Some applications of Taylor's Theorem.

The next result is just the second derivative test from calculus.

Theorem 1. *Let I be an open interval and $f\colon I \to \mathbb{R}$ twice differentiable with f'' continuous. Assume that $f'(x_0) = 0$ then,*

- *if $f''(x_0) < 0$ then x_0 is a local maximizer of f.*
- *if $f''(x_0) > 0$ then x_0 is a local minimizer of f.*

Problem 1. Prove this. *Hint:* It is enough to prove in the case $f''(x_0) > 0$.

(a) Use that f'' is continuous to show that there is $\delta > 0$ such that $f'' > 0$ on the interval $(x_0 - \delta, x_0 + \delta)$.

(b) Now use Taylor's Theorem to show for $x \in (x_0 - \delta, x_0 + \delta)$ that

$$f(x) \geq f(x_0)$$

and that equality holds if and only if $x = x_0$. $\square$

If we know that the second derivative has the same sign on an entire interval we can get a global maximum or minimum.

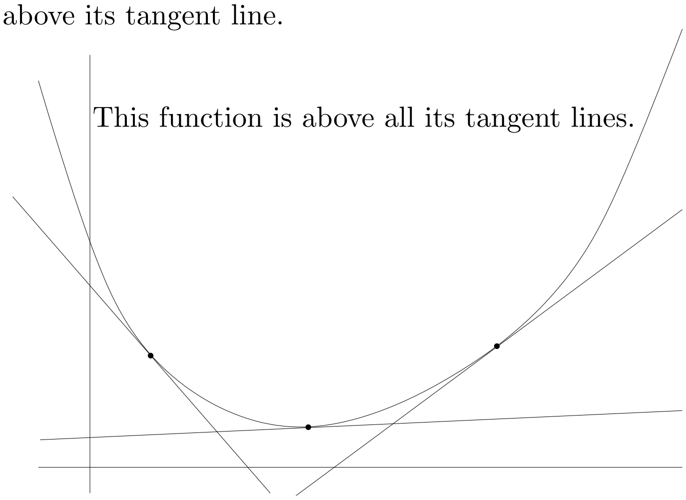
Theorem 2. *Let I be an interval and $x_0 \in I$ in the interior of I. Assume that f'' exists and $f'' \geq 0$ at interior points of I. Then $f'(x_0) = 0$ implies that*

$$f(x) \geq f(x_0)$$

for all $x \in I$. (And if $f'' < 0$ on the interior of I, then $f(x) \leq f(x_0)$ on I.)

Problem 2. Prove this in the case of $f'' > 0$. $\square$

The last result can be generalized to giving a condition for a function to always be above its tangent line.



Theorem 3. *Let $f'' \geq 0$ on an open interval I. Then the graph of f is above all its tangent lines. More precisely if $a \in I$, then*

$$f(a) + f'(a)(x - a) \leq f(x)$$

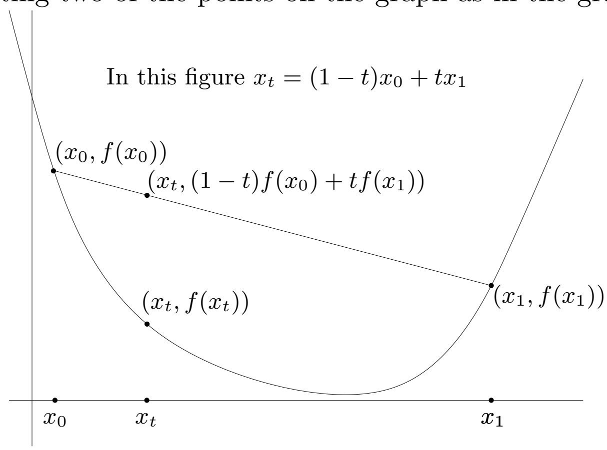
for all $x \in I$.

Problem 3. Prove this. $\square$

Let $f\colon I \to \mathbb{R}$ where I is an interval. Then f is ***convex*** if and only if for all $x_0, x_1 \in \mathbb{R}$ and $t \in [0,1]$ the inequality

$$f((1-t)x_0 + tx_1) \le (1-t)f(x_0) + tf(x_1)$$

holds. Geometrically this means that the graph of f lies below any of the chords connecting two of the points on the graph as in the graph below.



Theorem 4. *Let I be an open interval and $f\colon I \to \mathbb{R}$ be a function that is twice differentiable and with $f'' \ge 0$. Then f is convex on I.*

Problem 4. Prove this. *Hint:* To simplify notation let

$$x_t = (1-t)x_0 + tx_1.$$

By Theorem 3 we have that the graph of f is above its tangent line at x_t, which implies

$$\begin{aligned} f(x_t) + f'(x_t)(x_0 - x_t) &\le f(x_0) \\ f(x_t) + f'(x_t)(x_1 - x_t) &\le f(x_1). \end{aligned}$$

Show

$$\begin{aligned} x_0 - x_t &= -t(x_1 - x_0) \\ x_1 - x_t &= (1-t)(x_1 - x_0). \end{aligned}$$

And therefore

$$\begin{aligned} f(x_t) - tf'(x_t)(x_1 - x_0) &\le f(x_0) \\ f(x_t) + (1-t)f'(x_t)(x_1 - x_0) &\le f(x_1). \end{aligned}$$

Now manipulate these inequalities to complete the proof. $\square$

Proposition 5. *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 6. 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 7. Let x, y be real numbers. Then a ***convex combination***, also called a ***weighted average***, 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 5 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 8. Let $x_1, \ldots, x_n$ be real numbers. Then a ***convex combination*** (and again this is often called a ***weighted average***) 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 9. *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 5. Prove this. □

Remark 10. One way to think about the last lemma is that if x is a convex combination of $x_1, \ldots, x_{n+1}$, 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 $x_1, \ldots, x_n$. This is exactly the set up needed for induction proofs. □

Proposition 11. *Let x be a convex combination of $x_1, \dots, x_n$. Then*

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

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

Problem 6. Prove this. *Hint:* See Remark 6 (for the base case) and Remark 10 (for the induction stop).

Definition 12. 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

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

holds. □

Definition 13. 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 $x \ne y$ the inequality

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

holds. □

Remark 14. Anther way to say that f is strictly convex is that equality holds in the inequality (1) if and only if $x = y$. □

Problem 7. 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 15 (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*

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

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

Problem 8. Prove this. *Hint:* See the hint to Problem 6. □

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

Proposition 16. *Let f be twice differentiable on the open interval I with $f''(x) \ge 0$ for all $x \in I$. Then for any $a \in I$*

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

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 $x = a$. □

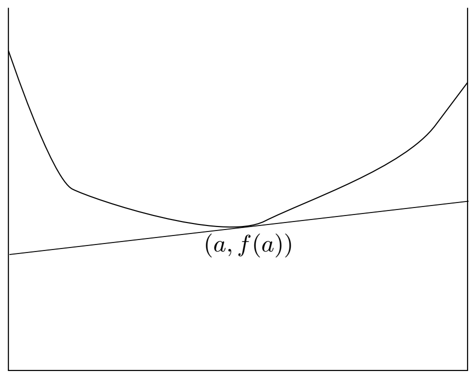


FIGURE 1. If $f'' \geq 0$, then the second order Taylor's theorem tells us

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

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.

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 16 tells us that if $f'' \geq 0$, then the graph of $y = f(x)$ lies above all its tangent lines. See Figure 1.

Theorem 17. *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 9. 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

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

Then we wish to show

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

From Proposition 16 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 18. *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 19 (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 10. 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 20. In different notation the generalized Arithmetic-Geometric inequality is

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

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

The can you may have seen before is

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

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

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

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

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

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

Here is a bit more challenging problem.

Problem 11. Show that if $f\colon I \to \mathbb{R}$ is a continuous function on an interval I, such that for all $x, y \in I$ the inequality

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x) + f(y)}{2} \tag{5}$$

holds, then f is convex on I. □

Remark 22. A function that satisfies the inequality (5) is called ***midpoint convex***. So the problem is asking you to show that a continuous midpoint convex function is convex. Without the assumption of continuity this is false. That is there are midpoint convex functions that are not convex. However they are not continuous. Examples of such functions are hard to produce and require using some version of the Axiom of Choice. □