References

- Convex Optimization by Stephen Boyd and Lieven Vandenberghe
- Lectures on Modern Convex Optimization by Aharon Ben-Tal and Arkadi Nemirovski
- Convex Analysis by R. T. Rockafellar, Vol. 28 of Princeton Math. Series, Princeton Univ. Press, 1970 (470 pages)
- Numerical Optimization by Nocedal, Jorge, Wright, Stephen
- Introduction to Nonlinear Optimization - Theory, Algorithms and Applications by Amir Beck

More exhaustive list at www.cse.iitb.ac.in/~cs709
Please check calendar page for all notes/specific references
www.cse.iitb.ac.in/~cs709/calendar.html

Developing Tools for Convexity Analysis of $f(x_1, x_2, ..x_n)$

Instructor: Prof. Ganesh Ramakrishnan

Summary of Optimization Principles for Univariate Functions

Detailed slides at https://www.cse.iitb.ac.in/~cs709/notes/enotes/2-08-01-2018-univariateprinciples.pdf, video at https://tinyurl.com/yc4d2aqg and Section 4.1.1 (pages 213 to 214) of the notes at https://www.cse.iitb.ac.in/~cs709/notes/BasicsOfConvexOptimization.pdf.

Maximum and Minimum values of univariate functions

Let $f : \mathcal{D} \rightarrow \Re$. Now f has

- An *absolute maximum* (or global maximum) value at point $c \in \mathcal{D}$ if

$$f(x) \leq f(c), \ \forall x \in \mathcal{D}$$

- An *absolute minimum* (or global minimum) value at $c \in \mathcal{D}$ if

$$f(x) \geq f(c), \ \forall x \in \mathcal{D}$$

- A *local maximum value* at c if there is an open interval $\mathcal{I}$ containing c in which $f(c) \geq f(x), \ \forall x \in \mathcal{I}$
- A *local minimum value* at c if there is an open interval $\mathcal{I}$ containing c in which $f(c) \leq f(x), \ \forall x \in \mathcal{I}$
- A *local extreme value* at c, if $f(c)$ is either a local maximum or local minimum value of f in an open interval $\mathcal{I}$ with $c \in \mathcal{I}$

First Derivative Test & Extreme Value Theorem

First derivative test for local extreme value of f, when f is differentiable at the extremum.

$f'(x) = 0$ for all local extreme values

First Derivative Test & Extreme Value Theorem

First derivative test for local extreme value of f, when f is differentiable at the extremum.

Claim

If $f(c)$ is a local extreme value and if f is differentiable at $x = c$, then $f'(c) = 0$.

The Extreme Value Theorem

Function has global extremes if (a) it is continuous
(b) the domain is bounded
(c) the domain is closed

First Derivative Test & Extreme Value Theorem

First derivative test for local extreme value of f, when f is differentiable at the extremum.

Claim

If $f(c)$ is a local extreme value and if f is differentiable at $x = c$, then $f'(c) = 0$.

The Extreme Value Theorem

Claim

A continuous function $f(x)$ on a closed and bounded interval $[a, b]$ attains a minimum value $f(c)$ for some $c \in [a, b]$ and a maximum value $f(d)$ for some $d \in [a, b]$. That is, a continuous function on a closed, bounded interval attains a minimum and a maximum value.

We must point out that either or both of the values c and d may be attained at the end points of the interval $[a, b]$.

Taylor's Theorem and n^{th} degree polynomial approximation

The n^{th} degree polynomial approximation of a function is used to prove a generalization of the mean value theorem, called the *Taylor's theorem*.

Claim

The Taylor's theorem states that if f and its first n derivatives $f', f'', \ldots, f^{(n)}$ are continuous on the closed interval $[a, b]$, and differentiable on (a, b), then there exists a number $c \in (a, b)$ such that

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2!}f''(a)(b-a)^2 + \ldots + \frac{1}{n!}f^{(n)}(a)(b-a)^n + \frac{1}{(n+1)!}f^{(n+1)}(c)(b-a)^{n+1}$$

Mean Value Theorem = Taylor's theorem with n = 0

approximation involves dropping last term

Taylor's Theorem and n^{th} degree polynomial approximation

The n^{th} degree polynomial approximation of a function is used to prove a generalization of the mean value theorem, called the *Taylor's theorem*.

Claim

The Taylor's theorem states that if f and its first n derivatives $f', f'', \ldots, f^{(n)}$ are continuous on the closed interval $[a, b]$, and differentiable on (a, b), then there exists a number $c \in (a, b)$ such that

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2!}f''(a)(b-a)^2 + \ldots + \frac{1}{n!}f^{(n)}(a)(b-a)^n + \frac{1}{(n+1)!}f^{(n+1)}(c)(b-a)^{n+1}$$

Mean Value Theorem = Taylor's theorem with n =0

Mean Value, Taylor's Theorem and words of caution

Note that if f fails to be differentiable at even one number in the interval, then the conclusion of the mean value theorem may be false. For example, if $f(x) = x^{2/3}$, then $f'(x) = \frac{2}{3\sqrt[3]{x}}$ and

Mean Value, Taylor's Theorem and words of caution

Note that if f fails to be differentiable at even one number in the interval, then the conclusion of the mean value theorem may be false. For example, if $f(x) = x^{2/3}$, then $f'(x) = \frac{2}{3\sqrt[3]{x}}$ and the theorem does not hold in the interval $[-3, 3]$, since f is not differentiable at 0 as can be seen in Figure 1.

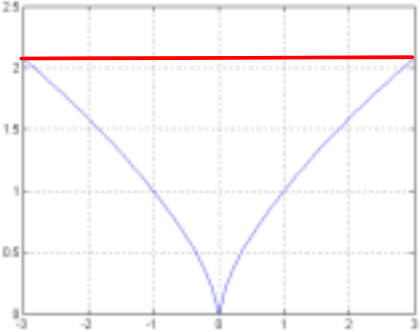


Figure 1:

Sufficient Conditions for Increasing and decreasing functions

A function f is said to be ...

- *increasing* on an interval $\mathcal{I}$ in its domain $\mathcal{D}$ if $f(t) < f(x)$ whenever $t < x$.
- *decreasing* on an interval $\mathcal{I} \in \mathcal{D}$ if $f(t) > f(x)$ whenever $t < x$.

Consequently:

Claim

Let $\mathcal{I}$ be an interval and suppose f is continuous on $\mathcal{I}$ and differentiable on $int(\mathcal{I})$. Then:

1. *if $f'(x) > 0$ for all $x \in int(\mathcal{I})$, then f is* (strictly) increasing

Sufficient Conditions for Increasing and decreasing functions

A function f is said to be ...

- *increasing* on an interval $\mathcal{I}$ in its domain $\mathcal{D}$ if $f(t) < f(x)$ whenever $t < x$.
- *decreasing* on an interval $\mathcal{I} \in \mathcal{D}$ if $f(t) > f(x)$ whenever $t < x$.

Consequently:

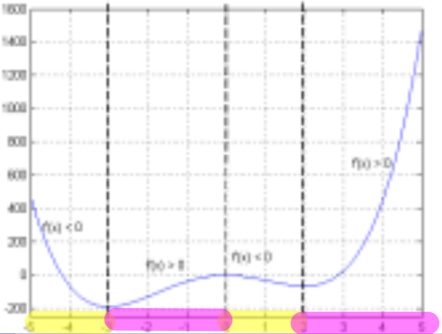
Claim

Let $\mathcal{I}$ be an interval and suppose f is continuous on $\mathcal{I}$ and differentiable on $int(\mathcal{I})$. Then:

1. *if $f'(x) > 0$ for all $x \in int(\mathcal{I})$, then f is increasing on $\mathcal{I}$;*
2. *if $f'(x) < 0$ for all $x \in int(\mathcal{I})$, then f is decreasing on $\mathcal{I}$;*
3. *if $f'(x) = 0$ for all $x \in int(\mathcal{I})$, iff, f is constant on $\mathcal{I}$.*

Illustration of Sufficient Conditions

Figure 2 illustrates the intervals in $(-\infty, \infty)$ on which the function $f(x) = 3x^4 + 4x^3 - 36x^2$ is decreasing and increasing. First we note that $f(x)$ is differentiable everywhere on $(-\infty, \infty)$ and compute $f'(x) = 12(x^3 + x^2 - 6x) = 12(x-2)(x+3)x$, which is negative in the intervals $(-\infty, -3]$ and $[0, 2]$ and positive in the intervals $[-3, 0]$ and $[2, \infty)$. We observe that f is decreasing in the intervals $(-\infty, -3]$ and $[0, 2]$ and while it is increasing in the intervals $[-3, 0]$ and $[2, \infty)$.



Necessary conditions for increasing/decreasing function

The conditions for increasing and decreasing properties of $f(x)$ stated so far are

Necessary conditions for increasing/decreasing function

The conditions for increasing and decreasing properties of $f(x)$ stated so far are not necesssary.

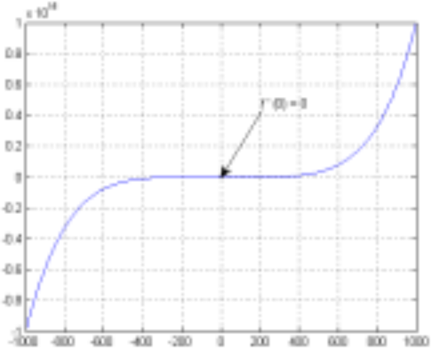


Figure 3:

Figure 3 shows that for the function $f(x) = x^5$, though $f(x)$ is increasing in $(-\infty, \infty)$, $f'(0) = 0$.

Another sufficient condition for increasing/decreasing function

Thus, a modified sufficient condition for a function f to be increasing/decreasing on an interval $\mathcal{I}$ can be stated as follows:

f'(.) > 0 everywhere except at a finite number of points where f'(.) = 0

Another sufficient condition for increasing/decreasing function

Thus, a modified sufficient condition for a function f to be increasing/decreasing on an interval $\mathcal{I}$ can be stated as follows:

Claim

Let $\mathcal{I}$ be an interval and suppose f is continuous on $\mathcal{I}$ and differentiable on $int(\mathcal{I})$. Then:

1. *if $f'(x) \geq 0$ for all $x \in int(\mathcal{I})$, and if $f'(x) = 0$ at only finitely many $x \in \mathcal{I}$, then f is increasing on $\mathcal{I}$;*
2. *if $f'(x) \leq 0$ for all $x \in int(\mathcal{I})$, and if $f'(x) = 0$ at only finitely many $x \in \mathcal{I}$, then f is decreasing on $\mathcal{I}$.*

For example, the derivative of the function $f(x) = 6x^5 - 15x^4 + 10x^3$ vanishes at 0, and 1 and $f'(x) > 0$ elsewhere. So $f(x)$ is increasing on $(-\infty, \infty)$.

Necessary conditions for increasing/decreasing function (contd.)

We have a slightly different necessary condition..

Claim

Let $\mathcal{I}$ be an interval, and suppose f is continuous on $\mathcal{I}$ and differentiable in int($\mathcal{I}$). Then:

1. *if f is increasing on $\mathcal{I}$, then $f'(x) \geq 0$ for all $x \in int(\mathcal{I})$;*
2. *if f is decreasing on $\mathcal{I}$, then $f'(x) \leq 0$ for all $x \in int(\mathcal{I})$.*

Critical Point

This concept will help us derive the general condition for local extrema.

Definition

[Critical Point]: *A point c in the domain $\mathcal{D}$ of f is called a critical point of f if either $f'(c) = 0$ or $f'(c)$ does not exist.*

The following general condition for local extrema extends the result in theorem 1 to general non-differentiable functions.

Claim

If $f(c)$ is a local extreme value, then c is a critical number of f.

The converse of above statement does not hold (see Figure 3); 0 is a critical number ($f'(0) = 0$), although $f(0)$ is not a local extreme value.

Critical Point and Local Extreme Value

Given a critical point c, the following test helps determine if $f(c)$ is a local extreme value:

Procedure

[Local Extreme Value]: *Let c be an isolated critical point of f*

1. *$f(c)$ is a local minimum if $f(x)$ is decreasing in an interval $[c-\epsilon_1, c]$ and increasing in an interval $[c, c+\epsilon_2]$ with $\epsilon_1, \epsilon_2 > 0$.*
2. *$f(c)$ is a local maximum if $f(x)$ is increasing in an interval $[c-\epsilon_1, c]$ and decreasing in an interval $[c, c+\epsilon_2]$ with $\epsilon_1, \epsilon_2 > 0$.*

First Derivative Test: Critical Point and Local Extreme Value

As an example, the function $f(x) = 3x^5 - 5x^3$ has

First Derivative Test: Critical Point and Local Extreme Value

As an example, the function $f(x) = 3x^5 - 5x^3$ has the derivative $f'(x) = 15x^2(x+1)(x-1)$. The critical points are

First Derivative Test: Critical Point and Local Extreme Value

As an example, the function $f(x) = 3x^5 - 5x^3$ has the derivative $f'(x) = 15x^2(x+1)(x-1)$. The critical points are 0, 1 and -1. Of the three, the sign of $f'(x)$ changes at 1 and -1, which are local minimum and maximum respectively. The sign does not change at 0, which is therefore not a local supremum.

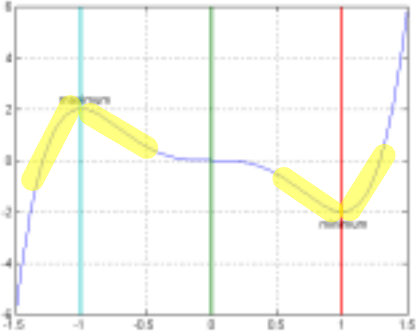


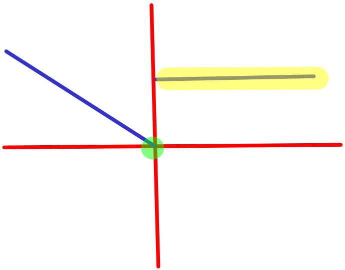
Figure 4:

First Derivative Test: Critical Point and Local Extreme Value

As another example, consider the function

$$f(x) = \begin{cases} -x & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Then,



First Derivative Test: Critical Point and Local Extreme Value

As another example, consider the function

$$f(x) = \begin{cases} -x & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Then,

$$f'(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x > 0 \end{cases}$$

Note that $f(x)$ is discontinuous at $x = 0$, and therefore $f'(x)$ is not defined at $x = 0$. All numbers $x \geq 0$ are critical numbers. $f(0) = 0$ is a local minimum, whereas $f(x) = 1$ is a local minimum as well as a local maximum $\forall x > 0$.

Strict Convexity and Extremum

- A differentiable function f is said to be *strictly convex* (or *strictly concave up*) on an open interval $\mathcal{I}$, *iff*, $f'(x)$ is increasing on $\mathcal{I}$.

Strict Convexity and Extremum

- A differentiable function f is said to be *strictly convex* (or *strictly concave up*) on an open interval $\mathcal{I}$, *iff*, $f'(x)$ is increasing on $\mathcal{I}$.
- Recall the graphical interpretation of the first derivative $f'(x)$; $f'(x) > 0$ implies that $f(x)$ is increasing at x.
- Similarly, $f'(x)$ is increasing when

Sufficient condition ==> f''(x) > 0
Sufficient condition ==> f''(x) >= 0
and f''(x) vanishes at a finite no.
of points

Necessary condition ==> f''(x) >=0

Strict Convexity and Extremum

- A differentiable function f is said to be *strictly convex* (or *strictly concave up*) on an open interval $\mathcal{I}$, *iff*, $f'(x)$ is increasing on $\mathcal{I}$. Definition (for a differentiable function)
- Recall the graphical interpretation of the first derivative $f'(x)$; $f'(x) > 0$ implies that $f(x)$ is increasing at x.
- Similarly, $f'(x)$ is increasing when $f''(x) > 0$. This gives us a sufficient condition for the strict convexity of a function:

Claim

If at all points in an open interval $\mathcal{I}$, $f(x)$ is doubly differentiable and if $f''(x) > 0$, $\forall x \in \mathcal{I}$, then the slope of the function is always increasing with x and the graph is strictly convex. This is illustrated in Figure 5.

- On the other hand, if the function is strictly convex and doubly differentiable in $\mathcal{I}$, then $f''(x) \geq 0$, $\forall x \in \mathcal{I}$. Necessary conditon for strict convexity for a differentiable function

Strict Convexity and Extremum (Illustrated)

Figure 5:

The function in [x1,x2] lies completely (strictly) below the line segment joining x1 to x2

Strict Convexity and Extremum: Slopeless interpretation (SI)

Claim

A function f is strictly convex on an open interval $\mathcal{I}$, iff

$$f(ax_1 + (1-a)x_2) < af(x_1) + (1-a)f(x_2) \quad (1)$$

whenver $x_1, x_2 \in \mathcal{I}$, $x_1 \neq x_2$ *and* $0 < a < 1$.

Strict Concavity

- A differentiable function f is said to be *strictly concave* on an open interval $\mathcal{I}$, *iff*, $f'(x)$ is decreasing on $\mathcal{I}$.
- Recall from theorem 4, the graphical interpretation of the first derivative $f'(x)$; $f'(x) < 0$ implies that $f(x)$ is decreasing at x.
- Similarly, $f'(x)$ is (strictly) monotonically decreasing when

$$f''(x) < 0$$

Strict Concavity

- A differentiable function f is said to be *strictly concave* on an open interval $\mathcal{I}$, *iff*, $f'(x)$ is decreasing on $\mathcal{I}$.
- Recall from theorem 4, the graphical interpretation of the first derivative $f'(x)$; $f'(x) < 0$ implies that $f(x)$ is decreasing at x.
- Similarly, $f'(x)$ is (strictly) monotonically decreasing when $f''(x) < 0$. This gives us a sufficient condition for the concavity of a function:

Claim

If at all points in an open interval $\mathcal{I}$, $f(x)$ is doubly differentiable and if $f''(x) < 0$, $\forall x \in \mathcal{I}$, then the slope of the function is always decreasing with x and the graph is strictly concave.

Strict Concavity

On the other hand, if the function is strictly concave and doubly differentiable in $\mathcal{I}$, then $f''(x) \leq 0, \ \forall x \in \mathcal{I}$. This is illustrated in Figure 6.

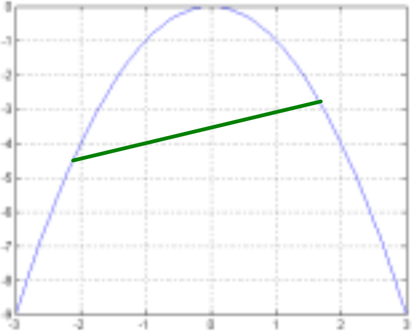


Figure 6:

Strict Concavity (slopeless interpretation)

There is also a slopeless interpretation of concavity as stated below:

Claim

A differentiable function f is strictly concave on an open interval $\mathcal{I}$, iff

$$f(ax_1 + (1-a)x_2) > af(x_1) + (1-a)f(x_2) \tag{2}$$

whenver $x_1, x_2 \in \mathcal{I}$, $x_1 \neq x_2$ *and* $0 < a < 1$.

The proof is similar to that for the slopeless interpretation of convexity.

Convex & Concave Regions and Inflection Point

Study the function $f(x) = x^3 - x + 2$.

Convex & Concave Regions and Inflection Point

Study the function $f(x) = x^3 - x + 2$. It's slope decreases as x increases to 0 ($f''(x) < 0$) and then the slope increases beyond $x = 0$ ($f''(x) > 0$). The point 0, where the $f''(x)$ changes sign is called the *inflection point*; the graph is strictly concave for $x < 0$ and strictly convex for $x > 0$. See Figure 7.

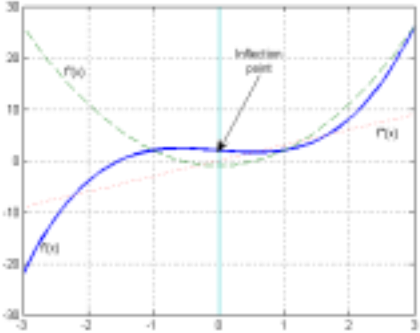


Figure 7:

Convex & Concave Regions and Inflection Point

Along similar lines, study the function $f(x) = \frac{1}{20}x^5 - \frac{7}{12}x^4 + \frac{7}{6}x^3 - \frac{15}{2}x^2$.

Convex & Concave Regions and Inflection Point

Along similar lines, study the function $f(x) = \frac{1}{20}x^5 - \frac{7}{12}x^4 + \frac{7}{6}x^3 - \frac{15}{2}x^2$.
It is strictly concave on $(-\infty, -1]$ and $[3, 5]$ and strictly convex on $[-1, 3]$ and $[5, \infty]$.
The inflection points for this function are at $x = -1$, $x = 3$ and $x = 5$.

First Derivative Test: Restated using Strict Convexity

The *first derivative test* for local extrema can be restated in terms of strict convexity and concavity of functions.

Expect convexity around a point of (local) minimum
And
Expect concavity around a point of (local) maximum

First Derivative Test: Restated using Strict Convexity

The *first derivative test* for local extrema can be restated in terms of strict convexity and concavity of functions.

Procedure

[First derivative test in terms of strict convexity]: *Let c be a critical number of f and* $f'(c) = 0$. *Then,*

1. $f(c)$ *is a local minimum if*

First Derivative Test: Restated using Strict Convexity

The *first derivative test* for local extrema can be restated in terms of strict convexity and concavity of functions.

Procedure

[First derivative test in terms of strict convexity]: *Let c be a critical number of f and* $f'(c) = 0$. *Then,*

1. *$f(c)$ is a local minimum if the graph of $f(x)$ is strictly convex on an open interval containing c.* sufficient condition for local min
2. *$f(c)$ is a local maximum if the graph of $f(x)$ is strictly concave on an open interval containing c.* sufficient condition for local max

Intuitively, relaxing strictness should give you sufficient conditions for local min/max ==> Revising with proofs for R^n case

Strict Convexity: Restated using Second Derivative

If the second derivative $f''(c)$ exists, then the strict convexity conditions for the critical number can be stated in terms of the sign of of $f''(c)$, making use of previous results. This is called the *second derivative test*.

Strict Convexity: Restated using Second Derivative

If the second derivative $f''(c)$ exists, then the strict convexity conditions for the critical number can be stated in terms of the sign of of $f''(c)$, making use of previous results. This is called the *second derivative test*.

Procedure

[Second derivative test]: *Let c be a critical number of f where $f'(c) = 0$ and $f''(c)$ exists.*

1. *If $f''(c) > 0$ then*

Strict Convexity: Restated using Second Derivative

If the second derivative $f''(c)$ exists, then the strict convexity conditions for the critical number can be stated in terms of the sign of of $f''(c)$, making use of previous results. This is called the *second derivative test*.

Procedure

[Second derivative test]: *Let c be a critical number of f where $f'(c) = 0$ and $f''(c)$ exists.*

1. *If $f''(c) > 0$ then $f(c)$ is a local minimum.* strict convexity
2. *If $f''(c) < 0$ then $f(c)$ is a local maximum.*
3. *If $f''(c) = 0$ then $f(c)$ could be a local maximum, a local minimum, neither or both. That is, the test fails.*

Convexity, Minima and Maxima: Illustrations

Study the functions $f(x) = x^4$, $f(x) = -x^4$ and $f(x) = x^3$:

Convexity, Minima and Maxima: Illustrations

Study the functions $f(x) = x^4$, $f(x) = -x^4$ and $f(x) = x^3$:

- If $f(x) = x^4$, then $f'(0) = 0$ and $f''(0) = 0$ and we can see that $f(0)$ is a local minimum.
- If $f(x) = -x^4$, then $f'(0) = 0$ and $f''(0) = 0$ and we can see that $f(0)$ is a local maximum.
- If $f(x) = x^3$, then $f'(0) = 0$ and $f''(0) = 0$ and we can see that $f(0)$ is neither a local minimum nor a local maximum. $(0,0)$ is an inflection point in this case.

Convexity, Minima and Maxima: Illustrations (contd.)

Study the functions: $f(x) = x + 2\sin x$ and $f(x) = x + \frac{1}{x}$:

- If $f(x) = x + 2\sin x$, then $f'(x) = 1 + 2\cos x$. $f'(x) = 0$ for $x = \frac{2\pi}{3}, \frac{4\pi}{3}$, which are the critical numbers. $f''\left(\frac{2\pi}{3}\right) = -2\sin\frac{2\pi}{3} = -\sqrt{3} < 0 \Rightarrow f\left(\frac{2\pi}{3}\right) = \frac{2\pi}{3} + \sqrt{3}$ is a local maximum value. On the other hand, $f''\left(\frac{4\pi}{3}\right) = \sqrt{3} > 0 \Rightarrow f\left(\frac{4\pi}{3}\right) = \frac{4\pi}{3} - \sqrt{3}$ is a local minimum value.
- If $f(x) = x + \frac{1}{x}$, then $f'(x) = 1 - \frac{1}{x^2}$. The critical numbers are $x = \pm 1$. Note that $x = 0$ is not a critical number, even though $f'(0)$ does not exist, because 0 is not in the domain of f. $f''(x) = \frac{2}{x^3}$. $f''(-1) = -2 < 0$ and therefore $f(-1) = -2$ is a local maximum. $f''(1) = 2 > 0$ and therefore $f(1) = 2$ is a local minimum.