CONVEXITY AND THE SIGN OF THE SECOND DERIVATIVE

1. Convexity

Definition 1. A function $f$ defined on an interval $I$ of real numbers is convex if, for any two points $(a, f(a))$ and $(b, f(b))$ lying on the graph of $f$, the secant line joining these two points lies above the graph of the function $y = f(x)$ for $a \leq x \leq b$.

The slope of a secant line joining the points $(a, f(a))$ and $(b, f(b))$ is $m = \frac{f(b) - f(a)}{b - a}$, and so the equation of the secant line joining these two points is

$$\frac{y - f(a)}{x - a} = m = \frac{f(b) - f(a)}{b - a}.$$ 

We can solve this equation for $y$. If $y = L_{a,b}(x)$ is the equation of this secant line, then

$$L_{a,b}(x) = f(a) + \left(\frac{x-a}{b-a}\right)(f(b) - f(a)).$$

Doing a little algebra, we see that

$$L_{a,b}(x) = f(a) + \left(\frac{x-a}{b-a}\right)(f(b) - f(a))$$

$$= f(a) + \left(\frac{x-a}{b-a}\right)f(b) - \left(\frac{x-a}{b-a}\right)f(a)$$

$$= f(a) - \left(\frac{x-a}{b-a}\right)f(a) + \left(\frac{x-a}{b-a}\right)f(b)$$

$$= \left[1 - \left(\frac{x-a}{b-a}\right)\right]f(a) + \left(\frac{x-a}{b-a}\right)f(b)$$

$$= \left(\frac{b-x}{b-a}\right)f(a) + \left(\frac{x-a}{b-a}\right)f(b).$$

(We can check that this formula is correct by noting that when $x = a$ we get $L_{a,b}(x) = f(a)$, and when $x = b$ we get $L_{a,b}(x) = f(b)$.)

It follows from Definition 1 that $f$ is convex on an interval $I$ if and only if for any two real numbers $a, b \in I$ with $a < b$, we have $f(x) \leq L_{a,b}(x)$, or equivalently $L_{a,b}(x) - f(x) \geq 0$, for $a \leq x \leq b$. Thus to show that $f$ is convex, we must show that

$$\left(\frac{b-x}{b-a}\right)f(a) + \left(\frac{x-a}{b-a}\right)f(b) - f(x) \geq 0 \quad \text{for} \quad a \leq x \leq b.$$ 

2. Non-negative second derivative implies convexity

We now use the Mean-Value Theorem (three times) to prove the following fact.

Theorem 1. Suppose that $f$ is a twice differentiable function defined on an interval $I$. If $f''(x) \geq 0$ for every $x$ in the interval, then the function $f$ is convex on this interval.

Proof. We do some more algebra as follows. Let $a < b$ be any two points in the interval $I$. First note that

$$1 = \left(\frac{x-a}{b-a}\right) + \left(\frac{b-x}{b-a}\right),$$

and so if we multiply both sides by $f(x)$ we get

$$f(x) = \left(\frac{x-a}{b-a}\right)f(x) + \left(\frac{b-x}{b-a}\right)f(x).$$

---

1Math 275 notes from the lecture on Monday, October 10, 2010.
It therefore follows that we can write the expression in the box above as

\[ L_{a,b}(x) - f(x) = \left( \frac{b-x}{b-a} \right) f(a) + \left( \frac{x-a}{b-a} \right) f(b) - f(x) \]

\[ = \left( \frac{b-x}{b-a} \right) f(a) + \left( \frac{x-a}{b-a} \right) f(b) - \left( \frac{x-a}{b-a} \right) f(x) - \left( \frac{b-x}{b-a} \right) f(x) \]

\[ = \left( \frac{b-x}{b-a} \right) \left[ f(a) - f(x) \right] + \left( \frac{x-a}{b-a} \right) \left[ f(b) - f(x) \right] \]

\[ = \left( \frac{x-a}{b-a} \right) \left[ f(b) - f(x) \right] - \left( \frac{b-x}{b-a} \right) \left[ f(x) - f(a) \right]. \]

(In the last line we replaced \([f(a) - f(x)]\) by \([f(x) - f(a)]\), and so we introduced a minus sign.) We can rewrite this last expression one more time my multiplying and dividing by \((b-x)\) and by \((x-a)\). We then get

\[ L_{a,b}(x) - f(x) = \left( \frac{x-a}{b-a} \right) \left[ f(b) - f(x) \right] - \left( \frac{b-x}{b-a} \right) \left[ f(x) - f(a) \right] \]

\[ = \frac{(x-a)(b-x)}{b-a} \left[ f(b) - f(x) \right] - \frac{(b-x)(x-a)}{b-a} \left[ f(x) - f(a) \right]. \]

Now suppose that \(a < x < b\). Then if we apply the Mean-Value Theorem to the function \(f\) on the interval \([x,b]\), we see that there is a number \(c_2\) with \(x < c_2 < b\) such that

\[ \left[ \frac{f(b)-f(x)}{b-x} \right] = f'(c_2). \]

Also, applying the Mean-Value Theorem to the function \(f\) on the interval \([a,x]\), we see that there is a number \(c_1\) with \(a < c_1 < x\) such that

\[ \left[ \frac{f(x)-f(a)}{x-a} \right] = f'(c_1). \]

Thus we can write

\[ L_{a,b}(x) - f(x) = \frac{(x-a)(b-x)}{b-a} f'(c_2) - \frac{(b-x)(x-a)}{b-a} f'(c_1) \]

\[ = \frac{(b-x)(x-a)}{b-a} \left[ f'(c_2) - f'(c_1) \right] \]

\[ = \frac{(b-x)(x-a)(c_2 - c_1)}{b-a} \left[ f''(c_2) - f''(c_1) \right]. \]

A final application of the Mean-Value Theorem, applied to the function \(f'\) on the interval \([c_1, c_2]\), shows that there is at least one point \(d\) with \(c_1 < d < c_2\) such that

\[ \frac{f'(c_2) - f'(c_1)}{c_2 - c_1} = f''(d). \]

Now we put all our calculations together. If \(a < x < b\), we have shown that there is a point \(d\) somewhere between \(a\) and \(b\) so that

\[ L_{a,b}(x) - f(x) = \frac{(b-x)(x-a)(c_2 - c_1)}{b-a} f''(d). \]

(1)

But when \(a < x < b\), all the factors on the right hand side of equation (1) are non-negative, and so \(L_{a,b}(x) - f(x) \geq 0\). This shows that \(f\) is convex, and completes the proof. \(\square\)

3. An extra credit exercise

We have shown that if the second derivative \(f''\) is non-negative, the function \(f\) is convex. The converse is also true. Try to prove this! That is, try to show that if \(f\) is twice differentiable and if \(L_{a,b}(x) - f(x) \geq 0\) for any two points \(a < b\) and any \(x\) such that \(a < x < b\), then \(f''(x) \geq 0\).

\textbf{HINT:} Try to fill in the following outline:

(a) Show that if \(a < x < b\), then \( \frac{f(x) - f(a)}{x-a} \leq \frac{f(b) - f(x)}{b-x} \).

(b) Show that if \(a < x_1 < x_2 < b\), then \( \frac{f(x_1) - f(a)}{x_1-a} \leq \frac{f(b) - f(x_2)}{b-x_2} \).

(c) Show that if \(a < b\) then \(f'(a) \leq f'(b)\).

(d) Show that \(f''(x)\) cannot be strictly negative.