{R}$. *Hint:* For the absolute value, use the triangle inequality. □

Next is a basic result about convex functions.

Theorem 11 (Jensen's inequality)**.** *If f is convex on the interval I, $x_1, \dots, x_n \in I$ and $\alpha_1, \dots, \alpha_n > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$, then*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

If f is strictly convex, then equality holds if and only if $x_1 = x_2 = \cdots = x_n$.

Problem 4. Prove this. *Hint:* See the hint to Problem 2. □

If would be nice to have an easily checked criterion that implies that f is convex. You most likely recall from calculus that a function is concave up, that is convex, if its second derivative is positive. As a first step in toward proving this we have

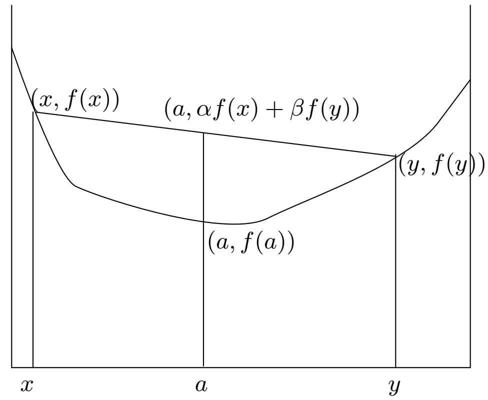

FIGURE 1. Here $a = \alpha x + \beta y$ Using that $(a, \alpha f(x) + \beta f(y)) = (\alpha x + \beta y, \alpha f(x) + \beta f(y)) = \alpha(x, f(x)) + \beta(y, f(y))$ is on the line segment connecting $(x, f(x))$ and $(y, f(y))$ we see that, geometrically, the inequality defining convex functions is equivalent to requiring that the graph $y = f(x)$ lies under the secant connecting any two points on the graph.

Proposition 12. *Let f be twice differentiable on the open interval I with $f''(x) \geq 0$ for all $x \in I$. Then for any $a \in I$*

$$f(x) \geq f(a) + f'(a)(x - a) \tag{2}$$

for all $x \in I$. If the stronger condition $f''(x) > 0$ holds for all $x \in I$ then equality holds in (2) if and only if $x = a$.

Proof. This is a straightforward application of Taylor's theorem. From Taylor's theorem with Lagrange's form of the remainder we have

$$f(x) = f(a) + f'(a)(x - a) + f''(\xi)\frac{(x-a)^2}{2} \geq f(a) + f'(a)(x - a)$$

as $f''(\xi)\dfrac{(x-a)^2}{2} \geq 0$ because $(x - a)^2 \geq 0$ and we are assuming $f'' \geq 0$. If $f'' > 0$ then equality can only hold if $x = a$. $\square$

Recall that $y = f(a) + f'(a)(x - a)$ is the equation of the tangent line to the graph of $y = f(x)$ at the point $(a, f(a))$. Therefore Proposition 12 tells us that if $f'' \geq 0$, then the graph of $y = f(x)$ lies above all its tangent lines. See Figure 2.

Theorem 13. *Let f be twice differentiable on the open interval I and with $f'' \geq 0$ on I. Then f is convex on I. If $f''(x) > 0$ for all $x \in I$, then f is strictly convex.*

Problem 5. Prove this. *Hint:* Let $x, y \in I$. If $x = y$ there is nothing to prove (as the inequality (1) reduces to $f(x) = f(x)$). So assume $x \neq y$. Let $\alpha, \beta > 0$ with $\alpha + \beta = 1$ and set

$$a = \alpha x + \beta y.$$

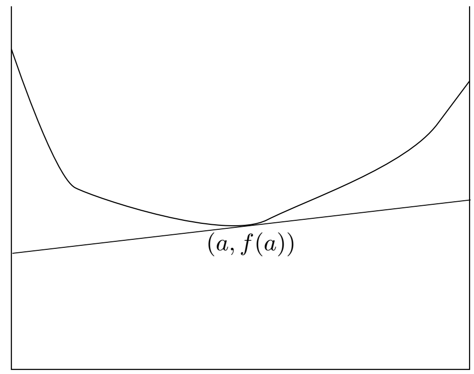

FIGURE 2. If $f'' \geq 0$, then the second order Taylor's theorem tells us

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + f''(\xi)\frac{(x-a)^2}{2} \\ &\geq f(a) + f'(a)(x-a) \end{aligned}$$

As $y = f(a) + f'(a)(x-a)$ is the equation of the tangent line to the graph of $y = f(x)$ at $(a, f(a))$ the graph of f is lies above all of its tangent lines. If $f''(\xi) > 0$ then equality can only if $x = a$, that is the graph $y = f(x)$ is strictly about the tangent line except at the point of tangency.

Then we wish to show

$$f(a) \leq \alpha f(x) + \beta f(y). \tag{3}$$

From Proposition 12 we know

$$f(x) \geq f(a) + f'(a)(x-a), \qquad f(y) \geq f(a) + f'(a)(y-a).$$

Multiply the first of these by α and the second by β and add to get an inequality for $\alpha f(x) + \beta f(y)$ and show that this simplifies to (3). Then show if $f'' > 0$ that this inequality is strict.

It is now easy to check (just by computing the second derivative and noting it is positive) the following

Proposition 14. *The following are strictly convex on the indicted intervals.*

(a) $f(x) = x^n$ *where* n *is an integer with* $n \geq 2$ *and* $I = (0, \infty)$.
(b) $f(x) = e^x$ *on* $I = \mathbb{R}$.
(c) $f(x) = -\ln(x)$ *on* $I = (0, \infty)$.
(d) $f(x) = x^{2n}$ *where* $n \geq 1$ *is an integer on* $I = \mathbb{R}$. *(Showing this is strictly convex takes a bit of work.)* □

We recall the ***Arithmetic-Geometric mean inequality***. This is that if a, b are positive real numbers, then

$$\sqrt{ab} \leq \frac{a+b}{2} \tag{4}$$

and equality holds if and only if $a = b$. The proof is simple

$$\frac{a+b}{2} - \sqrt{ab} = \frac{a - 2\sqrt{a}\sqrt{b} + b}{2} = \frac{\left(\sqrt{a} - \sqrt{b}\right)^2}{2} \geq 0$$

and equality can only hold if $\sqrt{a} = \sqrt{b}$. That is if only if $a = b$. The number $\sqrt{ab}$ is the ***geometric mean*** of a and b, while $\frac{a+b}{2}$ is the ***arithmetic mean*** of a and b, which is where the inequality gets its name. It can be greatly generalized.

Theorem 15 (Generalized Arithmetic-Geometric Mean Inequality)**.** *Let $\alpha_1, \ldots, \alpha_n > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$. Then for any positive real numbers $a_1, \ldots, a_n$ the inequality*

$$a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_n^{\alpha_n} \leq \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$$

holds. Equality holds if and only if all the a_j's are equality.

Problem 6. Prove this. *Hint:* We know that the function $f(x) = e^x$ is strictly convex on $\mathbb{R}$. That is for any real numbers $x_1, \ldots, x_n$ we have

$$f(\alpha_1 x_1 + \cdots \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$$

and equality holds if and only if all the x_j's are equal. Show this can be rewritten as

$$(e^{x_2})^{\alpha_1} (e^{x_2})^{\alpha_2} \cdots (e^{x_n})^{\alpha_n} \leq \alpha_1 e^{x_1} + \alpha_2 e^{x_2} + \cdots + \alpha_n e^{x_n}$$

and equality holds if and only if all the x_j's are equal.

Now given positive numbers $a_1, \ldots, a_n$ there are unique real numbers $x_1, \ldots, x_n$ with $a_j = e^{x_j}$ for all $j = 1, 2, \ldots, n$. (You can assume these x_j's exist.) And you take it from here.

Remark 16. In different notation the generalized Arithmetic-Geometric inequality is

$$\prod_{k=1}^{n} a_k^{\alpha_k} \leq \sum_{k=1}^{n} \alpha_k a_x$$

with equality holding if and only if all the a_k's are equal. □

The can you may have seen before is

$$\sqrt[n]{a_1 a_2 \cdots a_n} \leq \frac{a_1 + \cdots + a_n}{n}$$

coming form $\alpha_1 = \alpha_2 = \cdots = a_n = 1/n$ and equality holds if and only if all the a_j's are equal. The can of $n = 2$ is often useful. Then letting $\alpha = \alpha_1$ and $\beta = \alpha_2$ we have

$$a^\alpha b^\beta \leq \alpha a + \beta b$$

with equality holding if and only if $a = b$. (And as usual $\alpha, \beta > 0$ with $\alpha + \beta = 1$.)

Here is an example of the use of the generalized arithmetic geometric mean inequality

Example 17. For $x, y, z \geq 0$ maximize the product xyz subject to the constraint $x + y + z = c$, where c is a constant. We have

$$xyz = \left((xyz)^{1/3}\right)^3 \leq \left(\frac{x+y+z}{3}\right)^3 = \left(\frac{c}{3}\right)^3$$

and equality holds if and only of $x = y = z$. Thus the maximum is $(c/3)^3$ with equality if any only if $x = y = z = c/3$. $\square$