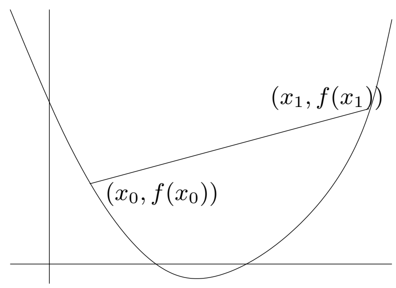
Convex functions and divided differences.

1. Convex functions and Jensen's inequality.

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

$$f((1-t)x_0 + tx_1) \le (1-t)f(x_0) + tf(x_1). \tag{1}$$

It is ***strictly convex*** if and only if equality in (1) and $t \ne 0, 1$ implies $x_0 = x_1$. Geometrically f being convex on an interval means that if lies below the segment connecting a pair of points on its graph.



Problem 1. Let $x_0, x_1, \dots, x_n \in I$ and $t_0, t_1, \dots, t_n \in [0,1]$ with $\sum_{k=1}^n t_k = 1$. Show for any convex f that

$$f(t_0x_0 + t_1x_1 + \cdots + t_nt_n) \le t_0f(x_0) + t_1f(x_1) + \cdots + t_nf(x_n). \qquad \square$$

Problem 2. In the previous problem assume that f is strictly convex and each $t_k > 0$. Show equality holds if and only if $x_0 = x_1 = \cdots = x_n$. $\square$

Problem 3. Let $f : (a,b)\colon \mathbb{R}$ be differentiable on (a,b) and assume f' is monotone increasing. Show that the graph of f is above any of its tangent lines. That is for all $x_*, x \in (a,b)$ the inequality

$$f(x_*) + f'(x_*)(x - x_*) \le f(x).$$

Hint: By the Mean Value Theorem

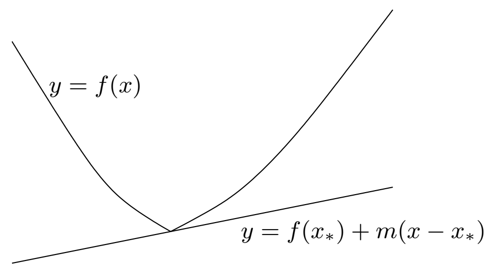
$$f(x) - f(x_*) = f'(\xi)(x - x_*)$$

where ξ is between x_0 and x. The monotonicity of f' implies that if $x > x_*$, then $f'(x_*) \le f'(\xi)$, and if $x < x_*$, then $f'(\xi) \le f'(x_*)$. $\square$

Definition 1. Let $f\colon I \to \mathbb{R}$ be a function on an interval I. The f has a ***lower support line*** at $x_* \in I$ if and only if there is a constant m such that

$$f(x_*) + m(x - x_*) \le f(x)$$

for all $x \in I$. Informally it we think of $y = f(x_*) + m(x - x_*)$ as a generalization of a tangent line to the graph of $f(x)$, this is saying that the graph $y = f(x)$ is above the tangent line, so this is an axiomatization of the property of differentiable convex functions given in Problem 3 $\square$



Note that a function can have several lower support lines at a point. For example the function $f(x) = |x|$ has as lower support lines at $x = 0$ all of the lines $y = mx$ for $-1 \le m \le 1$.

Proposition 2. *If $f(x)$ is a function on an open interval (a, b) that has a lower support line at each point, then f is convex on (a, b). In particular if f' exists and is monotone increasing on (a, b), then f is convex.*

Problem 4. Prove this. *Hint:* Let $x_0, x_1 \in (a, b)$ and for $t \in [a, b]$ let $x_t = (1 - t)x_0 + tx_1$. Let $y = f(x_t) + m(x - x_t)$ be a lower support function to f at x_t. Then

$$\begin{aligned} f(x_0) &\ge f(x_t) + m(x_0 - x_t) \\ &= f(x_t) - tm(x_1 - x_0) \\ f(x_1) &\ge f(x_t) + m(x_1 - x_0) \\ &= f(x_t) + (1 - t)m(x_1 - x_0). \end{aligned}$$

Use these inequalities in the expression $(1 - t)f(x_0) + tf(x_1)$. □

Problem 5. Show that if f is twice differentiable on (a, b) and $f'' \ge 0$, then f is convex on (a, b). Also show that if $f'' > 0$ then f is strictly convex (this will involve checking what happens in the previous results when f' is strictly increasing). □

Proposition 3 (Jensen's Inequality Form 1)**.** *Let $\varphi\colon (a, b) \to \mathbb{R}$ be defined on (a, b) and assume that f has a lower support line at each point of (a, b). Let (X, μ) be a measure space with $\mu(X) = 1$ and $f \in L^1(X)$ with $f(\xi) \in (a, b)$ for all $\xi \in X$. Then*

$$\varphi\left(\int_X f(\xi)\, d\mu(\xi)\right) \le \int_X \varphi(f(\xi))\, d\mu(\xi).$$

Problem 6. Prove this. *Hint:* Use $\mu(X) = 1$ to show that the number $x_* = \int_X f\, d\mu$ is in (a, b). Let $y = \varphi(x_*) + m(x - x_*)$ be a lower support function to φ at $x = x_*$. Then for all $\xi \in X$

$$\varphi(x_*) + m(f(\xi) - x_*) \le \varphi(f(\xi)).$$

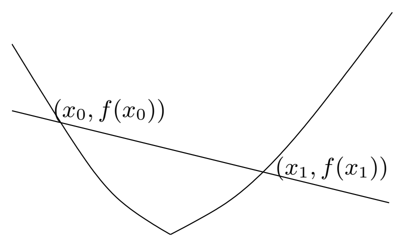
Show that integrating this over X (and again using $\mu(X) = 1$) gives

$$\varphi(x_*) + m\left(\int_X f(\xi)\, d\mu(\xi) - x_*\right) \le \int_X \varphi(f(\xi))\, d\mu(\xi).$$

and using the definition of x_* this reduces to $\varphi(x_*) \leq \int_X \varphi(f(\xi))\, d\mu(\xi)$ which is just what we wanted to prove. $\square$

Proposition 4. *Let f be convex on the interval I and let $x_0, x_1 \in I$ with $x_0 < x_1$. Let $y = ax + b$ be the line through $(x_0, f(x_1))$ and $(x_1, f(x_1))$. Then for $x \in I$,*

$$\begin{cases} f(x) \leq ax + b & \textit{for } x_0 \leq x \leq x_1, \\ f(x) \geq ax + b & \textit{for } x < x_0 \textit{ and } x > x_1. \end{cases}$$



Problem 7. Prove this. *Hint:* That $f(x) \leq ax + b$ for $x_0 \leq x \leq x_1$ follows from the definition of f being convex. Let $x \in I$ with $x < x_0$. Then x_0 is between x and x_1 so there is a s with $0 < s < 1$ and $x_0 = (1 - s)x_0 + sx_1$ As f is convex

$$f(x_0) \leq (1 - s)f(x) + tf(x_1)$$

As the line line $y = ax+b$ goes through the points $(x_0, f(x_1))$ and $(x_1, f(x_1))$ we have $f(x_0) = ax_0 + b$ and $f(x_1) = ax_1 + b$. Thus

$$ax_0 + b \leq (1 - s)f(x) + s(ax_1 + b),$$

which can be rearranged to give

$$a(x_0 - sx_1) + (1 - s)b \leq (1 - s)f(x)$$

now use $x_0 - sx_1 = (1 - s)x$ in this to complete the proof in the case $x < x_0$. The proof for $x > x_1$ is similar. $\square$

Lemma 5. *If f is convex on (a, b) and $x_1 \in (a, b)$, then the function*

$$M(x_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \textit{slope of line through } (x_0, f(x_1)) \textit{ and } (x_1, f(x_1))$$

is monotone increasing and bounded above on (a, x_1). (As we will see shortly, it is also monotone increasing on (x_1, b).)

Problem 8. Prove this. *Hint:* One method is to use Proposition 4. $\square$

Proposition 6. *If f is convex on the interval (a, b), then f has a lower support line at each point of (a, b).*

Problem 9. Prove this. *Hint:* Let $x_1 \in (a, b)$. Let $M(x_0)$ be as in Lemma 5 for x_0 and consider the line

$$y = f(x_1) + M(x_0)(x - x_0).$$

This line goes through $(x_0, f(x_1))$ and $(x_1, f(x_1))$ and if is below the graph of $y = f(x)$ except on the interval (x_0, x_1). Let

$$m = \lim_{x_0 \uparrow x_1} M(x_0).$$

This limit exists as $M(x_0)$ is a monotone increasing function which is bounded above. Now show $y = f(x_1) + m(x - x_1)$ is a lower support line to f at $x = x_1$. $\square$

Proposition 7 (Jensen's Inequality Usual Form)**.** *Let $\varphi\colon (a,b) \to \mathbb{R}$ be convex on (a,b) and let (X,μ) be a measure space with $\mu(X) = 1$ and $f \in L^1(X)$ with $f(\xi) \in (a,b)$ for all $\xi \in X$. Then*

$$\varphi\left(\int_X f(\xi)\, d\mu(\xi)\right) \le \int_X \varphi(f(\xi))\, d\mu(\xi).$$

Problem 10. Prove this. $\square$

2. Divided differences and convex functions.

Above we have shown that if a has $f'' \geq 0$, that it is convex. Here we use the notation a divided difference to generalize this to the case where f does not have to have two, or even one, derivative.

Definition 8. Let $f\colon (a,b) \to \mathbb{R}$ be a function and for $x_0, x_1 \in (a,b)$ with $x_0 \neq x_1$ define

$$f[x_0, x_1] = \frac{f(x_0) - f(x_1)}{x_0 - x_1}.$$

This is the ***divided difference*** of f. $\square$

The number $f[x_0, x_1]$ is the slope of the line through the points $(x_0, f(x_0))$ and $(x_1, f(x_1))$. And clearly when f is differentiable at x_0

$$f'(x_0) = \lim_{x_1 \to x_0} f[x_0, f_1].$$

Proposition 9. *Let $x_0, x_1 \in (a,b)$ then $f[x_0, x_1]$ is symmetric in x_0, x_1, that is $f[x_1, x_0] = f[x_0, x_1]$. And if f' exists on (a,b) there is a ξ between x_0 and x_1 with*

$$f[x_0, x_1] = f'(\xi).$$ $\square$

Problem 11. Prove this. *Hint:* The second statement just a restatement of the Mean Value Theorem. $\square$

The next result shows that the divided difference does for non-differentiable functions some of the jobs that the derivative does for differentiable.

Proposition 10. *Let $f\colon (a,b) \to \mathbb{R}$.*

(a) If $f[x_0, x_1] \equiv 0$, then f is constant on (a,b).
(b) The function f is monotone increasing if and only if $f[x_0, x_1] \geq 0$ for all $xx_0, x_1 \in (a,b)$ with $x_0 \neq x_1$. $\square$

Problem 12. Prove this. $\square$

We also want a general version of the second derivative.

Definition 11. Let $f\colon (a,b) \to \mathbb{R}$ for distinct points $x_0, x_1, x_2 \in (a,b)$ let

$$f[x_0,x_1,x_2] = \frac{f[x_0,x_1] - f[x_0,x_2]}{x_1 - x_2}.$$

This is the ***second divided difference*** of f. □

Proposition 12. *The second divided difference can be written as*

$$f[x_0,x_1,x_2] = \frac{(x_1-x_2)f(x_0) + (x_2-x_0)f(x_1) + (x_0-x_1)f(x_2)}{(x_1-x_2)(x_1-x_0)(x_2-x_1)}$$

This function is symmetric in x_0, x_1, x_2. That is for any permutation, σ, of $\{0,1,2\}$ we have $f[x_{\sigma(0)}, x_{\sigma(1)}, x_{\sigma(2)}] = f[x_0,x_1,x_2]$.

Problem 13. Prove this. *Hint:* To show that invariance it is enough to check invariance under the transposition that interchanges x_1 and x_2, and the cyclic permutation $x_0 \mapsto x_1$, $x_1 \mapsto x_2$, and $x_1 \mapsto x_0$ as these to permutations generate the full permutation. □

There is anther useful formula for the second divided difference that relates it to convexity.

Proposition 13. *The second divided difference satisfies*

$$f[x_0,x_1,x_2] = \frac{1}{(x_2-x_1)(x_1-x_0)}\left(\frac{x_2-x_1}{x_2-x_0}f(x_0) - f(x_1) + \frac{x_1-x_0}{x_2-x_0}f(x_2)\right).$$

Moreover

$$\frac{x_2-x_1}{x_2-x_0} + \frac{x_1-x_0}{x_2-x_0} = 1$$

and

$$\frac{x_2-x_1}{x_2-x_0}x_0 + \frac{x_1-x_0}{x_2-x_0}x_2 = x_1.$$

Or to make it look even more like convexity set

$$\alpha = \frac{x_2-x_1}{x_2-x_0}, \qquad \beta = \frac{x_1-x_0}{x_2-x_0}$$

then

$$\frac{1}{(x_2-x_1)(x_1-x_0)}\left(\alpha f(x_0) - f(x_1) + \beta f(x_2)\right)$$

with

$$\alpha + \beta = 1, \qquad \alpha x_0 + \beta x_2 = x_1$$

and

$$x_0 < x_1 < x_2 \qquad \implies \qquad \alpha, \beta > 0.$$

Problem 14. Prove this. □

After this proposition the following should not be a surprise.

Proposition 14. *The function f on (a,b) is is convex if and only if $f[x_0,x_1,x_2] \geq 0$ for all distinct $x_0, x_1, x_2 \in (a,b)$. It is strictly convex if and only if $f[x_0,x_1,x_2] > 0$.*

Problem 15. Prove this. $\square$

Problem 16. Let $g(x) = ax^2 + bx + c$ be a polynomial of degree at most two. Show

$$g[x_0, x_1, x_2] = 2a = g''(x). \qquad \square$$

Proposition 15. *Let f be a twice differentiable function on (a, b) that vanishes at the three distinct points x_0, x_1, x_2. Then there is a point ξ between* $\min(x_0, x_1, x_2)$ *and* $\max(x_0, x_1, x_2)$ *with*

$$f''(\xi) = 0.$$

Problem 17. Prove this. *Hint:* Assume $x_0 < x_1 < x_2$. By Rolle's Theorem there is a ξ_0 between x_0 and x_1 with $f'(\xi_0) = 0$. Likewise there is ξ_1 between x_1 and x_2 with $f'(\xi_1) =$. Now apply Rolle's Theorem yet again, this time to the function f' on the interval (ξ_0, ξ_1). $\square$

Proposition 16. *Let f be twice differentiable on (a, b) and let x_0, x_1, x_2 be distinct points of this interval. There there is a point ξ between* $\min(x_0, x_1, x_2)$ *and* $\max(x_0, x_1, x_2)$ *with*

$$f[x_0, x_1, x_2] = f''(\xi).$$

Problem 18. Prove this. *Hint:* Let $g(x) = ax^2 + bx + c$ be the polynomial of degree at most with $g(x_j) = f(x_j)$ for $j = 0, 1, 2$. The $g''(x) = 2a$ is constant. Let $h = f - g$. Then $h(x_j) = 0$ for $j = 0, 1, 2$. Therefore

$$f[x_0, x_1, x_2] - g[x_0, x_1, x_2] = h[x_0, x_1, x_2] = 0.$$

But, Problem 16, $g[x_0, x_1, x_2] = g''(x) = 2a$. By Proposition 15 there is a ξ between $\min(x_0, x_1, x_2)$ and $\max(x_0, x_1, x_2)$ with $h''(\xi) = 0$. Then

$$f[x_0, x_1, x_2] - 2a = 0 = h''(\xi) = f''(\xi) - g''(\xi) = f''(\xi) = 2a. \qquad \square$$

Problem 19. (a) Show $f(x) = e^x$ is strictly convex on $\mathbb{R}$.

(b) Show that if $\alpha, \beta > 0$ and $\alpha + \beta = 1$ then for any $x, y \in \mathbb{R}$

$$(e^x)^\alpha (e^y)^\beta \le \alpha e^x + \beta e^y$$

and that inequality holds if and only if $x = y$.

(c) Prove the general arithmetic-geometric mean inequality: if $\alpha, \beta > 0$ and $\alpha + \beta = 1$ then for any positive numbers a, b

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

and equality holds if and only if $a = b$. (In the special case that $\alpha = \beta = 1/2$ this becomes usual arithmetic-geometric mean inequality $\sqrt{ab} \le \frac{a+b}{2}$.) *Hint:* Let $a = e^x$ and $b = e^y$. $\square$