Monotonic functions*

Sunil Kumar Singh

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 2.0†

The term "monotonic" conveys the meaning of maintaining order or the sense of "no change". In the context of function, we think a monotonic function as the one whose successive values are increasing, decreasing or constant. There is a sense of maintaining order of function values as independent variable changes. These aspects are pictorially evident on the graph of a function. In a general case, a function may or may not maintain its order of change in its domain i.e. in the overall context. However, we can always identify monotonic behavior in an appropriately chosen subset of domain – unless it is a point function or a singleton.

Consider the graph of sine function. As a whole, the function is not monotonic as the order of the function is not preserved over the domain of the function, which is "R". However, if we consider an interval, say, between "0" and "$\pi/2$", then we find that function keeps increasing with the increasing independent variable. Therefore, sine function is monotonic in this interval.

*Version 1.10: Sep 3, 2008 9:44 pm +0000
†http://creativecommons.org/licenses/by/2.0/

Monotonic function

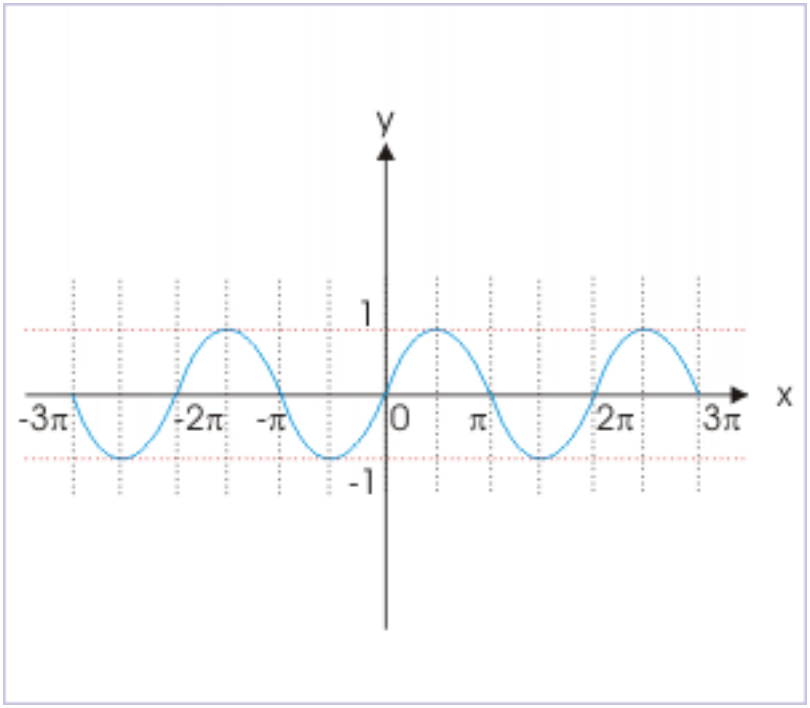

Figure 1: The sine function is monotonic in certain intervals.

On the other hand, a linear polynomial function represents a straight line, which maintains its monotonic nature through out its domain. The monotonic nature of a function, therefore, is investigated in a suitable interval, which is either domain or its subset. We shall refer this interval as X to illustrate the concept in this module. From the point of view of monotonic behavior, we classify function in following categories :

1: Constant function : Function values does not change as independent variable varies.

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) = f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

2: Strictly increasing: Function value change as independent variable varies in accordance with following condition :

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) < f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

3: Non-decreasing or increasing : Function value change as independent variable varies in accordance with following condition :

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \leq f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

4: Strictly decreasing: Function value change as independent variable varies in accordance with following condition :

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) > f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

5: Non-increasing or decreasing : Function value changes as independent variable varies in accordance with following condition :

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \geq f(x_2), \quad \text{for all} \quad x_1, x_2 \in X.$$

There is one ambiguity in the definition of classification presented above. According to the definition, a constant function is an increasing, decreasing or both kinds of function. Clearly, this interpretation is wrong and is an exception. An increasing or non-decreasing class actually captures the notion of an overall increasing function, which is intermittently constant and thereby distinguishes this class from strictly increasing order. Similarly, a decreasing or non-increasing class actually captures the notion of an overall decreasing function, which is intermittently constant and thereby distinguishes this class from strictly decreasing order.

Note : It may confound clarity, but we should know that there is another classification. In this classification (i) "strictly increasing" is known simply as "increasing", (ii) "strictly decreasing" is known simply as "decreasing", (iii) "increasing" is known as "monotonically increasing" and (iv) "decreasing" is known as "monotonically decreasing". Clearly, this classification is not the same as what is given here. The best way to deal with this situation is to ignore this confusion and be explicit in what we mean. Saying "strictly increasing" for example ensures that equality of function values is not allowed. Similarly, saying "non-decreasing" ensures that function values do not decrease. We shall try to adhere to this explicit classification to the extent possible.

Derivative and nature of function

We shall learn subsequently that first derivative of a function is defined in terms of ratio of the differences between two successive values of function and that of independent variable, however small is the difference in independent variable. It is easy to visualize, then, that if we know the nature of derivative in a given interval, then we can determine monotonic behavior of the function as well.

$$f\prime(x) = \underset{h \to 0}{Lim} \frac{f(x+h) - f(x)}{h}$$

Depending on the monotonic nature of function, the relative values of f(h) and f(x+h) are different and so the sign of first derivative.

1 Strictly increasing function

The successive value of function increases as the value of the independent variable increases. In other words, the preceding values are less than successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_{2,} \quad \text{then} \quad f(x_1) < f(x_2)$$

Increasing function

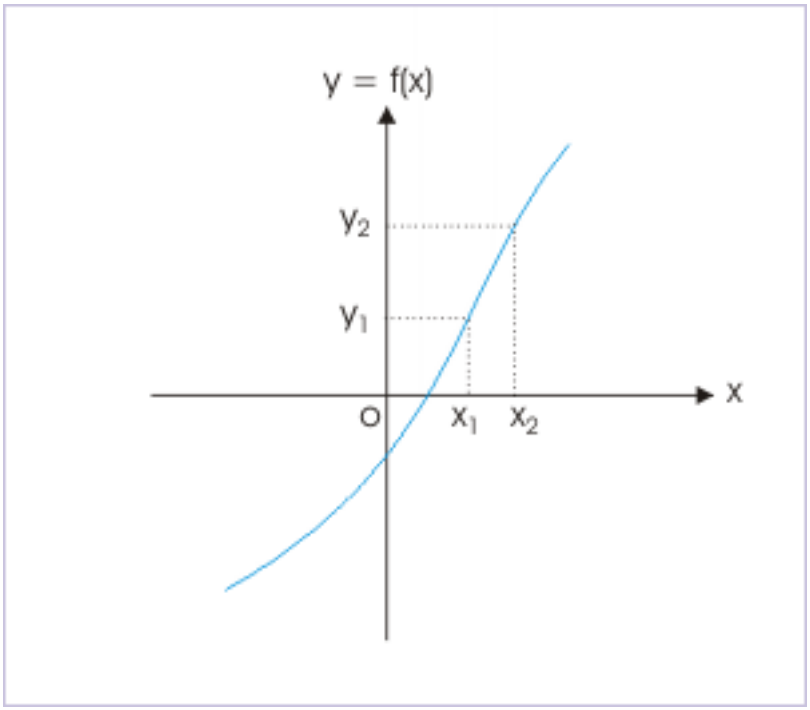

Figure 2: The successive function value is greater than previous value.

Example 1
Problem : Determine monotonic nature of the function in the interval [0,∞).

$$y = x^2$$

Solution : Let x_1 and x_2 belong to the interval [0,∞) such that $x_1 < x_2$. Multiplying inequality with x_1 (a positive number) does not change the nature of inequality :

$$\Rightarrow x_1^2 < x_1 x_2$$

Multiplying inequality with x_2 (a positive number) does not change the nature of inequality :

$$\Rightarrow x_1 x_2 < x_2^2$$

Combining two inequalities,

$$\Rightarrow x_1^2 < x_2^2$$

$$\Rightarrow f(x_1) < f(x_2)$$

Thus, given function is strictly decreasing in [0,∞).

As f(x_1)<f(x_2) for all x_1,x_2∈X, the difference "f(x+h) − f(x)" is positive for "h", however small. This implies that the first derivative of function is positive. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for couple of x values, while curve is continuously increasing in

the interval. It means that first derivative can be equal to zero for few points in the interval in which it is strictly increasing. This is clear from the figure given here,

Strictly increasing function

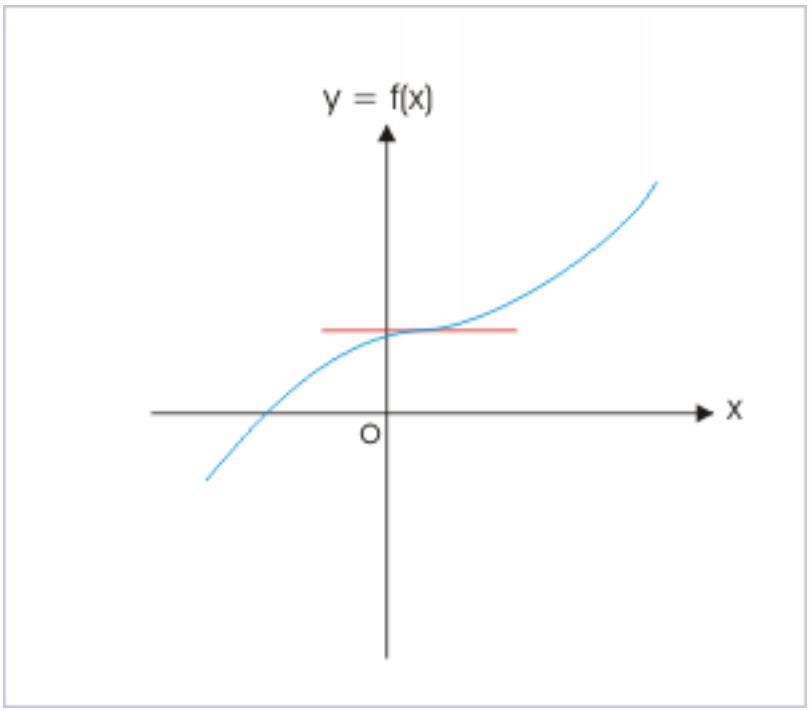

Figure 3: The first derivative can be zero at certain point(s) - but not in continuous interval.

Thus, for strictly increasing function,

$$f\prime(x) \geq 0; \quad \text{Equality sign holds for points only - not on a continuous section in X}$$

For strictly increasing function, if $x_1 < x_2$, then f(x_1) < f(x_2), for all x_1,x_2∈X. It means that all distinct x values correspond to distinct y values and vice-versa. Therefore, strictly increasing function is one-one function i.e. a bijection and hence "invertible". In other words, if a function has strict increasing order, then it is invertible. Mathematically, we say that if f'(x) ≥0; (equality holding for points only), x∈X, then function is invertible in X. For example, consider sine function,

$$f(x) = \sin x$$

$$\Rightarrow f\prime(x) = \cos x$$

We know that cosx is positive in the interval [-π/2,π/2]. Hence sine function is a strictly increasing function in [-π/2,π/2] and is invertible. Recall that inverse sine function is defined in this interval.

The order of a function provides an easy technique to determine range of a continuous function, corresponding to a given domain interval. For example, if domain of a continuously increasing function, f(x), is [x_1, x_2], then the least value of the function is f(x_1) and greatest value of the function is f(x_2). Hence, range of the function is :

$$\text{Range} = [f(x_1), f(x_2)]$$

We shall study this aspect of finding range in detail in a separate module.

2 Non-decreasing function or increasing

The successive value of function increases or remains constant as the value of the independent variable increases. In other words, the preceding values are less than or equal to successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \leq f(x_2)$$

Non decreasing function

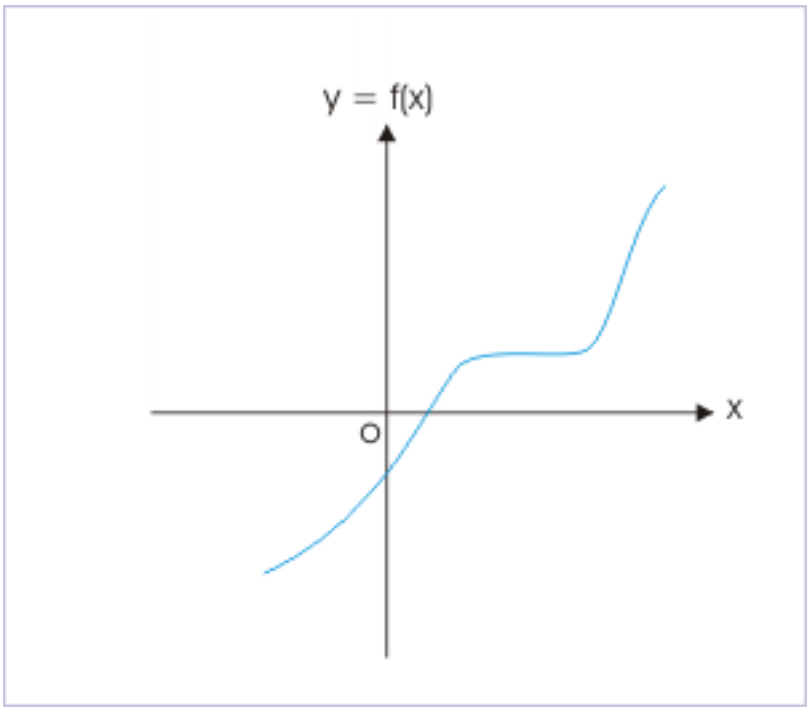

Figure 4: The successive function value is greater than or equal to previous value.

As f(x_1)≤f(x_2) for all x_1,x_2∈X, the difference "f(x+h) − f(x)" is non-negative for "h", however small. This implies that the first derivative of function is non-negative. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for a subset of X, while curve is increasing overall in the interval. It means that first derivative can be equal to zero points or sub-intervals in which it is increasing. Thus, for non-decreasing function,

$$f\prime(x) \geq 0; \quad \text{Equality sign holds for few points or a continuous section in X}$$

For increasing function, if x_1<x_2, then f(x_1) ≤ f(x_2), for all x_1,x_2∈X. This means that there may be same function values for different values of x. This is "many one" relation and as such function is not invertible in X.

3 Strictly decreasing function

The successive value of function decreases as the value of the independent variable increases. In other words, the preceding values are greater than successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) > f(x_2)$$

Strictly decreasing function

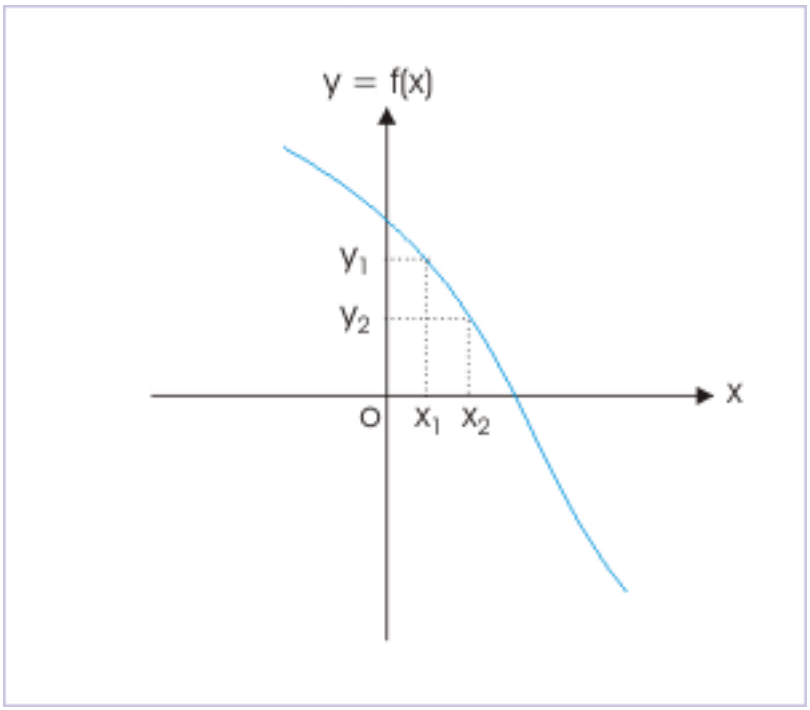

Figure 5: The successive function value is less than previous value.

Example 2
Problem : Determine monotonic nature of the function in the interval (-∞,0].

$$y = x^2$$

Solution : Let x_1 and x_2 belong to the interval [0,∞) such that $x_1 < x_2$. Multiplying inequality with x_1 (a negative number) changes the nature of inequality :

$$\Rightarrow x_1^2 > x_1 x_2$$

Multiplying inequality with x_2 (a negative number) changes the nature of inequality :

$$\Rightarrow x_1 x_2 > x_2^2$$

Combining two inequalities,

$$\Rightarrow x_1^2 > x_2^2$$

$$\Rightarrow f(x_1) > f(x_2)$$

Thus, given function is strictly decreasing in (-∞,0].

As f(x_1)>f(x_2) for all x_1,x_2∈X, the difference "f(x+h) – f(x)" is negative for "h", however small. This implies that the first derivative of function is negative. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for couple of x values, while curve is continuously decreasing in the interval. It means that first derivative can be equal to zero for few points in the interval in which it is strictly decreasing. Thus, for strictly decreasing function,

$$f\prime(x) \leq 0; \quad \text{Equality sign holds for points only - not on a continuous section in X}$$

For strictly decreasing function, if x_1<x_2, then f(x_1) > f(x_2), for all x_1,x_2∈X. It means that all distinct x values correspond to distinct y values and vice-versa. Therefore, strictly decreasing function is one-one function i.e. a bijection and hence "invertible". In other words, if a function has strict decreasing order, then it is invertible. Mathematically, we say that if f'(x) ≤ 0 (equality holding for points only); x∈X, then function is invertible in X. For example, consider sine function,

$$f(x) = \sin x$$

$$\Rightarrow f\prime(x) = \cos x$$

We know that cosx is negative in the interval [$\pi/2$, $3\pi/2$]. Hence sine function is a strictly decreasing function in [$\pi/2$, $3\pi/2$] and is invertible. Recall though that inverse sine function is not defined in this interval, but in basic interval about origin [-$\pi/2$,$\pi/2$].

The order of a function provides an easy technique to determine range of a continuous function, corresponding to a given domain interval. For example, if domain of a continuously decreasing function, f(x), is [x_1, x_2], then the least value of the function is f(x_2) and greatest value of the function is f(x_1). Hence, range of the function is :

$$\text{Range} = [f(x_2), f(x_1)]$$

We shall study this aspect of finding range in detail in a separate module.

4 Non-increasing function or decreasing

The successive value of function decreases or remains constant as the value of the independent variable increases. In other words, the preceding values are greater than or equal to successive values that follow. Mathematically,

$$\text{If} \quad x_1 < x_2 \quad \text{then} \quad f(x_1) \geq f(x_2)$$

Non-increasing function or decreasing

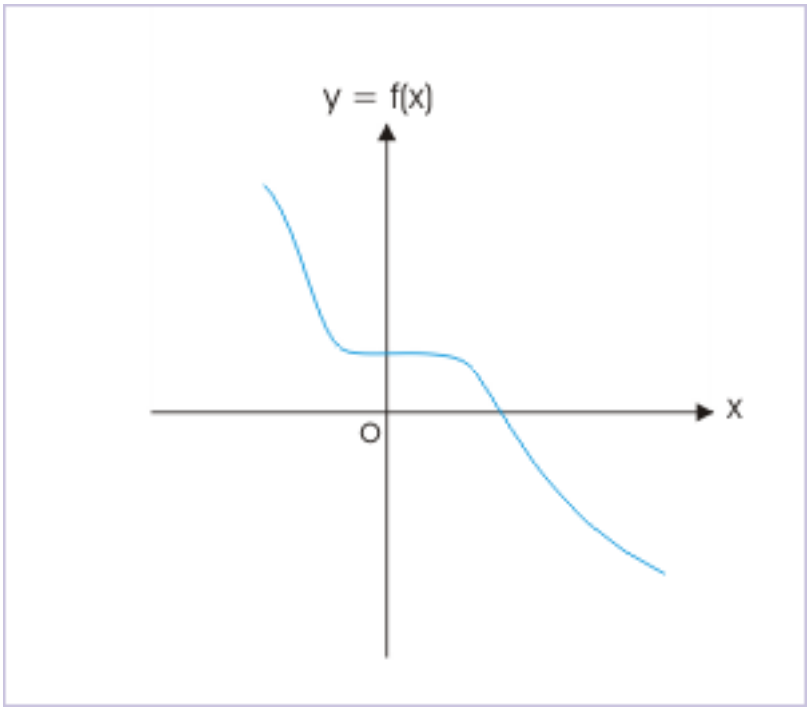

Figure 6: The successive function value is less than or equal to previous value.

As f(x_1)$\geq$ f(x_2) for all x_1,$x_2$$\in$X, the difference "f(x+h) − f(x)" is non-positive for "h", however small. This implies that the first derivative of function is non-positive. If we think of possibility, then we can realize that tangent to the function curve can be parallel to x-axis for a subset of X, while curve is decreasing overall in the interval. It means that first derivative can be equal to zero at points or in sub-intervals in which it is decreasing. Thus, for non-decreasing function,

$$f\prime(x) \leq 0; \quad \text{Equality sign holds for few points or a continuous section in X}$$

For decreasing function, if x_1<x_2, then f(x_1) $\geq$ f(x_2), for all x_1,$x_2$$\in$X. This means that there may be same function values for different values of x. This is "many one" relation and as such function is not invertible in X.