Generalized local test for local extrema in single-variable functions

Eleftherios Gkioulekas

University of Texas Rio Grande Valley

April 14, 2018

Introduction

- The challenge: *generalize the second derivative test for classifying critical points of single variable functions as local minima or maxima in situations where both the first derivatives are zero.*
- Details are published in
  - E. Gkioulekas: "Generalized local test for local extrema in single-variable functions", *International Journal of Mathematical Education in Science and Technology* **45** (2014), 118-131
- Another generalization of the second derivative test
  - Y. Wu: "Improved second derivative test for relative extrema", *International Journal of Mathematical Education in Science and Technology* **38** (2007), 1121-1123.
- Second derivative test for multi-variable calculus requires solving an eigenvalue problem.
  - T. Apostol: "Calculus. II", John Wiley & Sons, New York NY, 1969.
- Second derivative test for constrained optimization problems.
  - M. Nerenberg: "The second derivative test for constrained extremum problems", *International Journal of Mathematical Education in Science and Technology* **22** (1991), 303-308.

Motivation

- The *second derivative test* can be used to classify most critical points that are candidates for local minimum or local maximum with zero first derivative.
- Disadvantages of the second derivative test:
  - It cannot classify critical points where the function loses differentiability or points that are not interior to the domain of the function that is being optimized. According to the Fermat theorem, these are also potential candidates for local minima and maxima.
  - Finding the second derivative can sometimes require an unreasonable amount of effort.
  - After long calculations to find the second derivative, one might discover that for some critical point the second derivative test is inconclusive.
- These disadvantages are why it is better to use a table of signs to determine the sign of the first derivative over all relevant intervals, and rely instead on the *first derivative test*
- Advantages of the second derivative test:
  - *The second derivative test can be especially advantageous for* ***trigonometric functions*** *where, due to periodicity, applying the first derivative test is too awkward.*
  - It can also be used in more advanced coursework in the context of proving theorems that are dependant on the second derivative test. (classification of local bifurcations)
  - Difficult to generalize first derivative test to multivariable calculus.
- Can higher derivatives be used locally to resolve the inconclusive case of the second derivative test?
- Answer: **yes** for *single-variable* functions. Otherwise: open question.

Preliminaries: Generalized monotonicity theorems. I

- Let $f : A \to \mathbb{R}$ be a real-valued function with domain $A \subseteq \mathbb{R}$ and let $I \subseteq A$. We use the notation $f \uparrow A$ to state that f is strictly increasing on I, and $f \downarrow I$ to state that f is strictly decreasing on I.
- The corresponding definitions are given by:

$$f \uparrow I \iff \forall x_1, x_2 \in I : (x_1 < x_2 \implies f(x_1) < f(x_2)),$$
$$f \downarrow I \iff \forall x_1, x_2 \in I : (x_1 < x_2 \implies f(x_1) > f(x_2)).$$

- Derivative determines monotonicity; sharper results

$$\left.\begin{array}{l} \forall x \in (a,b) : f'(x) > 0 \\ f \text{ is continuous on } x = b \end{array}\right\} \implies f \uparrow (a,b],$$
$$\left.\begin{array}{l} \forall x \in (a,b) : f'(x) > 0 \\ f \text{ is continuous on } x = a \end{array}\right\} \implies f \uparrow [a,b),$$
$$\left.\begin{array}{l} \forall x \in (a,b) : f'(x) < 0 \\ f \text{ is continuous on } x = b \end{array}\right\} \implies f \downarrow (a,b],$$
$$\left.\begin{array}{l} \forall x \in (a,b) : f'(x) < 0 \\ f \text{ is continuous on } x = a \end{array}\right\} \implies f \downarrow [a,b).$$

Preliminaries: Generalized monotonicity theorems. II

- *Differentiability and a non-zero derivative are required only for the interior of isolated intervals, whereas at the boundary points of the intervals, all that is required is continuity. Given that we get monotonicity over a set including the boundary points where the function is continuous*
- The proof for these sharper theorems is the standard proof via the mean-value theorem.
- The key is to realize that the mean-value theorem itself requires differentiability only at the interior of the interval, with continuity being sufficient at the interval endpoints.
- After applying the mean-value theorem, the resulting first derivative will always be evaluated on the interior of the interval, therefore it is only necessary to know the sign of the first derivative at the open interval (a, b) for all cases.
- The above results are needed to establish a strong version of the first derivative test, and furthermore in the proof of the higher derivative generalizations of the second derivative tests.

Preliminaries: "We get there eventually" theorem

- We also require the following results from the theory of limits:

$$\lim_{x \to \sigma} f(x) > 0 \implies \exists \delta > 0 : \forall x \in N(\sigma, \delta) \cap A : f(x) > 0, \tag{1}$$

$$\lim_{x \to \sigma} f(x) < 0 \implies \exists \delta > 0 : \forall x \in N(\sigma, \delta) \cap A : f(x) < 0. \tag{2}$$

- Here σ is a generalized accumulation point of the domain $A = \mathrm{dom}(f)$ of the function f (i.e. $\sigma = x_0 \in \mathbb{R}$, or $\sigma = x_0^+$, or $\sigma = x_0^-$, or $\sigma = +\infty$, or $\sigma = -\infty$), and $N(\sigma, \delta)$ is a generalized neighborhood of σ defined as:

$$N(\sigma, \delta) = \begin{cases} (x_0 - \delta, x_0 + \delta) - \{x_0\} & , \text{ if } \sigma = x_0 \\ (x_0 - \delta, x_0) & , \text{ if } \sigma = x_0^- \\ (x_0, x_0 + \delta) & , \text{ if } \sigma = x_0^+ \\ (1/\delta, +\infty) & , \text{ if } \sigma = +\infty \\ (-\infty, -1/\delta) & , \text{ if } \sigma = -\infty \end{cases} . \tag{3}$$

- It is worth noting that the converse of this result cannot be proven without weakening the inequalities into weak inequalities.
- The converse statements do not become relevant to the problem at hand.

The first derivative test

- **Theorem 1**
  Let $f : A \to \mathbb{R}$ be a function and let $x_0 \in A$ be a point in the interior of the domain A of f such that it satisfies all of the following conditions:
  1. *f differentiable on $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$*
  2. *f continuous on x_0*

  Then, it follows that

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0 \end{cases} \Longrightarrow x_0 \text{ is local minimum of } f,$$

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0 \end{cases} \Longrightarrow x_0 \text{ is local maximum of } f,$$

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0 \end{cases} \Longrightarrow x_0 \text{ is not extremum of } f,$$

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0 \end{cases} \Longrightarrow x_0 \text{ is not extremum of } f.$$

  - Mainstream Calculus texts are not careful in their formulation of the first derivative test.
  - On the other hand, the formulation of the first derivative test proved by Apostol is equivalent to the statement given above.

Now: A chicken crosses the road!!

- ## Lemma 2

Let $f : A \to \mathbb{R}$ be a function with domain $A \subseteq \mathbb{R}$ and let $x_0 \in A$ be a point interior to A such that it satisfies all of the following conditions:

1. *f is 3 times differentiable on $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$*
2. *f is 2 times differentiable on x_0*
3. *f'' continuous on x_0*
4. *$f'(x_0) = f''(x_0) = 0$*

Then, the statements below follow:

$$\left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0 \end{array} \right. \implies \left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0, \end{array} \right.$$

$$\left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'''(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) < 0 \end{array} \right. \implies \left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0, \end{array} \right.$$

$$\left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'''(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0 \end{array} \right. \implies \left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0, \end{array} \right.$$

$$\left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) < 0 \end{array} \right. \implies \left\{ \begin{array}{l} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0. \end{array} \right.$$

Proof: Why did the chicken cross the road? Part I

Proof.

- Assume that:

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0. \end{cases}$$

- Let $x \in (x_0 - \delta, x_0)$ be given. Apply the mean-value theorem to f' on the interval $[x, x_0]$:

$$\exists x_1 \in (x, x_0) : f''(x_1) = \frac{f'(x) - f'(x_0)}{x - x_0}. \tag{4}$$

- Apply the mean-value theorem to f'' on the interval $[x_1, x_0]$:

$$\exists x_2 \in (x_1, x_0) : f'''(x_2) = \frac{f''(x_1) - f''(x_0)}{x_1 - x_0}. \tag{5}$$

- Combining the above two equations we find that

$$\begin{aligned} f'(x) &= f'(x) - f'(x_0) && \text{[because } f'(x_0) = 0\text{]} \\ &= f''(x_1)(x - x_0) && \text{[via Eq. (4)]} \\ &= [f''(x_1) - f''(x_0)](x - x_0) && \text{[because } f''(x_0) = 0\text{]} \\ &= f'''(x_2)(x_1 - x_0)(x - x_0) < 0 && \text{[via Eq. (5)]} \end{aligned}$$

  because $f'''(x_2) < 0$ and $x_1 - x_0 < 0$ and $x - x_0 < 0$. We conclude that $\forall x \in (x_0 - \delta, x_0) : f'(x) < 0$.

Proof: Why did the chicken cross the road? Part II

- Let $x \in (x_0, x_0 + \delta)$ be given. Apply the mean-value theorem on f' at the interval $[x_0, x]$:
$$\exists x_1 \in (x_0, x) : f''(x_1) = \frac{f'(x) - f'(x_0)}{x - x_0}.$$
- Apply the mean-value theorem on f'' at the interval $[x_0, x_1]$:
$$\exists x_2 \in (x_0, x_1) : f'''(x_2) = \frac{f''(x_1) - f''(x_0)}{x_1 - x_0}.$$
- From the above equations, it follows again that
$$f'(x) = f'''(x_2)(x_1 - x_0)(x - x_0) > 0,$$
because now we have $f'''(x_2) > 0$ and $x_1 - x_0 > 0$ and $x - x_0 > 0$, instead. We conclude that $\forall x \in (x_0, x_0 + \delta) : f'(x) > 0$.
- Similar proofs can be given for the other three statements.

□

Odd-order derivative test

- Notation: $[n] = \{k \in \mathbb{N} \mid 1 \leq k \leq n\} = \{1, 2, 3, \ldots, n\}$

- ## Lemma 3

  Let $f : A \to \mathbb{R}$ be a function, let $x_0 \in A$ be an interior point, and let $n \in \mathbb{N} - \{0\}$ such that all of the following conditions are satisfied:

  1. *f is $2n+1$ times differentiable on $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$*
  2. *f is $2n$ times differentiable on x_0*
  3. *$f^{(2n)}$ continuous on x_0*

  Then it follows that:

$$\left.\begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) > 0 \end{array}\right\} \implies x_0 \text{ is local minimum of } f,$$

$$\left.\begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) < 0 \end{array}\right\} \implies x_0 \text{ is local maximum of } f,$$

$$\left.\begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) > 0 \end{array}\right\} \implies x_0 \text{ is not extremum of } f,$$

$$\left.\begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) < 0 \end{array}\right\} \implies x_0 \text{ is not extremum of } f.$$

Generalization of the 2nd derivative test. I

- The overall idea is that if it should happen that at a critical point $x_0 \in A$, in the domain A of some function f, we have $f'(x_0) = f''(x_0) = 0$, we continue checking higher-order derivatives such as $f'''(x_0), \ldots, f^{(n)}(x_0)$ until we encounter a non-zero derivative.
- Case 1: The first nonzero derivative is **even**

- Theorem 4

Let $f : A \to \mathbb{R}$ be a function, let $x_0 \in A$ be an interior point, and let $n \in \mathbb{N} - \{0\}$. We assume that all of the following statements are satisfied:

1. *f is $2n - 1$ times differentiable on $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$*
2. *f is $2n$ times differentiable on x_0*

Then, it follows that

$$\left.\begin{array}{l} \forall \alpha \in [2n-1] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n)}(x_0) > 0 \end{array}\right\} \implies x_0 \text{ is local minimum of } f,$$

$$\left.\begin{array}{l} \forall \alpha \in [2n-1] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n)}(x_0) < 0 \end{array}\right\} \implies x_0 \text{ is local maximum of } f.$$

Generalization of the 2nd derivative test. II

Proof.

- We begin with establishing the first statement. Assume that $\forall \alpha \in [2n-1] : f^{(\alpha)}(x_0) = 0$ and $f^{(2n)}(x_0) > 0$. First, we note that

$$
\begin{aligned}
f^{(2n)}(x_0) &= \lim_{x \to x_0} \frac{f^{(2n-1)}(x) - f^{(2n-1)}(x_0)}{x - x_0} && \text{[by definition]} \\
&= \lim_{x \to x_0} \frac{f^{(2n-1)}(x)}{x - x_0} > 0 && \text{[via } f^{(2n-1)}(x_0) = 0\text{]} \\
&\implies \exists \delta > 0 : \forall x \in N(x_0, \delta) : \frac{f^{(2n-1)}(x)}{x - x_0} > 0 \\
&\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n-1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n-1)}(x) > 0. \end{cases}
\end{aligned}
$$

- The first implication uses the "We get there eventually" theorem, where the differentiability assumptions given by the theorem allows us to replace $N(x_0, \delta) \cap A$ with $N(x_0, \delta)$.
- Combining the above result with the assumption $\forall \alpha \in [2n-1] : f^{(\alpha)}(x_0) = 0$ via the Lemma 3, we conclude that x_0 is a local minimum.
- The next statement can be derived via a similar argument.

$\square$

Generalization of the 2nd derivative test. III

- Case 2: The first nonzero derivative is **odd**

- ## Theorem 5

  Let $f : A \to \mathbb{R}$ be a function, let $x_0 \in A$ be an interior point, and let $n \in \mathbb{N} - \{0\}$. We assume that all of the following statements are satisfied:

  1. *f is $2n$ times differentiable on $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$*
  2. *f is $2n + 1$ times differentiable on x_0*

  Then it follows that

$$\left.\begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n+1)}(x_0) \neq 0 \end{array}\right\} \Longrightarrow x_0 \text{ is not extremum of } f.$$

Generalization of the 2nd derivative test. IV

Proof.

- **Proof:** Let us assume with no loss of generality that $f^{(2n+1)}(x_0) > 0$. Then it follows that

$$
\begin{aligned}
f^{(2n+1)}(x_0) &= \lim_{x \to x_0} \frac{f^{(2n)}(x) - f^{(2n)}(x_0)}{x - x_0} && \text{[by definition]} \\
&= \lim_{x \to x_0} \frac{f^{(2n)}(x)}{x - x_0} > 0 && [\text{via } f^{(2n)}(x_0) = 0] \\
&\Longrightarrow \exists \delta > 0 : \forall x \in N(x_0, \delta) : \frac{f^{(2n)}(x)}{x - x_0} > 0 \\
&\Longrightarrow \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n)}(x) > 0 \end{cases} \\
&\Longrightarrow \begin{cases} f^{(2n-1)} \downarrow (x_0 - \delta, x_0] \\ f^{(2n-1)} \uparrow [x_0, x_0 + \delta) \end{cases} \\
&\Longrightarrow \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n-1)}(x) > f^{(2n-1)}(x_0) = 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n-1)}(x) > f^{(2n-1)}(x_0) = 0. \end{cases}
\end{aligned}
$$

- The third implication uses the *generalized monotonicity theorems*. This transition to closed intervals, on the x_0 side, is in turn necessary to complete the fourth implication.
- Combining the above result with the assumption $\forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0$, we conclude via Lemma 3 that x_0 is neither a local minimum nor a local maximum. $\square$

A trigonometric example. I

- Polynomial functions do not make for interesting application problems for the second derivative test and the generalizations given by Theorem 4 and Theorem 5.
- Trigonometric functions, on the other hand, lead to an infinite set of critical points, and using the first derivative test on such problems tends to be cumbersome.
- Simple trigonometric functions can be usually handled via the second derivative test with relative efficiency, but it is possible to define trigonometric functions where the second derivative test fails.
- A unique challenge posed by trigonometric problems is that some set theory is needed to organize and keep track of the critical points.

A trigonometric example. II

Example 6

Find all local minima and maxima of the function defined by $f(x) = \sin(4x)\cos^4(x)$.

Final Answer.

- Critical points are organized in the following disjoint sets:

$$\forall \alpha \in [5] : S_{2,\alpha} = \left\{ \frac{\pi}{10} + \frac{(5k+\alpha)\pi}{5} \;\middle|\; k \in \mathbb{Z} \right\}.$$

- Using the second derivative test, we can show that:

$$\begin{aligned}
&\text{For } a = 1 : \quad \forall x \in S_{2,1} : x \text{ local minimum;}\\
&\text{For } a = 3 : \quad \forall x \in S_{2,3} : x \text{ local maximum;}\\
&\text{For } a = 4 : \quad \forall x \in S_{2,4} : x \text{ local minimum;}\\
&\text{For } a = 5 : \quad \forall x \in S_{2,5} : x \text{ local maximum.}
\end{aligned}$$

- On the other hand, for the case $\alpha = 2$, we find

$$\forall x \in S_{2,2} : \begin{cases} \forall \alpha \in [4] : f^{(\alpha)}(x) = 0 \\ f^{(5)}(x) \neq 0 \end{cases} \Longrightarrow \forall x \in S_{2,2} : x \text{ not a local extremum.} \square$$

Conclusion

- The main advantage of the generalization of the second derivative test given by Theorem 4 and Theorem 5 is that they are local results where all relevant derivatives need only be evaluated at the critical points.
- The theorems are robust, since any non-constant analytic function will eventually yield a non-zero high-order derivative at each critical point.
- Construction of a decent exercise set leveraging these theorems would be worth exploring

Thank you!