MATH 1010E University Mathematics
Lecture Notes (week 7)
Martin Li

# 1 Optimization II: Extreme Value Theorem

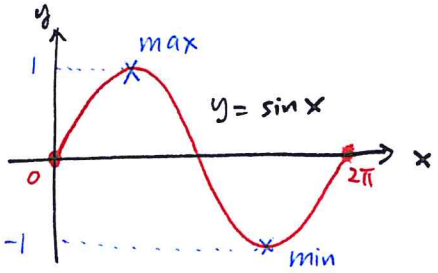
Recall from last lecture that we are interested in solving the following *optimization problem*:

$$\max/\min \qquad f = f(x) \qquad \text{where } x \in I$$

where $I$ could be a an open/closed and finite/infinite interval.

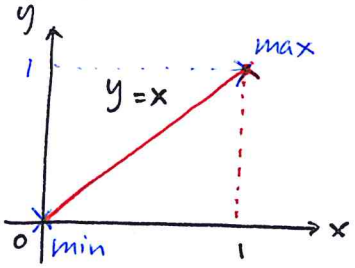
Let's look at some typical scenarios that could happen when we are looking for the maximum and minimum.

(1) Let $f(x) = \sin x$ and $I = [0, 2\pi]$. The maximum is 1 and the minimum



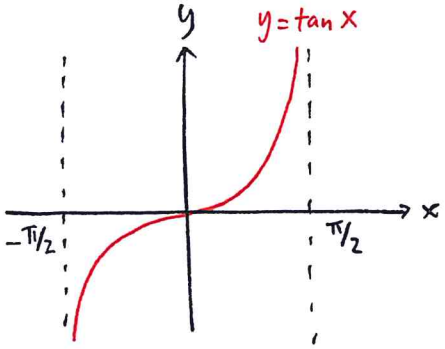

is $-1$. Both of them are achieved in the interior of $I$ and we have $f' = 0$ at these points.

(2) Let $f(x) = x$ and $I = [0, 1]$. The maximum is 1 and the minimum is



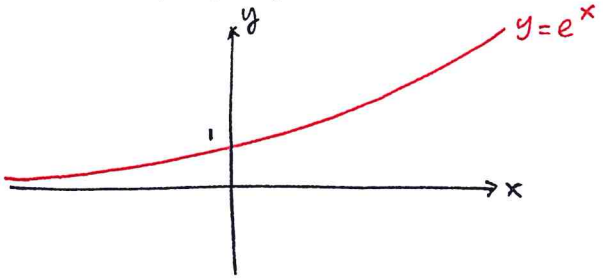

0. Both of them are achieved at the boundary points $x = 0$ and $x = 1$, where $f' \neq 0$ at these points.

(3) Let $f(x) = \tan x$ and $I = (-\pi/2, \pi/2)$ be an open interval. Neither the



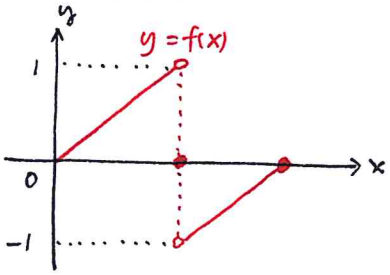

maximum nor minimum exist in this case because $\lim_{x\to\pm\pi/2}\tan x = \pm\infty$.

(4) Let $f(x) = e^x$ and $I = \mathbb{R} = (-\infty, +\infty)$ be the infinite interval. Neither




the maximum nor minimum exist in this case because $\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$, which is not achieved at any $x$.

(5) Let $f(x)$ be the following function defined on $I = [0, 2]$ as below: Neither the maximum nor minimum exist because the "maximum"




value 1 and "minimum" value i−1 s not achieved by any point. Note that $f$ is *not* continuous at $x = 1$.

So the general question about the optimization problem is that "when can we find the maximum and minimum?". This is answered by the following theorem.

**Theorem 1.1 (Extreme Value Theorem)** *Any continuous function $f : [a, b] \to \mathbb{R}$ on a finite, closed interval must achieve its minimum and maximum.*

The proof of the theorem is omitted since that uses the Bolzano-Weierstrass theorem about the the compactness of a closed finite interval $[a, b]$, which is beyond the scope of this class (interested students may consult any textbook on mathematical analysis).

The Extreme Value Theorem is the foundation of why the first derivative test holds. Let us restate this test.

**Theorem 1.2 (First order condition/First Derivative Test)** *If $f : [a, b] \to \mathbb{R}$ is continuous on the closed interval $[a, b]$ and differentiable on the open interval $(a, b)$, then if $x_0 \in (a, b)$ is an interior minimum or maximum point, then we have*

$$f'(x_0) = 0.$$

*Proof*: Let $x_0 \in (a, b)$ be an interior minimum point (the proof for maximum point is similar and left as an exercise), i.e.

$$f(x_0) \leq f(x) \qquad \text{for all } x \in [a, b]. \tag{1.1}$$

Our goal is to show that $f'(x_0) = 0$. From the assumption, we already know that $f'(x_0)$ exists. We just have to prove that its value is 0. Since $f'(x_0)$ exists, we have

$$\lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h}. \tag{1.2}$$

Now, for the right hand limit, we consider all $h > 0$ and

$$\lim_{h \to 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \geq 0$$

since the numerator $f(x_0 + h) - f(x_0) \geq 0$ by (1.1). Similarly, for the left hand limit, we consider all $h < 0$ and

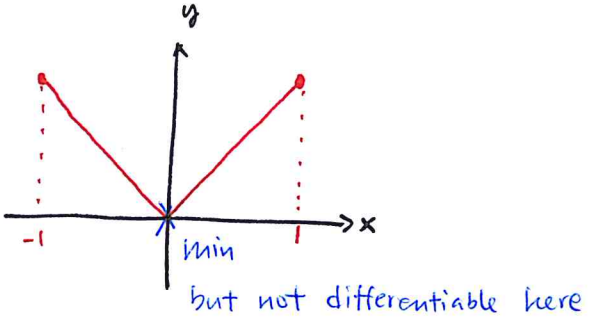
$$\lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h} \leq 0$$

since the numerator $f(x_0 + h) - f(x_0) \geq 0$ by (1.1). Combining these two inequalities with (1.2), we have

$$f'(x_0) \geq 0 \qquad \text{and} \qquad f'(x_0) \leq 0,$$

which can only be true when $f'(x_0) = 0$. This proves the first derivative test.

**Remark 1.3** *The first derivative test can only be applied if $f$ is* differentiable *and the minimum or maximum lies in the interior. For example, if $f(x) = |x|$ on $[-1, 1]$, then the minimum occurs at $x = 0$ where the function is not differentiable so we don't have $f'(0) = 0$. If the minimum or maximum occurs at an end point, e.g. see example (2) on the first page, then we do not have $f'(x_0) = 0$ either.*



## 2 Optimization III: Local min/max and 2nd Derivative Test

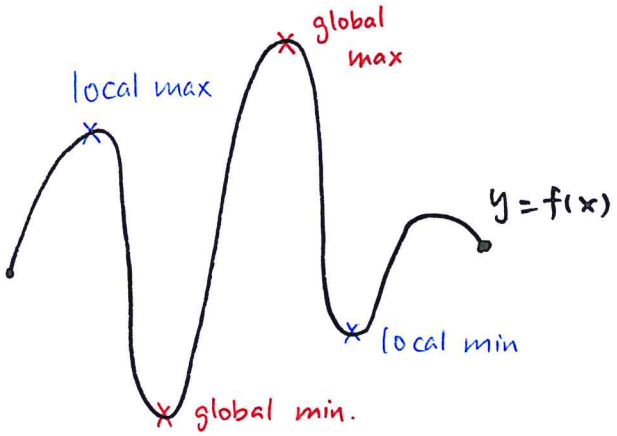
Before we talk about the second derivative test, we need to introduce the concept of local minimum and local maximum.

**Definition 2.1** *Let $f : [a, b] \to \mathbb{R}$ be a function. A point $x_0 \in [a, b]$ is said to be a* local minimum *if there exists $\epsilon > 0$ such that*

$$f(x_0) \leq f(x) \qquad \text{for all } x \in [a, b] \text{ such that } |x - x_0| < \epsilon.$$

*Similarly, we define* local maximum *by reversing the inequality to $f(x_0) \geq f(x)$.*

In other words, local minimum are minimum *compared only to nearby points.* A *global minimum* has to be a minimum compared to ALL the points, including points far away from $x_0$. The picture below illustrates this concept:




Now we can state the second derivative test.

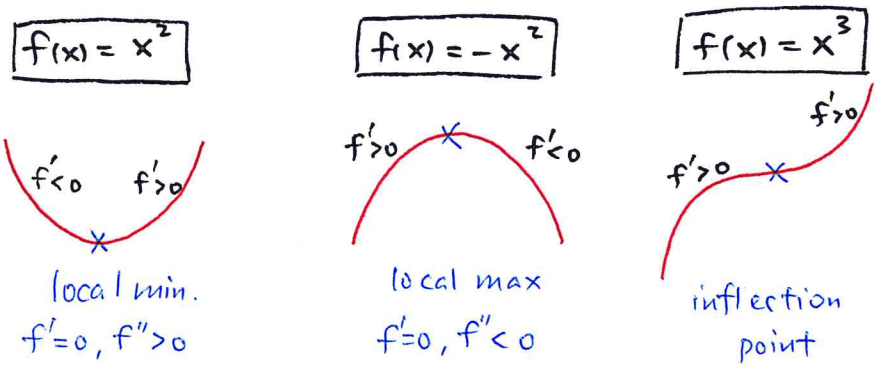
**Theorem 2.2 (Second Derivative Test)** *Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function and $x_0 \in (a, b)$. Suppose*

*(i) $f'(x_0) = 0$, i.e. $x_0$ is a critical point; and*

*(ii) $f''(x_0) > 0$ (respectively $f''(x_0) < 0$),*

*then $x_0$ is a local minimum (respectively local maximum).*

Note that if $f''(x_0) = 0$, then we cannot say anything about the critical point $x_0$. It is inconclusive from the second derivative test in this case. The following pictures show the model behavior for each of these cases:




The proof of the second derivative test needs more tools - *mean value theorems*, which will be discussed in the next section.

## 3 Mean Value Theorems

There are a couple versions of mean value theorems on the market. A prototype of mean value theorem is the following:

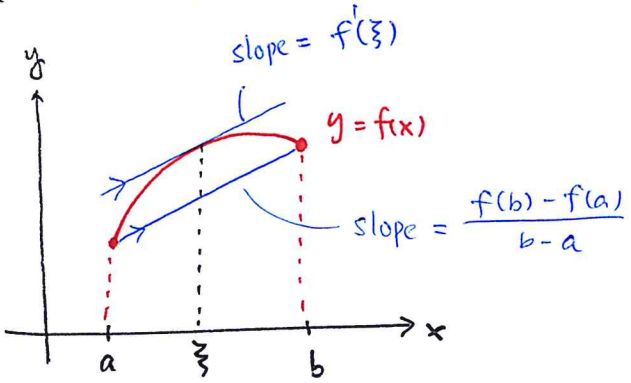
**Theorem 3.1 (Lagrange's Mean Value Theorem)** *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on $(a, b)$, then there exists some $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

**Remark 3.2** *Note that the conclusion of the theorem does not tell us where is this "$\xi$" located exactly. We just know that it lies somewhere strictly* between *$a$ and $b$. Notice that this fact is particularly useful when $a$ is very close to $b$, in which case $\xi$ is also very close to either $a$ or $b$. This fact will be used later when we discuss Taylor approximations.*

Geometrically, Lagrange's Mean Value Theorem says the following: suppose we draw a line connecting the end points of a graph (the blue line segment), if you move the blue line in a parallel manner, then it would *touch* the graph at some point again. The derivative of $f$ at that point is

equal to the slope of the blue line.




To prove Lagrange's Mean Value Theorem, we first look at a special case where $f(a) = f(b)$, which is another mean value theorem discovered by Rolle.

**Theorem 3.3 (Rolle's Mean Value Theorem)** *Under the same hypothesis as in Lagrange's Mean Value Theorem with the additional condition that $f(a) = f(b)$, then there exists $\xi \in (a, b)$ such that $f'(\xi) = 0$.*

*Proof*: The proof is just an application of the Extreme Value Theorem together with the first order condition. First of all, since $f$ is continuous on the closed finite interval $[a, b]$, the Extreme Value Theorem applies and both global minimum and maximum exist.

Let $m = \min f$ and $M = \max f$. We consider two possible cases:

Case 1: $m = M$. This implies that $f(x) \equiv m = M$ is a constant function since

$$m \leq f(x) \leq M \qquad \text{for all } x \in [a, b].$$

Hence, $f'(x) \equiv 0$ for all $x \in (a, b)$. Therefore, we can take any $\xi \in (a, b)$.

Case 2: $m \neq M$. We cannot have both minimum and maximum achieved at the end points since $f(a) = f(b)$ would imply $m = M$, which is just Case 1. Therefore, either the maximum or the minimum is achieved at some $\xi \in (a, b)$ in the interior. By the first order condition, we have $f'(\xi) = 0$. So we are done.

*Proof of Lagrange's Mean Value Theorem*: Define a function $\varphi : [a, b] \to \mathbb{R}$ by

$$\varphi(x) := [f(b) - f(a)]x - (b - a)f(x).$$

Note that $\varphi$ is continuous on $[a, b]$ and differentiable on $(a, b)$ (why?). Moreover,

$$\varphi(a) = af(b) - bf(a) = \varphi(b).$$

Therefore, we can apply Rolle's Mean Value Theorem to conclude that there exists $\xi \in (a,b)$ such that $\varphi'(\xi) = 0$. On the other hand, a direct calculation gives

$$\varphi'(x) = f(b) - f(a) - (b-a)f'(x),$$

and hence

$$\varphi'(\xi) = 0 \qquad \Leftrightarrow \qquad f'(\xi) = \frac{f(b)-f(a)}{b-a}.$$

This proves Lagrange's Mean Value Theorem.

**Question**: Use Lagrange's Mean Value Theorem to prove the following:

**Theorem 3.4 (Cauchy's Mean Value Theorem)** *If $f, g, : [a,b] \to \mathbb{R}$ are two functions which are continuous on $[a,b]$ and differentiable on $(a,b)$, and that $g'(x) \neq 0$ for all $x \in (a,b)$, then there exists $\xi \in (a,b)$ such that*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

**Remark 3.5** *If $g(x) = x$, then it is just Lagrange's Mean Value Theorem. Note that the right hand side above is always well defined, i.e. $g(a) \neq g(b)$. Otherwise, Rolle's Mean Value Theorem would imply the existence of some $\xi \in (a,b)$ such that $g'(\xi) = 0$ which contradicts the assumption on $g$.*

**Question**: Point out the flaw in the following proof of Cauchy's Mean Value Theorem:

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{\frac{f(b)-f(a)}{b-a}}{\frac{g(b)-g(a)}{b-a}} = \frac{f'(\xi)}{g'(\xi)}.$$

You can find the correct proof in the textbook "University Mathematics".

# 4 Application of Mean Value Theorem I: Taylor's Theorem

Consider a differentiable function $f : (a,b) \to \mathbb{R}$, and fix some $x_0 \in (a,b)$, applying Lagrange's Mean Value Theorem with $a = x_0$ and $b = x > x_0$, there exists some $\xi \in (x_0, x)$ such that

$$\frac{f(x)-f(x_0)}{x-x_0} = f'(\xi)$$

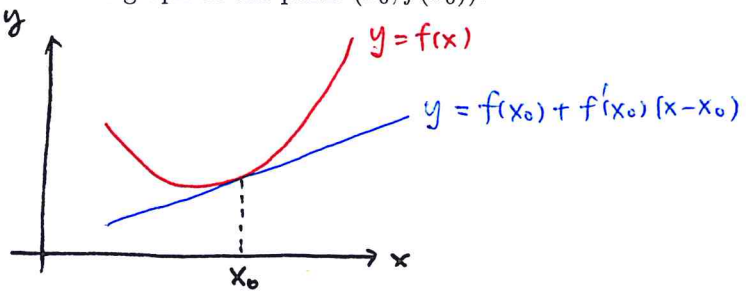
which we can rewrite as

$$f(x) = f(x_0) + f'(\xi)(x - x_0) \qquad \text{for all } x.$$

Note that the right hand side looks like a linear function in $x$ but it's actually not because $\xi$ actually depends (in some nonlinear way) on $x$ (and $x_0$ of course). However, when $x \approx x_0$, it is reasonable to have $f'(\xi) \approx f'(x_0)$ since $\xi \approx x_0$ as well (Caution: we need continuity of $f'$ at $x_0$ here), therefore

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) \qquad \text{for } x \approx x_0.$$

Now the right hand side is indeed a linear polynomial in $x$, but it would not be exactly the same as $f(x)$ and it's only a good approximation near the "center" $x_0$. The graph of the linear function on the right hand side actually is the tangent line to the graph at the point $(x_0, f(x_0))$.




Next, we can ask the question whether we can approximate $f(x)$ by a polynomial of higher degree (e.g. a quadratic polynomial) near $x_0$. In fact, the answer is YES, as long as $f(x)$ can be differentiated sufficiently many times. This is the famous Taylor's Theorem.

**Theorem 4.1 (Taylor's Theorem)** *Let $f : (a, b) \to \mathbb{R}$ be a function which has $(n+1)$-derivatives and fix some $x_0 \in (a, b)$, then for all $x \in (a, b)$, there exists some $\xi$ (depending on $x$) lying strictly between $x_0$ and $x$ such that*

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\ &\quad + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}. \end{aligned}$$

*We call the first part* Taylor polynomial of $f$ of degree $n$ at $x = x_0$

$$TP_n(x) := f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

*and the second part the* error term

$$E_n(x) := \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

**Remark 4.2** *Note that the theorem holds for ALL $x \in (a,b)$, not just for those $x$ close to $x_0$. However, if we want the error term $E_n(x)$ to be small, then we require $x \approx x_0$ to make $(x-x_0)^{n+1}$ in the error term small.*

When $n = 0$, Taylor's theorem says

$$f(x) = f(x_0) + f'(\xi)(x-x_0),$$

which is just Lagrange's Mean Value Theorem.

When $n = 1$, Taylor's theorem says

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(\xi)}{2!}(x-x_0)^2,$$

therefore, for $x \approx x_0$,

$$f(x) \approx f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2,$$

which is a quadratic polynomial in $x$ approximating $f(x)$ near $x_0$.

**Question**: Check that $TP_n(x)$ approximates $f(x)$ at $x = x_0$ up to order $n$, i.e. show that

$$TP_n(x_0) = f(x_0), \quad (TP_n)'(x_0) = f'(x_0), \quad \cdots \quad (TP_n)^{(n)}(x_0) = f^{(n)}(x_0).$$

**Example 4.3** *Find $TP_3(x)$ about $x = 0$ for the function $f(x) = \frac{1}{1+x}$.*

**Solution**: Recall that

$$TP_3(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3.$$

Computing the successive derivatives, we get

$$f(x) = \frac{1}{1+x}, \quad f'(x) = \frac{-1}{(1+x)^2}, \quad f''(x) = \frac{2}{(1+x)^3}, \quad f'''(x) = \frac{-6}{(1+x)^4}.$$

Substituting $x = 0$, we obtain

$$f(0) = 1, \quad f'(0) = -1, \quad f''(0) = 2, \quad f'''(0) = -6.$$

Therefore, the Taylor polynomial of $f$ of degree 3 at $x = 0$ is

$$TP_3(x) = 1 - x + x^2 - x^3.$$

**Example 4.4** *Find $TP_3(x)$ about $x = 1$ for the function $f(x) = \frac{1}{1+x}$.*

**Solution**: This time, we have

$$TP_3(x) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3.$$

Using the calculation of derivatives above, we find that the Taylor polynomial of $f$ of degree 3 at $x = 1$ is

$$TP_3(x) = \frac{1}{2} - \frac{1}{4}(x-1) + \frac{1}{8}(x-1)^2 - \frac{1}{16}(x-1)^3.$$

Taylor's Theorem can be used to prove some inequalities.

**Example 4.5** *Use Taylor's Theorem to prove that*

$$e^x > 1 + x + \frac{x^2}{2} \qquad \text{for all } x > 0.$$

**Solution**: By Taylor's Theorem for $f(x) = e^x$ at $x = 0$ with $n = 1$, since $f^{(n)}(x) = e^x$ for $n = 0, 1, 2, 3, \ldots$, we have

$$e^x = 1 + x + \frac{e^\xi}{2}x^2,$$

for some $\xi \in (0, x)$ (note that $x > 0$). Now, since $\xi > 0$, we have $e^\xi > 1$, this implies that

$$e^x = 1 + x + \frac{e^\xi}{2}x^2 > 1 + x + \frac{1}{2}x^2$$

which is the inequality we want to prove.

We can also use the Taylor's theorem to prove the second derivative test. Remember the second order Taylor approximation tells us that

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \qquad \text{for } x \approx x_0.$$

If we have $f'(x_0) = 0$ and $f''(x_0) > 0$, then

$$f(x) \approx f(x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \geq f(x_0) \qquad \text{for } x \approx x_0.$$

This hints that $x_0$ is a local minimum. However, this is not a rigorous proof.

**Question**: Make the argument rigorous if we assume that $f$ is twice differentiable AND $f''$ is continuous at $x_0$.

The original second derivative test holds even without the extra assumption on the continuity of $f''$. One has to be more careful in the proof though (see textbook "University Mathematics" for a complete proof in this case).