Caltech Division of the Humanities and Social Sciences

Calculus and Maximization I

KC Border

Fall 2000
v. 2018.10.04::10.03

1 Maxima and Minima

A function $f\colon X \to \boldsymbol{R}$ attains a **global maximum** or **absolute maximum** over X at $x^* \in X$ if

$$f(x^*) \geqslant f(x) \qquad \text{for all } x \in X.$$

We may also say that x^* is a **global maximizer** or **global maximum point** of f on X, or that x^* **maximizes** f over X. The value $f(x^*)$ is called the **global maximum** of f on (or over) X.[1]

The function f achieves a **minimum** at x^* if $f(x^*) \leqslant f(x)$ for every $x \in X$. An **extremum** of f is a point where f attains either a maximum or a minimum. Notice that f achieves a maximum at x^* if and only if $-f$ achieves a minimum there. Consequently, anything we say about maxima can be converted into a statement about minima by appropriately changing signs. In particular, all the definitions listed here regarding maxima have corresponding definitions for minima.

A function f has a **strict maximum** on X at $x^* \in X$ if

$$f(x^*) > f(x) \qquad \text{for all } x \in X \text{ satisfying } x \neq x^*.$$

We may sometimes use the term **weak maximum** instead of maximum to emphasize that we do not necessarily mean a strict maximum.

When $X \subset \boldsymbol{R}^{\mathrm{n}}$, we say that f has a **local maximum** or **relative maximum** at x^* or that x^* is a **local maximizer** of f on X if there is a neighborhood U of x^* in X such that x^* maximizes f on U. That is,

$$\text{there exists } \varepsilon > 0 \text{ such that } f(x^*) \geqslant f(x) \text{ whenever } x \in X \text{ and } \|x - x^*\| < \varepsilon.$$

Likewise a **strict local maximizer** satisfies

$$\text{there exists } \varepsilon > 0 \text{ such that } f(x^*) > f(x) \text{ whenever } x \in X \text{ and } 0 < \|x - x^*\| < \varepsilon.$$

[1] I used to refer to the point x^* as a maximum of f. Roko Aliprantis has convinced me that this is confusing at best and wrong at worst. For instance, what is the maximum of the cosine function, 1 or 0? Most people will answer 1, but this value is attained at 0.

I shall endeavor to avoid this perhaps erroneous practice, but I may backslide. Incidentally, I am not alone. A quick random sample of my bookshelves reveals that Luenberger [10] and Varian [17] also refer to x^* as a maximum of f, while Apostol [3], Debreu [4], and Mas-Colell, Whinston, and Green [12] do not. Some authors equivocate, e.g., Marsden [11]. The venerable Hancock [5] says that "f is a maximum for x^*."

Often we are interested in maxima and minima of f on a subset of its domain. A common way to define such a subset is in terms of inequality constraints of the form $g_j(x) \geqslant \alpha_j$, where $g_j : X \to \boldsymbol{R}$, $j = 1, \dots, m$. (Or, we may write $g(x) \geqq a$, where $g = (g_1, \dots, g_m) : X \to \boldsymbol{R}^{\mathrm{m}}$, and $a = (\alpha_1, \dots, \alpha_m)$.) We say that a point x **satisfies the constraints** if it belongs to X and $g(x) \geqq a$. The set of points satisfying the constraints, $[g \geqq a] = \{x \in X : g(x) \geqq a\}$, is called the **constraint set**. The function f itself may be referred to as the **objective function** to distinguish it from the **constraint functions** $g_1, \dots, g_m$.

We say that x^* is a **constrained maximizer** of f if x^* satisfies the constraints, $g(x^*) \geqq a$, and $f(x^*) \geqslant f(x)$ for every x satisfying the constraints. The point x is a **local constrained maximizer** if it satisfies the constraints and there is a neighborhood U of x such that $f(x^*) \geqslant f(x)$ for every $x \in U$ satisfying the constraints. The point x^* is an **interior** maximizer of f if its lies in the relative interior of the constraint set.

2 Calculus Review

We start by recalling that the notation

$$\lim_{x \to a} f(x) = c$$

where c is a real number, means that for every $\varepsilon > 0$, there is some $\delta > 0$ such that if $0 < |x - a| < \delta$, then $|f(x) - c| < \varepsilon$. The reason we restrict attention to x with $0 < |x - a|$ is so that we can divide by $x - a$, as in the next definition. We may also write

$$\lim_{x \to a} f(x) = \infty$$

if every $M \in \boldsymbol{R}$, there is some $\delta > 0$ such that if $0 < |x - a| < \delta$, then $f(x) > M$. Similarly

$$\lim_{x \to a} f(x) = -\infty$$

if for every $M \in \boldsymbol{R}$, there is some $\delta > 0$ such that if $0 < |x - a| < \delta$, then $f(x) < M$.

1 Definition (Derivative of a function) *Let $f : (a, b) \to \boldsymbol{R}$ be a real function of one real variable. If the limit*

$$\lim_{v \to 0} \frac{f(x + v) - f(x)}{v}$$

*exists (as a finite real number), then we say that f is **differentiable** at x and that the limit is the **derivative** of f at x, denoted $f'(x)$ or $Df(x)$.*

If $\lim_{v \to 0} \frac{f(x+v) - f(x)}{v} = \pm\infty$ we may on occasion say that $f'(x) = \pm\infty$, but we shall not say that f is differentiable in this case.

*If the function f' has a derivative at x, it is called the **second derivative of f at x** and is denoted $f''(x)$ or $D^2 f(x)$. Proceeding inductively, if the $n-1$st derivative of f has a derivative at x it is called the **nth derivative** of f at x and is denoted $f^{(n)}(x)$ (note the parentheses around the n) or $D^n f(x)$.*

Finally, if $\lim_{v\uparrow 0} \frac{f(x+v)-f(x)}{v}$ *exists, the limit is the* ***left-hand (one-sided) derivative of f at x*** *and if* $\lim_{v\downarrow 0} \frac{f(x+v)-f(x)}{v}$ *exists, the limit is the* ***right-hand (one-sided) derivative of f at x***.[2]

According to this definition, a function has a derivative only at interior points of its domain, and the derivative is always finite.

The next result is immediate from the definition, but is worth giving a name.

2 Squeezing Lemma *Suppose* $f \geqslant h \geqslant g$ *everywhere and* $f(x) = h(x) = g(x)$. *If* f *and* g *are differentiable at* x, *then* h *is also differentiable at* x, *and* $f'(x) = h'(x) = g'(x)$.

Next we present the well known Mean Value Theorem, see e.g., Apostol [3, Theorem 4.5, p. 185], and an easy corollary, cf. [3, Theorems 4.6 and 4.7].

3 Mean Value Theorem *Suppose* f *is continuous on* $[a, b]$ *and differentiable on* (a, b). *Then there exists* $c \in (a, b)$ *satisfying*

$$f(b) - f(a) = f'(c)(b - a).$$

This result has some corollaries relating derivatives and monotonicity of functions.

4 Definition (Monotonicity) *Let* $f\colon [a, b] \to \boldsymbol{R}$. *We say that* f *is*

strictly increasing *on* (a, b) *if* $a < x < y < b$ *implies* $f(x) < f(y)$.

increasing *or* ***nondecreasing*** *on* (a, b) *if* $a < x < y < b$ *implies* $f(x) \leqslant f(y)$. *The term* ***isotone*** *is occasionally used to mean this.*

strictly decreasing *on* (a, b) *if* $a < x < y < b$ *implies* $f(x) > f(y)$.

decreasing *or* ***nonincreasing*** *on* (a, b) *if* $a < x < y < b$ *implies* $f(x) \geqslant f(y)$. *The term* ***antitone*** *is occasionally used to mean this.*[3]

monotone *on* (a, b) *if it is either increasing on* (a, b) *or decreasing on* (a, b).

Some authors, notably Landau [9, Definition 27, p. 88], say that f *is* ***increasing at c*** *if there exists some* $\varepsilon > 0$ *such that* $c - \varepsilon < x < c < y < c + \varepsilon$ *implies* $f(x) \leqslant f(c) \leqslant f(y)$.

f *is* ***decreasing at c*** *if there exists some* $\varepsilon > 0$ *such that* $c - \varepsilon < x < c < y < c + \varepsilon$ *implies* $f(x) \geqslant f(c) \geqslant f(y)$.

---

[2] Here, $\lim_{v\downarrow 0}$ only considers $v > 0$. That is,

$$\lim_{v\downarrow 0} \frac{f(x+v) - f(x)}{v} = c$$

means that for every $\varepsilon > 0$, there exists some $\delta > 0$ such that $0 < v < \delta$ implies $|f(x + v) - f(x)/v| < \varepsilon$. Corresponding definitions apply for infinite limits. For $\lim_{v\uparrow 0}$ we use $-\delta < v < 0$.

[3] Topkis [16] points out that the negation of the statement "f is increasing" is *not* "f is nonincreasing." For instance the sine function is not an increasing function, nor is it a nonincreasing function in my terminology. This does not seem to lead to much confusion however.

5 Corollary (Derivatives and Monotonicity) *Suppose $f\colon [a,b] \to \boldsymbol{R}$ is continuous on $[a,b]$ and differentiable on (a,b). If $f'(x) \geqslant 0$ for all $x \in (a,b)$, then f is nondecreasing on $[a,b]$. If $f'(x) > 0$ for all $x \in (a,b)$, then f is strictly increasing on $[a,b]$.*

Similarly if $f'(x) \leqslant 0$ for all $x \in (a,b)$, then f is nonincreasing on $[a,b]$. If $f'(x) < 0$ for all $x \in (a,b)$, then f is strictly decreasing on $[a,b]$.

If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant on $[a,b]$.

It is extremely important in the above theorem that $f'(x) > 0$ for all $x \in (a,b)$ in order to conclude that f is strictly increasing. If we know only that $f'(x_0) > 0$, we cannot conclude even that f is monotone! The following example is well known, see e.g., Marsden [11, Exercise 7.1.3, p. 209], but was introduced to me by Ket Richter.

6 Example (Nonmonotonicity with $f'(x_0) > 0$) Consider the function on $\boldsymbol{R}$ given by

$$f(x) = x + 2x^2 \sin \frac{1}{x^2}.$$

Then f is differentiable everywhere on $\boldsymbol{R}$, and $f'(0) = 1$, but f is not monotone on any open interval around 0. To see that f is differentiable, the only difficulty is at zero. But observe that f is squeezed between $g(x) = x+2x^2$ and $h(x) = x-2x^2$, which have the property that $g \geqslant f \geqslant h$ everywhere, $g(0) = f(0) = h(0) = 0$, and $g'(0) = h'(0) = 1$, so by the Squeezing Lemma 2, f is differentiable at zero and $f'(0) = 1$. For nonzero x, $f'(x) = 1 + 4x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}$, which is continuous and attains arbitrarily large positive and negative values in every neighborhood of zero. Therefore f cannot be monotone on a neighborhood of zero. See Figure 1. Note that the derivative of f is discontinuous at zero.[4]

Note that this function is increasing at zero in Landau's sense, but is not monotone on any open interval containing zero. □

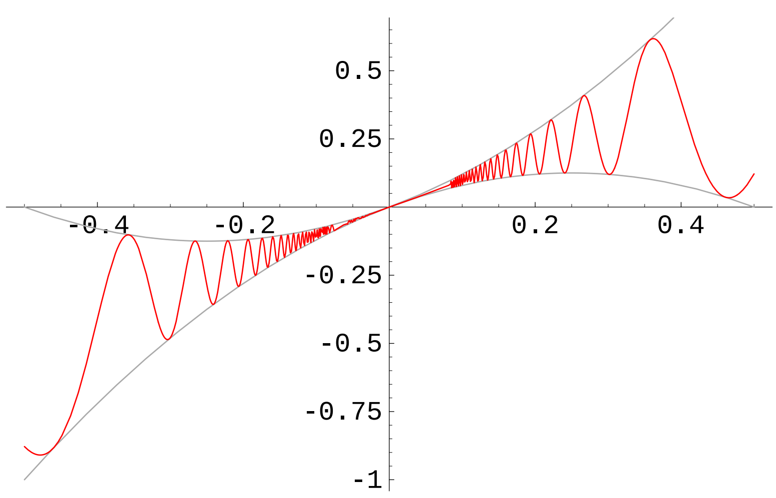

Figure 1. The nonmonotone function $x + 2x^2 \sin \frac{1}{x^2}$.

The next lemma is not hard to see. Compare it to Theorem 10 below.

[4]The idea of the derivative as the slope of a tangent line is weird in this case.

7 Lemma *Let $f: (a,b) \to \boldsymbol{R}$ be differentiable at x with $f'(x) > 0$. Then f is increasing at x (in Landau's sense). Likewise if $f'(x) < 0$. Then f is decreasing at x.*

Many of the results on maximization can be derived from Taylor's Theorem, which has two useful forms. The first is a generalization of the Mean Value Theorem, and assumes $n-1$ continuous derivatives on an open set and the existence everywhere of the n^{th} derivative. The version here is taken from Apostol [2, Theorem 5–14, p. 96].

8 Taylor's Theorem *Let $f: (a,b) \to \boldsymbol{R}$ be $n-1$ times continuously differentiable on (a,b) and assume that f has an n^{th} derivative at each point of (a,b). Fix a point x in (a,b). For any $v \neq 0$ such that $x+v$ belongs to (a,b), there is a point u strictly between 0 and v such that*

$$f(x+v) = f(x) + \sum_{k=1}^{n-1} \frac{f^{(k)}(x)}{k!} v^k + \frac{f^{(n)}(x+u)}{n!} v^n.$$

The other useful form of Taylor's Theorem is Young's form. It too assumes $n-1$ continuous derivatives on an open set, but assumes only that the n^{th} derivative exists at a point, and has a remainder term. This statement is a slight rewording of Serfling [14, Theorem C, p. 45], who cites Hardy [6, p. 278].

9 Young's Form of Taylor's Theorem *Let $f: (a,b) \to \boldsymbol{R}$ be $n-1$ times continuously differentiable on (a,b) and assume that f has an n^{th} derivative at the point x in (a,b). For any v such that $x+v$ belongs to (a,b),*

$$f(x+v) = f(x) + \sum_{k=1}^{n} \frac{f^{(k)}(x)}{k!} v^k + \frac{r(v)}{n!} v^n,$$

where the remainder term $r(v)$ satisfies

$$\lim_{v \to 0} r(v) = 0.$$

3 Extrema of a function of one variable

The main reference for these results is Apostol [3, pp. 181–195, 273–280].

3.1 Necessary first order conditions

We present the so-called **first order necessary conditions** for an interior extremum. They are called first order conditions because they involve first derivatives.

10 Theorem (Necessary First Order Conditions) *If $x^* \in (a,b)$ is a local extremum of f and f is differentiable at x^*, then $f'(x^*) = 0$.*

Proof: This is immediate from Lemma 7, but here is a proof from scratch.[5]

[5] This proof also serves to prove Lemma 7. In fact, this result is used to prove Rolle's Theorem, from which the Mean Value Theorem 3 is derived. Logically, this theorem should have been presented first.

If f has a local maximum at x^*, then $f(x^*) - f(x) \geqslant 0$ for every x in some neighborhood of x^*. Thus for x in that neighborhood,

$$\frac{f(x^*+v) - f(x^*)}{v} \quad \text{is} \quad \begin{cases} \leqslant 0 & \text{for} \quad v > 0 \\ \geqslant 0 & \text{for} \quad v < 0. \end{cases}$$

Therefore since the limit $f'(x^*) = \lim_{v \to 0} \frac{f(x^*+v)-f(x^*)}{v}$ exists, we must have simultaneously $f'(x^*) \leqslant 0$ and $f'(x^*) \geqslant 0$, which means, of course, that $f'(x^*) = 0$. A similar argument applies if x^* is a local minimizer. ∎

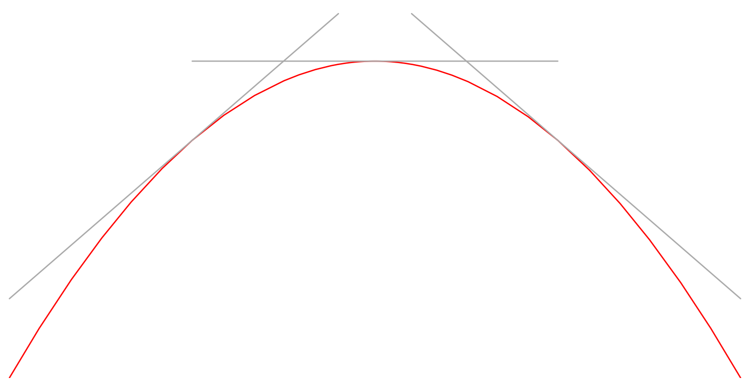

Figure 2. A nicely behaved maximum.

11 Theorem (Necessary Boundary First Order Conditions) *If a is a local maximizer of f on $[a, b]$ and f has a right-hand derivative at a, then $f'(a) \leqslant 0$.*

If b is a local maximizer of f on $[a, b]$ and f has a left-hand derivative at b, then $f'(b) \geqslant 0$.

The signs are reversed for minimizers.

Proof: The proof is essentially just half of the above argument. If f has a local maximum at a, then $f(x) - f(a) \leqslant 0$ for every $x \in [a, b]$ in some neighborhood of a. In particular $x \geqslant a$. Thus for $v > 0$ so that $x + v$ is in that neighborhood, $\frac{f(a+v)-f(a)}{v} \leqslant 0$, so $\lim_{v \downarrow 0} \frac{f(a+v)-f(a)}{v} \leqslant 0$.

The other cases are similar. ∎

Figure 2 also demonstrates this result for boundary minimizers.

3.2 Sufficient first order conditions

The next result is a straightforward consequence of the Mean Value Theorem 3, cf. Apostol [3, Theorems 4.6 and 4.7].

12 Theorem (Sufficient First Order Conditions) *Suppose f is continuous on $[a, b]$, and differentiable on (a, b) except perhaps at $c \in (a, b)$.*

- *If $f'(x) > 0$ for $x \in (a, c)$ and $f'(x) < 0$ for $x \in (c, b)$, then c is a strict maximizer of f on $[a, b]$.*
- *If $f'(x) \geqslant 0$ for $x \in (a, c)$ and $f'(x) \leqslant 0$ for $x \in (c, b)$, then c maximizes f on $[a, b]$.*
- *If $f'(x) < 0$ for $x \in (a, c)$ and $f'(x) > 0$ for $x \in (c, b)$, then c is a strict minimizer of f on $[a, b]$.*
- *If $f'(x) \leqslant 0$ for $x \in (a, c)$ and $f'(x) \geqslant 0$ for $x \in (c, b)$, then c is a minimizer of f on $[a, b]$.*

The conditions of Theorem 12 are sufficient conditions for the existence of a maximum at a point, but are hardly necessary.

13 Example (An unruly maximum) This is almost the same as Example 6. Consider the function

$$f(x) = \begin{cases} -x^2(2 + \sin(\frac{1}{x})) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

(see Figure 3). Clearly f is differentiable for all nonzero x, and by the Squeezing Lemma 2, f is differentiable at zero since it is squeezed between $-x^2$ and $-3x^2$. In fact,

$$f'(x) = \begin{cases} -2x(2 + \sin(\frac{1}{x})) + \cos(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Observe that f achieves a strict local maximum at zero, but its derivative switches signs infinitely often on both sides of zero (since for small x the $\cos(\frac{1}{x})$ term determines the sign of $f'(x)$). In particular, the function is neither increasing on any interval $(a, 0)$ nor decreasing on any interval $(0, b)$. This example appears in Sydsaeter [15]. □

3.3 Second order and n^{th} order sufficient conditions

The next theorem has slightly weaker hypotheses than the standard statement, which assumes that f is n times continuously differentiable on an interval, see, e.g., Apostol [2, Theorem 7–7, p. 148].

14 Theorem (Higher order sufficient conditions) *Let $f\colon (a, b) \to \boldsymbol{R}$ be $n{-}1$ times continuously differentiable on (a, b) and assume that it has an n^{th} derivative at the interior point x^*. Suppose in addition that*

$$f'(x^*) = f''(x^*) = \cdots = f^{(n-1)}(x^*) = 0 \quad \text{and} \quad f^{(n)}(x^*) \neq 0.$$

- *If n is even, and $f^{(n)}(x^*) < 0$, then f has a strict local maximum at x^*.*
- *If n is even, and $f^{(n)}(x^*) > 0$, then f has a strict local minimum at x^*.*
- *If n is odd, then f has neither a local maximum nor a local minimum at x^*.*

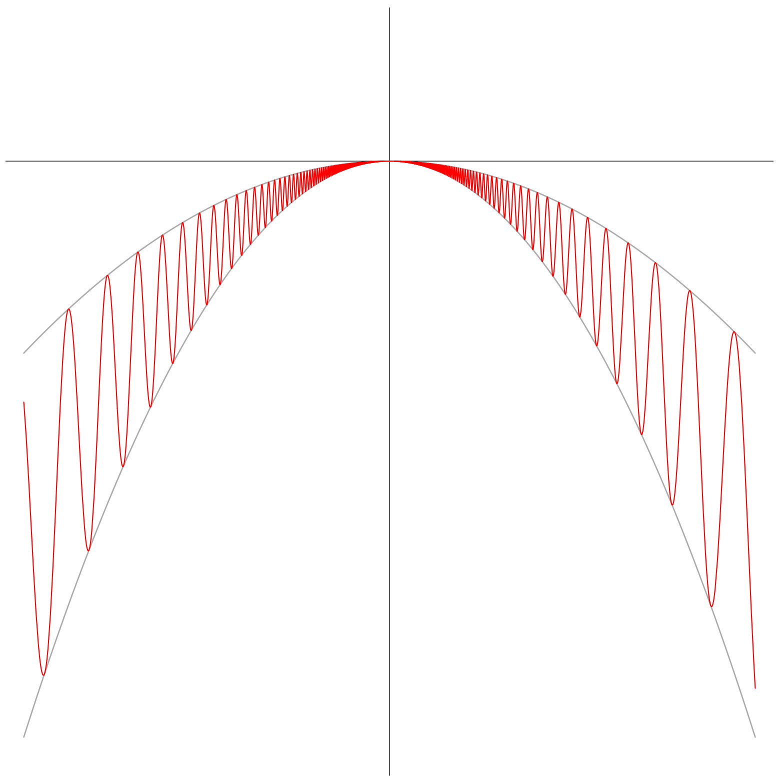

Figure 3. An unruly maximum at 0 for $-x^2\left(2+\sin(\frac{1}{x})\right)$.

Proof: From Young's form of Taylor's Theorem 9, for any v with $x^* + v \in (a, b)$ we have

$$f(x^* + v) - f(x^*) = \frac{f^{(n)}(x^*) + r(v)}{n!} v^n,$$

where $\lim_{v \to 0} r(v) = 0$. Thus there is some $\delta > 0$ such that for $|v| < \delta$, we have $|r(v)| < |f^{(n)}(x^*)|$. This implies that for $|v| < \delta$ the sign of $f^{(n)}(x^*) + r(v)$ is the same as the sign of $f^{(n)}(x^*)$, and the conclusion follows. ∎

15 Corollary (Necessary Second Order Conditions) *Let $f\colon (a, b) \to \boldsymbol{R}$ is continuously differentiable on a neighborhood of x^* and suppose $f''(x^*)$ exists. If x^* is a local maximizer of f, then $f''(x^*) \leqslant 0$. If x^* is a local minimizer of f, then $f''(x^*) \geqslant 0$.*

Proof: Assume first that x^* is a local maximizer. Then by Theorem 10 we must have $f'(x^*) = 0$, so by Theorem 14, if $f''(x^*) > 0$, then it is a *strict* local minimizer, so it cannot be a local maximizer. By contraposition then, $f''(x^*) \leqslant 0$. For the case of a local minimizer reverse the signs. ∎

16 Example (All derivatives vanishing at a strict maximum) It is possible for f to have derivatives of all orders that all vanish at a strict local maximizer. E.g., define

$$f(x) = \begin{cases} -e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Then 0 is a strict global maximizer of f. Furthermore f has derivatives of all orders everywhere,[6] and all these derivatives are zero.

I would graph this function for you, but it its graph near zero is visually indistinguishable from a horizontal line. For instance, according to Mathematica, $f(0.01) \approx -1.1 \times 10^{-4343}$, $f(0.0001) \approx -6.5 \times 10^{-43,429,449}$, and $f(0.000\,01)$ is beyond the numerical capability of my Intel Core Duo processor.) **Update:** I'm not sure when that last staement was written, but my late-2014 Mac Mini with an Intel Core i5 and Mathematica 11 reports that $f(0.000\,01) \approx -1.786 \times 10^{434,294,481,903}$ and $f(0.000\,001) \approx -1.5 \times 10^{43,429,448,190,325}$ before overflowing on $f(0.000\,000\,1)$.

This example appears in Apostol [2, Exercise 5–4, p. 98] and in Sydsaeter [15]. It is closely related to the function

$$f(x) = \begin{cases} e^{\frac{1}{x^2 - 1}} & |x| \leqslant 1 \\ 0 & |x| \geqslant 1, \end{cases}$$

which was shown by Cauchy to have continuous derivatives of all orders.[7] □

[6] Indeed, for $x \neq 0$, we have $f^{(n)}(x) = \frac{e^{-x^{-2}}}{p(x)}$, where p is a polynomial.

[7] At least Horváth [8, p. 166] attributes this function to Cauchy. See Aliprantis and Burkinshaw [1, Problem 25.2, p. 220] for a proof of the properties of Cauchy's function.

4 Concave and convex functions

A function f defined on an interval of $\boldsymbol{R}$ is

- **concave** if for all x, y and all $0 < \alpha < 1$,

$$f((1-\alpha)x + \alpha y) \geqslant (1-\alpha)f(x) + \alpha f(y)$$

- **strictly concave** if for all $x \neq y$ and all $0 < \alpha < 1$,

$$f((1-\alpha)x + \alpha y) > (1-\alpha)f(x) + \alpha f(y)$$

- **convex** if for all x, y and all $0 < \alpha < 1$,

$$f((1-\alpha)x + \alpha y) < (1-\alpha)f(x) + \alpha f(y)$$

- **strictly convex** if for all $x \neq y$ and all $0 < \alpha < 1$,

$$f((1-\alpha)x + \alpha y) < (1-\alpha)f(x) + \alpha f(y)$$

These functions are especially nice in the theory of optimization for the following reasons.

17 Theorem *If f is concave and x^* is a local maximizer of f, then it is a global maximizer. If f is convex and x^* is a local minimizer of f, then it is a global minimizer.*

18 Theorem *If f is concave and differentiable at x^* and if $f'(x^*) = 0$, then x^* is a global maximizer of f. If f is strictly concave and differentiable at x^* and if $f'(x^*) = 0$, then x^* is a strict global maximizer of f.*

If f is convex and differentiable at x^* and if $f'(x^*) = 0$, then x^* is a global minimizer of f. If f is strictly convex and differentiable at x^* and if $f'(x^*) = 0$, then x^* is a strict global minimizer of f.

Concave and convex functions can be characterized in terms of their derivatives.

19 Theorem *Let f be a real function on an interval, and assume that f is differentiable everywhere. Then f is concave if and only if f' is nonincreasing, and f is strictly concave if and only if f' is strictly decreasing.*

Similarly, f is convex if and only if f' is nondecreasing, and f is strictly convex if and only if f' is strictly increasing.

If f is twice differentiable everywhere, then f is concave if and only if $f'' \leqslant 0$ everywhere. If $f'' < 0$ everywhere, then f is strictly concave.

Similarly, f is convex if and only if $f'' \geqslant 0$ everywhere. If $f'' > 0$ everywhere, then f is strictly convex.

We cannot strengthen the last conclusion. For example, $f(x) = -x^2$ is a strictly concave function with $f''(0) = 0$.

References

[1] C. D. Aliprantis and O. Burkinshaw. 1999. *Problems in real analysis*, 2d. ed. San Diego: Academic Press. Date of publication, 1998. Copyright 1999. 7

[2] T. M. Apostol. 1957. *Mathematical analysis: A modern approach to advanced calculus*. Addison-Wesley series in mathematics. Reading, Massachusetts: Addison Wesley. 2, 3.3, 16

[3] ——— . 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell. 1, 2, 3, 3.2

[4] G. Debreu. 1959. *Theory of value: An axiomatic analysis of economic equilibrium*. Number 17 in Cowles Foundation Monographs. New Haven: Yale University Press. http://cowles.econ.yale.edu/P/cm/m17/m17-all.pdf 1

[5] H. Hancock. 1960. *Theory of maxima and minima*. New York: Dover. Reprint of 1917 edition published by Ginn–Blaisdell. 1

[6] G. H. Hardy. 1952. *A course of pure mathematics*, 10th. ed. Cambridge: Cambridge University Press. 2

[7] J.-B. Hiriart-Urruty and C. Lemaréchal. 2001. *Fundamentals of convex analysis*. Grundlehren Text Editions. Berlin: Springer–Verlag.

[8] J. Horváth. 1966. *Topological vector spaces and distributions*, volume 1. Reading, Mass.: Addison Wesley. 7

[9] E. Landau. 1960. *Foundations of analysis*. New York: Chelsea. Translation of *Grundlagen der Analysis*, published in 1930. 4

[10] D. G. Luenberger. 1995. *Microeconomic theory*. New York: McGraw-Hill. 1

[11] J. E. Marsden. 1974. *Elementary classical analysis*. San Francisco: W. H. Freeman and Company. 1, 2

[12] A. Mas-Colell, M. D. Whinston, and J. R. Green. 1995. *Microeconomic theory*. Oxford: Oxford University Press. 1

[13] R. T. Rockafellar. 1970. *Convex analysis*. Number 28 in Princeton Mathematical Series. Princeton: Princeton University Press.

[14] R. J. Serfling. 1980. *Approximation theorems of mathematical statistics*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley. 2

[15] K. Sydsaeter. 1974. Letter to the editor on some frequently occurring errors in the economic literature concerning problems of maxima and minima. *Journal of Economic Theory* 9(4):464–466. DOI: 10.1016/0022-0531(74)90046-5 13, 16

[16] D. M. Topkis. 1998. *Supermodularity and complementarity*. Princeton: Princeton University Press. 3

[17] H. R. Varian. 1992. *Microeconomic analysis*, 3d. ed. New York: W. W. Norton & Co. 1