Lecture 22 - 11/26 - Second derivatives.

Given $f : [a, b] \to \mathbb{R}$ of class C^1 on (a, b), we have seen that f' exists and is continuous on (a, b). If f' is in turn differentiable at $x_0 \in (a, b)$ (in the sense that $\lim_{h\to 0}(f'(x_0 + h) - f'(x_0))/h)$ exists, we say that f is **twice differentiable** at x_0 and denote $f''(x_0) = (f')'(x_0)$ (also called $\frac{d^2 f}{dx^2}(x_0)$, or second-order derivative of f at x_0). If f is twice differentiable at every point of (a, b) and $f''(x)$ is continuous on (a, b), we say that f is **twice continuously differentiable on** (a, b) (or, **of class** C^2 on (a, b)).

The second derivative f'' is useful for: (i) finding local minima or maxima; (ii) determining the concavity of the graph of f.

Theorem 85 (Local properties of f'')**.** *Suppose f is differentiable on a neighbordhood of x_0, and suppose $f''(x_0)$ exists. Define $g(x) := f(x_0) + f'(x_0)(x - x_0)$ (the best linear approximation to f near x_0).*

(a) Suppose $f'(x_0) = 0$. If $f''(x_0) > 0$ (resp. < 0), then x_0 is a strict local minimum (resp. maximum).

(b) If x_0 is a local minimum (resp. maximum), then $f''(x_0) \ge 0$ (resp. ≤ 0).

(c) If $f''(x_0) > 0$ (resp. < 0), there exists a neighborhood of x_0 where for all x, $f(x) \ge g(x)$ (resp. $\le g(x)$).

(d) If $f(x) \ge g(x)$ (resp. $\le g(x)$) for all x in a neighborhood of x_0, then $f''(x_0) \ge 0$ (resp. ≤ 0).

Proof. Proof of (a). If $f''(x_0) > 0$, then for x close enough to x_0, we write

$$f'(x) = f''(x_0)(x - x_0) + o(x - x_0),$$

so since $f''(x_0) > 0$, there exists $\delta > 0$ such that $f'(x) > 0$ for $x \in (x_0, x_0 + \delta)$ and $f'(x) < 0$ for $x \in (x_0 - \delta, x_0)$. In particular, f is strictly decreasing on the left of x_0 and strictly increasing on the right of x_0, thus f has a local minimum at x_0. The case $f''(x_0)$ is similar.

Proof of (b). By contradition, if x_0 is local minimum and $f''(x_0) < 0$, then by (a), x_0 is also a strict local maximum, which is impossible.

Proof of (c). Apply part (a) to the function $f(x) - g(x)$.

Proof of (d). Apply part (b) to the function $f(x) - g(x)$. □

Example 50.

1. *In (b), it is not necessary to have $f''(x_0) > 0$ at a strict local minimum: the function $f(x) = x^4$ has a strict local minimum at $x_0 = 0$, yet $f''(0) = 0$.*
2. *Moreover, when $f''(x_0) = 0$, we cannot tell a priori if the graph is above or below the tangent line near x_0, as this could be neither case, see for instance $f(x) = x^3$ at $x = 0$.*

Convexity/concavity. Another way to look at the second derivative is when describing the position of a graph with respect to its local chords (lines joining two points on the curve).

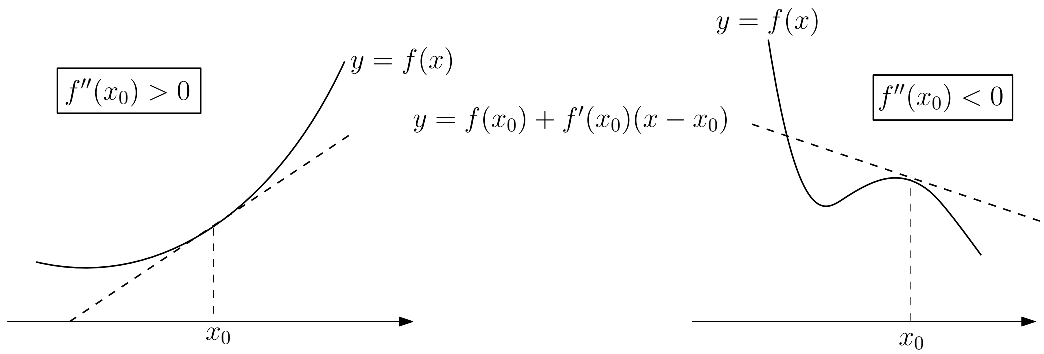

Figure 13: Local considerations when $f''(x_0) \neq 0$.

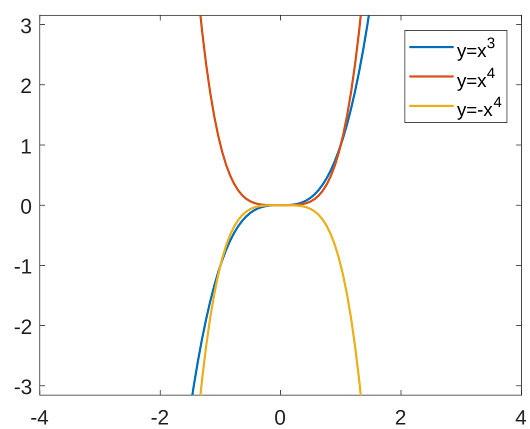

Figure 14: Local considerations when $f''(x_0) = 0$. Until we look at higher-order derivatives, no conclusion can be made regarding the relative position of f with respect to its tangent line near x_0. Here, all three functions satisfy $f(0) = f'(0) = f''(0) = 0$ so the tangent line at $x = 0$ is $y = 0$ but the relative positions can be anything near $x = 0$.

Theorem 86. *Suppose $f : [a,b] \to \mathbb{R}$ is of class C^2 on (a,b). Let $(x_1, x_2) \subset (a,b)$ and define*[17] $g(x) := f(x_1) + (x - x_1)\frac{f(x_2)-f(x_1)}{x_2-x_1}$.

(*i*) *If $f''(x) > 0$ for every $x \in (x_1, x_2)$, then $f(x) < g(x)$ for every $x \in (x_1, x_2)$.*

(*ii*) *If $f''(x) < 0$ for every $x \in (x_1, x_2)$, then $f(x) > g(x)$ for every $x \in (x_1, x_2)$.*

Proof. We only prove (i), as (ii) is similar. Set $h = f - g$. Note that $h(x_1) = h(x_2) = 0$ and $h''(x) = f''(x) > 0$ for every $x \in (x_1, x_2)$. Since h is continuous on $[x_1, x_2]$ so achieves its maximum there. If there is $x \in (x_1, x_2)$ such that $h(x) \geq 0$, then h has a local maximum x_0 inside (x_1, x_2), but from the previous theorem, $h''(x_0) \leq 0$, which contradicts $h'' > 0$. □

The past two theorems tell us that if $f''(x) > 0$ throughout an interval, then the graph of f lies above its tangents and below its chords. More generally, we call a function $f : [a,b] \to \mathbb{R}$ **convex** on $[a,b]$ if for every $x_1 < x_2$ in $[a,b]$, and $t \in (0,1)$,

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

Similarly, f is **concave** on $[a,b]$ if for every $x_1 < x_2$ in $[a,b]$, and $t \in (0,1)$,

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2). \tag{20}$$

One may define **strictly convex** and **strictly concave** by making the inequalities strict in the last two equations above.

Noticing that in the last two right-hand sides,

$$tf(x_1) + (1-t)f(x_2) = g(tx_1 + (1-t)x_2),$$

where $g(x) = f(x_1) + (x - x_1)\frac{f(x_2)-f(x_1)}{x_2-x_1}$, Theorem 86 is equivalent to saying that

(i) If $f''(x) > 0$ for every $x \in [a,b]$, then f is strictly convex on $[a,b]$.

(ii) If $f''(x) < 0$ for every $x \in [a,b]$, then f is strictly concave on $[a,b]$.

Example 51. *1. The function $f(x) = x^2$ is strictly convex on $\mathbb{R}$ since $f''(x) = 2$ for all x.*

2. The function $f(x) = \exp(x)$ is strictly convex on $\mathbb{R}$ since $\exp''(x) = \exp(x) > 0$ for all x.

Graphically, a function f is convex on $[a,b]$ if its **epigraph** $\text{epi}(f) = \{(x,y) : x \in [a,b], y \geq f(x)\}$ is a convex domain of $\mathbb{R}^2$ (in the sense that, if $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ belong to $\text{epi}(f)$, then the segment P_1P_2 belongs entirely to $\text{epi}(f)$).

[17] g is the unique affine function passing through $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Exercises for Lecture 22:

1. Where are the following functions convex/concave ?
   (a) $f(x) = (x^2+1)\exp(x)$, $x \in \mathbb{R}$.
   (b) $f(x) = \sqrt{x}$, $x \in (0,\infty)$.
2. Suppose that $f : (a,b) \to (c,d)$ is C^2 and invertible (in particular, $f'(x) \neq 0$ for every $x \in (a,b)$).
   (a) Express $(f^{-1})''(y)$ solely in terms of $f'(f^{-1}(y))$ and $f''(f^{-1}(y))$.
   (b) If f is strictly convex on (a,b), is it always true that f^{-1} is strictly concave on (c,d) ? Prove or disprove.
3. (a) Prove that $\ln(x)$ is increasing and concave on $(0,\infty)$.
   (b) Prove by induction on n using (a) that for any positive numbers $x_1, \dots x_n$,

$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{1}{n}(x_1 + \cdots + x_n).$$

   [Hint: this is equivalent to showing that $\frac{1}{n}(\ln(x_1) + \cdots + \ln(x_n)) \leq \ln\left(\frac{1}{n}(x_1 + \dots x_n)\right)$. Think about how to inductively use (20)]