Optimization using Calculus

Convexity and Concavity of Functions of One and Two Variables

D Nagesh Kumar, IISc

Objective

- To determine **the convexity and concavity of functions**

Convex Function (Function of one variable)

- A real-valued function *f* defined on an interval (or on any convex subset *C* of some vector space) is called **convex**, if for any two points *x* and *y* in its domain *C* and any *t* in [0,1], we have

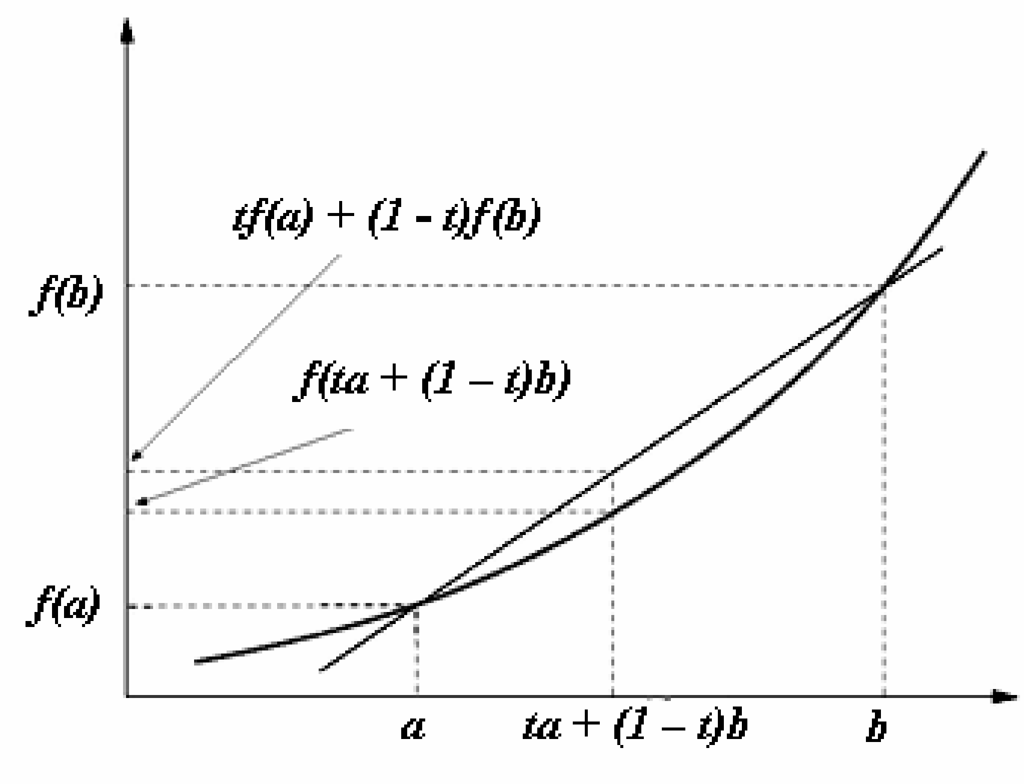
$$f(ta+(1-t)b) \le tf(a)+(1-t)f(b))$$

- In other words, a function is convex if and only if its epigraph (the set of points lying on or above the graph) is a convex set. A function is also said to be *strictly convex* if

$$f(ta+(1-t)b) < tf(a)+(1-t)f(b)$$

for any *t* in (0,1) and a line connecting any two points on the function lies completely above the function.

A convex function



Testing for convexity of a single variable function

- A function is convex if its slope is non decreasing or $\partial^2 f / \partial x^2 \geq 0$. It is strictly convex if its slope is continually increasing or $\partial^2 f / \partial x^2 > 0$ throughout the function.

Properties of convex functions

- A convex function *f*, defined on some convex open interval *C*, is continuous on *C* and differentiable at all or at many points. If *C* is closed, then *f* may fail to be continuous at the end points of *C*.

- A continuous function on an interval *C* is convex if and only if

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2} \quad \text{for all a and b in } C.$$

- A differentiable function of one variable is convex on an interval if and only if its derivative is monotonically non-decreasing on that interval.

Properties of convex functions (contd.)

- A continuously differentiable function of one variable is convex on an interval if and only if the function lies above all of its tangents: for all *a* and *b* in the interval.

- A twice differentiable function of one variable is convex on an interval if and only if its second derivative is non-negative in that interval; this gives a practical test for convexity.

- More generally, a continuous, twice differentiable function of several variables is convex on a convex set if and only if its Hessian matrix is positive semi definite on the interior of the convex set.

- If two functions f and g are convex, then so is any weighted combination $af + b\,g$ with non-negative coefficients a and b. Likewise, if f and g are convex, then the function $\max\{f, g\}$ is convex.

Concave function (function of one variable)

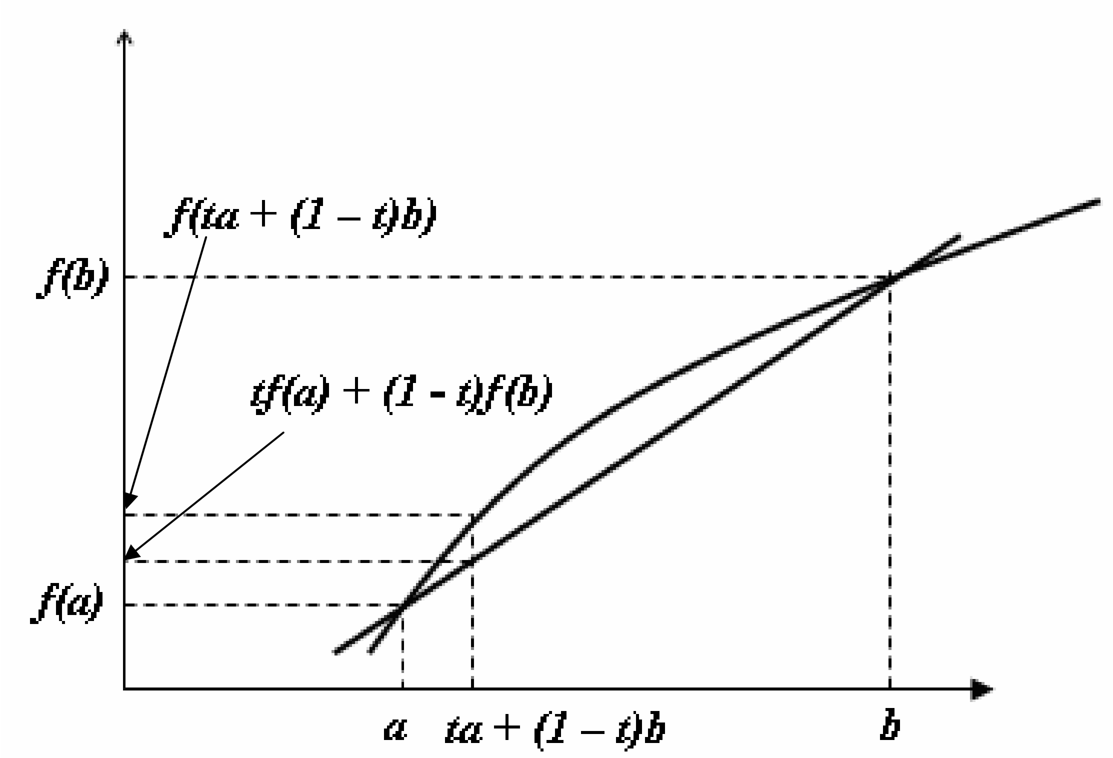
- A differentiable function f is **concave** on an interval if its derivative function f' is decreasing on that interval: a concave function has a decreasing slope.
- A function $f(x)$ is said to be **concave** on an interval if, for all a and b in that interval,

$$\forall t \in [0,1],\ f(ta + (1-t)b) \geq tf(a) + (1-t)f(b)$$

- Additionally, $f(x)$ is **strictly concave** if

$$\forall t \in [0,1],\ f(ta + (1-t)b) > tf(a) + (1-t)f(b)$$

A concave function



Testing for concavity of a single variable function

- A function is convex if its slope is non increasing or $\partial^2 f / \partial x^2 \leq 0$. It is strictly concave if its slope is continually decreasing or $\partial^2 f / \partial x^2 < 0$ throughout the function.

Properties of concave functions

- A continuous function on C is concave if and only if

$$f\left(\frac{a+b}{2}\right) \geq \frac{f(a)+f(b)}{2}$$

for any a and b in C.

- Equivalently, $f(x)$ is concave on $[a, b]$ if and only if the function $-f(x)$ is convex on every subinterval of $[a, b]$.

- If $f(x)$ is twice-differentiable, then $f(x)$ is concave if and only if $f''(x)$ is non-positive. If its second derivative is negative then it is strictly concave, but the opposite is not true, as shown by $f(x) = -x^4$.

Example

Consider the example in lecture notes 1 for a function of two variables. Locate the stationary points of $f(x) = 12x^5 - 45x^4 + 40x^3 + 5$

and find out if the function is convex, concave or neither at the points of optima based on the testing rules discussed above.

Solution

$$f'(x) = 60x^4 - 180x^3 + 120x^2 = 0$$

$$=> x^4 - 3x^3 + 2x^2 = 0$$

or $\quad x = 0, 1, 2$

Consider the point $x = x^* = 0$

$$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = 0 \text{ at } x^* = 0$$

$$f'''(x^*) = 720(x^*)^2 - 1080x^* + 240 = 240 \text{ at } x^* = 0$$

Example …contd.

Since the third derivative is non-zero $x = x^* = 0$ is neither a point of maximum or minimum but it is a point of inflection. Hence the function is neither convex nor concave at this point.

Consider $x = x^* = 1$

$$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = -60 \text{ at } x^* = 1$$

Since the second derivative is negative, the point $x = x^* = 1$ is a point of local maxima with a maximum value of $f(x) = 12 - 45 + 40 + 5 = 12$. At this point the function is concave since $\partial^2 f / \partial x^2 < 0$.

Example …contd.

Consider $x = x^* = 2$

$f''(x^*) = 240(x^*)^3 - 540(x^*)^2 + 240x^* = 240$ at $x^* = 2$

Since the second derivative is positive, the point $x = x^* = 2$ is a point of local minima with a minimum value of $f(x) = -11$. At this point the function is convex since $\partial^2 f / \partial x^2 > 0$.

Functions of two variables

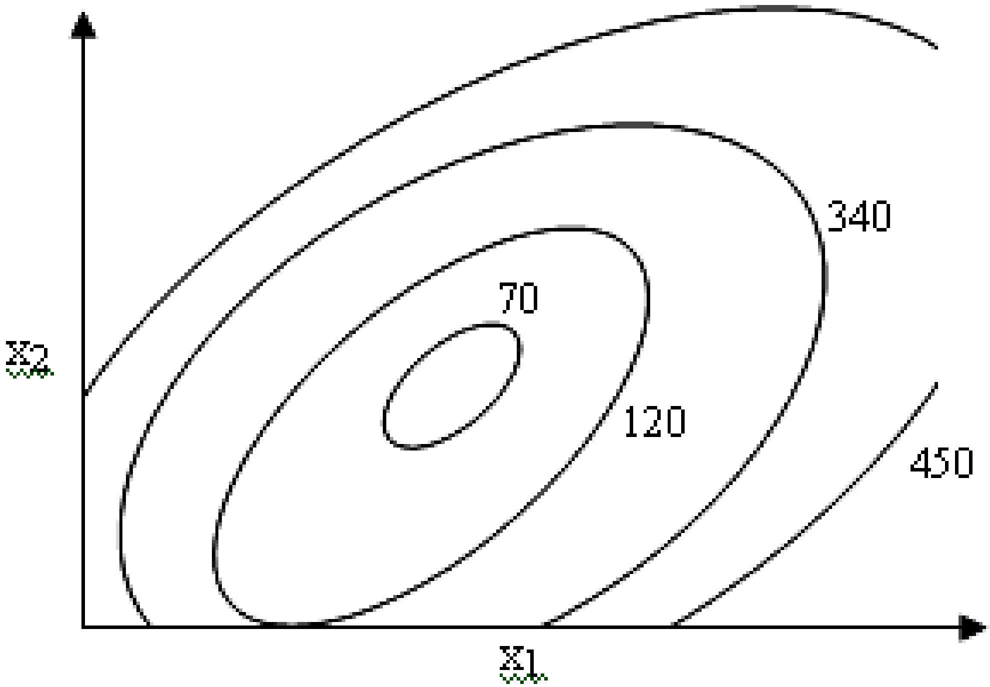
- A function of two variables, $f(\mathbf{X})$ where X is a vector = $[x_1, x_2]$, is strictly convex if

$$f(tX_1 + (1-t)X_2) < tf(X_1) + (1-t)f(X_2)$$

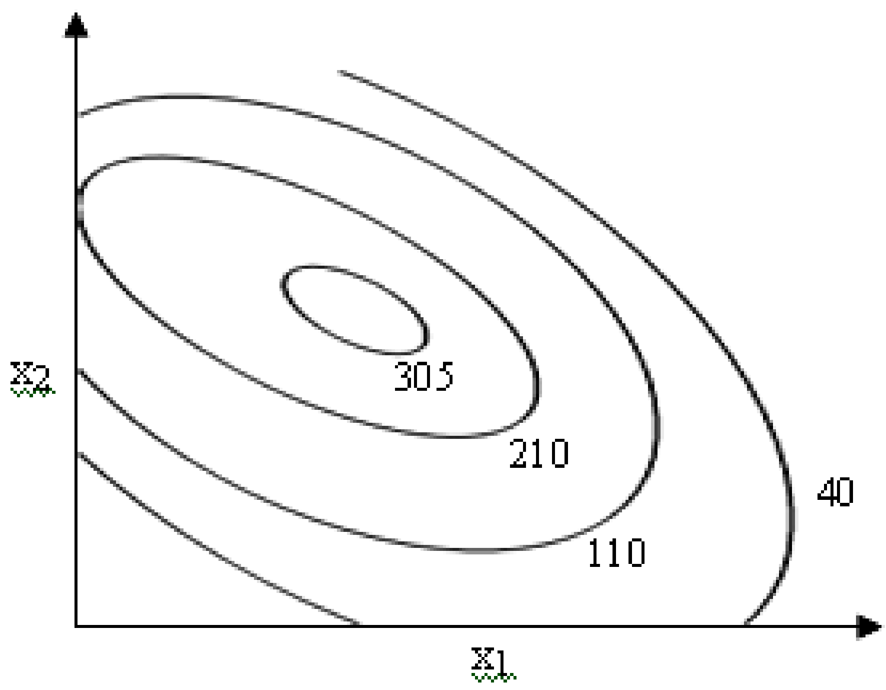
- where $\mathbf{X_1}$ and $\mathbf{X_2}$ are points located by the coordinates given in their respective vectors. Similarly a two variable function is strictly concave if

$$f(tX_1 + (1-t)X_2) > tf(X_1) + (1-t)f(X_2)$$

Contour plot of a convex function



Contour plot of a concave function



Sufficient conditions

- To determine convexity or concavity of a function of multiple variables, the eigen values of its Hessian matrix is examined and the following rules apply.
  - If all eigen values of the Hessian are positive the function is strictly convex.
  - If all eigen values of the Hessian are negative the function is strictly concave.
  - If some eigen values are positive and some are negative, or if some are zero, the function is neither strictly concave nor strictly convex.

Example

Consider the example in lecture notes 1 for a function of two variables. Locate the stationary points of *f*(X) and find out if the function is convex, concave or neither at the points of optima based on the rules discussed in this lecture.

$$f(\mathbf{X}) = 2x_1^3/3 - 2x_1x_2 - 5x_1 + 2x_2^2 + 4x_2 + 5$$

$$\Delta_x f = \begin{bmatrix} \frac{\partial f}{\partial x_1}(X^*) \\ \frac{\partial f}{\partial x_2}(X^*) \end{bmatrix} = \begin{bmatrix} 2x_1^2 - 2x_2 - 5 \\ -2x_1 + 4x_2 + 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$X_1 = [-1,-3/2]$ $X_2 = [3/2,-1/4]$

The Hessian of $f(\mathbf{X})$ is

$$\frac{\partial^2 f}{\partial x_1^{\,2}} = 4x_1 ; \frac{\partial^2 f}{\partial x_2^{\,2}} = 4; \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = -2$$

$$\mathbf{H} = \begin{bmatrix} 4x_1 & -2 \\ -2 & 4 \end{bmatrix}$$

$$|\lambda \mathbf{I} - \mathbf{H}| = \begin{vmatrix} \lambda - 4x_1 & 2 \\ 2 & \lambda - 4 \end{vmatrix}$$

At X_1

$$|\lambda \mathbf{I} - \mathbf{H}| = \begin{vmatrix} \lambda + 4 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda + 4)(\lambda - 4) - 4 = 0$$

$$\lambda^2 - 16 - 4 = 0$$

$$\lambda^2 = 12$$

$$\lambda_1 = +\sqrt{12} \qquad \lambda_2 = -\sqrt{12}$$

Since one eigenvalue is positive and one negative, X_1 is neither a relative maximum nor a relative minimum. Hence at X_1 the function is neither convex or concave.

Example (contd..)

At $X_2 = [3/2, -1/4]$

$$|\lambda \mathbf{I} - \mathbf{H}| = \begin{vmatrix} \lambda - 6 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda - 6)(\lambda - 4) - 4 = 0$$

$$\lambda^2 - 10\lambda + 20 = 0$$

$$\lambda_1 = 5 + \sqrt{5} \quad \lambda_2 = 5 - \sqrt{5}$$

Since both the eigenvalues are positive, X_2 is a local minimum, and the function is convex at this point as both the eigenvalues are positive.

Thank you