Convex functions

Neculai Andrei

Research Institute for Informatics,
Center for Advanced Modeling and Optimization,
8-10, Averescu Avenue, Bucharest 1, Romania,
E-mail: nandrei@ici.ro

Abstract. In this work we present the definition and the most important concepts and properties of convex functions. Convex functions play an important role in the study of optimization. These functions have many and important properties which can be used in developing suitable optimality conditions and computational schemes for optimization problems. One of the most important properties is that any local minimum of a convex function over a convex set is also a global minimum.
The subjects covered in this study include: definition and properties of convex functions, operations with convex functions, and first order and second order characterization of convexity. Finally, we present some important examples of convex functions.

1. Definition and properties of convex functions

Definition 1. *Let $S \subset R^n$ be a nonempty convex set. Function $f : S \to R$ is said to be convex on S if for any $x_1, x_2 \in S$ and all $0 \le \alpha \le 1$, we have*

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2).$$

If

$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha) f(x_2),$$

for all $x_1 \ne x_2$, then f is called a strictly convex function on S. If there is a constant $c > 0$ such that for any $x_1, x_2 \in S$

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) - \frac{1}{2}c\alpha(1-\alpha)\|x_1 - x_2\|^2,$$

then f is called a uniformly convex (or strongly convex) function on S.

The geometrical interpretation of convexity is very simple. For a convex function *the function values are below the corresponding chord*, that is, the values of convex function at points on the line segment $\alpha x_1 + (1-\alpha)x_2$ are less than or equal to the height of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

A function is convex if and only if it is convex when restricted to any line that intersects its domain. Rephrased, f is convex if and only if for all $x \in S$ and for all v, the function $h(t) = f(x + tv)$ is convex on $\{t : x + tv \in S\}$. This property is very useful in testing whether a function is convex by restricting it to a line.

If f is a convex (strictly convex) function, then $-f$ is said to be a concave (strictly concave) function.

It is important to extend a convex function to all R^n by defining its value to be ∞ outside its domain. If f is a convex function, then we define its *extended-valued extension* $\widetilde{f} : R^n \to R \cup \{\infty\}$ by

$$\widetilde{f}(x) = \begin{cases} f(x) & x \in S, \\ +\infty & x \notin S. \end{cases}$$

In the following we assume that all convex functions are implicitly extended. This allow us to express convexity as: for $0 < \alpha < 1$,

$$\widetilde{f}(\alpha x_1 + (1-\alpha)x_2) \le \alpha \widetilde{f}(x_1) + (1-\alpha)\widetilde{f}(x_2)$$

for *any* x_1 and x_2.

Any convex function can be described by an epigraph. Let $S \subset R^n$ be a nonempty set. The set $\{(x, f(x)) : x \in S\} \subset R^{n+1}$ is said to be the *graph of the function* f. The graph of any function f defines two sets in R^{n+1}: the epigraph, which consists of points above the graph of f and the hypograph, which consists of points below the graph of f.

Definition 2. *Let $S \subset R^n$ be a nonempty set, and $f : S \subset R^n \to R$. The epigraph of f is a subset of R^{n+1} defined by*

$$epif = \{(x,\alpha) : f(x) \le \alpha, x \in S, \alpha \in R\}.$$

The hypograph of f is a subset of R^{n+1} defined by

$$hypf = \{(x,\alpha) : f(x) \ge \alpha, x \in S, \alpha \in R\}.$$

Now we show that f is convex if and only if its epigraph is a convex set.

Theorem 1. *Let $S \subset R^n$ be a nonempty convex set, and $f : S \subset R^n \to R$. Then f is convex if and only if $epif$ is a convex set.*

Proof. Assume that f is convex. Let $x_1, x_2 \in S$ and (x_1,α_1), (x_2,α_2) be in $epif$. Then, for any $\theta \in (0,1)$ we have

$$f(\theta x_1 + (1-\theta)x_2) \le \theta f(x_1) + (1-\theta) f(x_2) \le \theta\alpha_1 + (1-\theta)\alpha_2.$$

Since S is a convex set, $\theta x_1 + (1-\theta)x_2 \in S$. Therefore,

$$(\theta x_1 + (1-\theta)x_2, \theta\alpha_1 + (1-\theta)\alpha_2) \in epif,$$

i.e. $epif$ is convex.

To show sufficiency, assume that $epif$ is convex, and consider $x_1, x_2 \in S$ and $(x_1, f(x_1))$, $(x_2, f(x_2)) \in epif$. From convexity of $epif$, for $\theta \in (0,1)$ we have that

$$(\theta x_1 + (1-\theta)x_2, \theta f(x_1) + (1-\theta) f(x_2)) \in epif.$$

Therefore,

$$f(\theta x_1 + (1-\theta)x_2) \le \theta f(x_1) + (1-\theta) f(x_2),$$

for each $\theta \in (0,1)$, i.e. f is convex. ♦

The epigraph of a function f is in close connection with the lower semi-continuity of f. Both these concepts $epif$ and lower semi-continuity are key concepts in mathematical programming.

Definition 3. *A function f is lower semi-continuous if, for each $x, y \in R^n$,*

$$\liminf_{y \to x} f(y) \geq f(x).$$

Theorem 2. *Let $f : R^n \to R \cup \{\infty\}$. The following statements are equivalent:*

1) f is lower semi-continuous on R^n;

2) $epif$ is a closed set in $R^n \times R$;

3) *the level sets $L_t(f) = \{x \in R^n : f(x) \leq t, t \in R\}$ are closed for all $t \in R$.*

Proof. (1) $\Rightarrow$ (2): Let $(y_k, t_k) \in epif$ a sequence converging to (x,t) for $k \to \infty$. Since $f(y_k) \leq t_k$ for all k, it follows that $t = \lim_{k\to\infty} t_k \geq \liminf_{y_k \to x} f(y_k) \geq f(x)$, i.e. $(x,t) \in epif$.

(2) $\Rightarrow$ (3): The level set $L_t(f)$ is the intersection of two closed sets: $epif$ and $(R^n \times \{t\})$. Obviously the intersection is closed.

(3) $\Rightarrow$ (1): Now suppose that f is not lower semi-continuous at some x, which means there exists a sequence $\{y_k\}$ converging to x such that $f(y_k)$ converges to $r < f(x) \leq +\infty$. Consider $t \in (r, f(x))$. When k is large enough, we have $f(y_k) \leq t < f(x)$, i.e. $L_t(f)$ does not contain its limit x. Therefore, $L_t(f)$ is not closed. ♦

With this the following definition of closed function can be presented.

Definition 4. *A function $f : R^n \to R \cup \{\infty\}$ is said to be closed if it is lower semi-continuous everywhere, or its epigraph is closed, or if its level sers are closed.*

Now, we will introduce the concept of monotone function, which is very useful for characterization of a convex function with monotonicity.

Definition 5. *Let $F : S \subset R^n \to R^n$ and $S_0 \subset S$. Then*

1) F is monotone on S if for any $x, y \in S_0$,

$$(F(x) - F(y))^T (x - y) \geq 0.$$

2) F is strictly monotone on S if for any $x, y \in S_0$, $x \neq y$,

$$(F(x) - F(y))^T (x - y) > 0.$$

3) F is uniformly monotone (or strongly monotone) on S if for any $x, y \in S_0$, there is a constant $c > 0$ so that

$$(F(x) - F(y))^T (x - y) \geq c\|x - y\|^2.$$

The following theorem presents a result concerning the level set L_α of a convex function. It is shown that for any α real, L_α is a convex set.

Theorem 3. *Let $S \subset R^n$ be a nonempty convex set, $f : S \to R$ a convex function and α a real number. Then the level set $L_\alpha = \{x : f(x) \leq \alpha\}$ is a convex set.*

Proof. Let us consider two points $x_1, x_2 \in L_\alpha$. Then $x_1, x_2 \in S$ and $f(x_1) \le \alpha$, $f(x_2) \le \alpha$. Consider the point $x = \lambda x_1 + (1-\lambda)x_2$, where $\lambda \in (0,1)$. From convexity of S it follows that $x \in S$. Since f is convex we have

$$f(x) \le \lambda f(x_1) + (1-\lambda) f(x_2) \le \lambda\alpha + (1-\lambda)\alpha = \alpha.$$

Therefore, $x \in L_\alpha$, i.e. L_α is a convex set. ♦

2. Operations with convex functions

We describe some operations that preserve convexity or concavity of functions or give us the possibility to construct new convex and concave functions. Firstly we consider the simple operations such as scaling, addition and pointwise suppremum, and then consider the continuity.

Theorem 4. *a) Let f be a convex function on a convex set $S \subset R^n$ and $\alpha \ge 0$ a real number, then αf is a convex function on S.*
b) Let f_1 and f_2 be convex functions on a convex set S, then $f_1 + f_2$ is a convex function on S.
c) Let f_i, $(i = 1,\dots,m)$ be convex functions on a convex set S and real nonnegative numbers $\alpha_i \ge 0$, $(i = 1,\dots,m)$, then $\sum_{i=1}^m \alpha_i f_i$ is a convex function on S.

Proof. We proof only b). Indeed, let $x_1, x_2 \in S$ and $0 < \alpha < 1$, then

$$f_1(\alpha x_1 + (1-\alpha)x_2) + f_2(\alpha x_1 + (1-\alpha)x_2) \le \alpha\left[f_1(x_1) + f_2(x_1)\right] + (1-\alpha)\left[f_1(x_2) + f_2(x_2)\right].$$ ♦

Theorem 5. *Let f_i, $(i = 1,\dots,m)$ be convex functions on a convex set S, then their pointwise maximum $f(x) = \max\{f_1(x),\dots, f_m(x)\}$ is also a convex function.*

Proof. For simplicity, let us consider $m = 2$. If $0 \le \alpha \le 1$, then

$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= \max\{f_1(\alpha x_1 + (1-\alpha)x_2), f_2(\alpha x_1 + (1-\alpha)x_2)\} \\ &\le \max\{\alpha f_1(x_1) + (1-\alpha) f_1(x_2), \alpha f_2(x_1) + (1-\alpha) f_2(x_2)\} \\ &\le \alpha \max\{f_1(x_1), f_2(x_1)\} + (1-\alpha)\max\{f_1(x_2), f_2(x_2)\} \\ &= \alpha f(x_1) + (1-\alpha) f(x_2), \end{aligned}$$

which prove the convexity of f. ♦

The continuity of a convex function is an important concept. If a convex function whose domain of definition is not open is continuous is not sure. We know that a convex function is continuous on the relative interior of its domain; it can have discontinuities only on its relative boundary. The following theorem shows the main results on continuity of convex functions.

Theorem 6. *Let $S \subset D \subset R^n$ be an open convex set and $f : D \to R$ be convex. Then f is continuous on S.*

Proof. Let us consider an arbitrary point $x_0 \in S$. Since S is an open convex set, we can find $n+1$ points $x_1,\dots,x_{n+1} \in S$ such that the interior of the convex hull

$$C = \left\{ x : x = \sum_{i=1}^{n+1} a_i x_i ,\ a_i \geq 0, \sum_{i=1}^{n+1} a_i = 1 \right\}$$

is not empty and $x_0 \in \operatorname{int} C$. Now, let $a = \max_{1 \leq i \leq n+1} f(x_i)$. Then, for any $x \in C$,

$$f(x) = f\left(\sum_{i=1}^{n+1} a_i x_i \right) \leq \sum_{i=1}^{n+1} a_i f(x_i) \leq a.$$

Therefore f is bounded over C.

On the other hand, since $x_0 \in \operatorname{int} C$, there is a $\delta > 0$ such that $B(x_0, \delta) \subset C$, where $B(x_0, \delta) = \{x : \|x - x_0\| \leq \delta\}$. Hence for arbitrary $h \in B(0, \delta)$ and $\lambda \in [0,1]$, we have the following representation of x_0:

$$x_0 = \frac{1}{1+\lambda}(x_0 + \lambda h) + \frac{\lambda}{1+\lambda}(x_0 - h).$$

Since f is convex on C, it follows that

$$f(x_0) \leq \frac{1}{1+\lambda} f(x_0 + \lambda h) + \frac{\lambda}{1+\lambda} f(x_0 - h).$$

Therefore,

$$f(x_0 + \lambda h) - f(x_0) \geq \lambda\big(f(x_0) - f(x_0 - h)\big) \geq -\lambda(a - f(x_0)).$$

On the other hand,

$$f(x_0 + \lambda h) = f(\lambda(x_0 + h) + (1-\lambda)x_0) \leq \lambda f(x_0 + h) + (1-\lambda) f(x_0),$$

i.e.

$$f(x_0 + \lambda h) - f(x_0) \leq \lambda\big(f(x_0 + h) - f(x_0)\big) \leq \lambda(a - f(x_0)).$$

Therefore, from the above inequalities we get

$$|f(x_0 + \lambda h) - f(x_0)| \leq \lambda |f(x_0) - a|.$$

Now, for given $\varepsilon > 0$, select $\delta' \leq \delta$ such that $\delta' |f(x_0) - a| \leq \varepsilon \delta$. Consider $d = \lambda h$ with $\|h\| = \delta$, then $d \in B(0, \delta)$ and $|f(x_0 + d) - f(x_0)| \leq \varepsilon$, which prove the theorem. ♦

3. First order condition

The following theorem gives the first order condition of differential convex functions.

Theorem 7. *Let $S \subset R^n$ be a nonempty open convex set and $f : S \to R$ be a differentiable function. Then:*

1) f is convex if and only if, for any $x, y \in S$:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x). \tag{1.1}$$

2) f is strictly convex on S if and only if, for any $x, y \in S$, with $y \neq x$:

$$f(y) > f(x) + \nabla f(x)^T (y - x). \tag{1.2}$$

3) f is strongly convex (or uniformly convex) on S if and only if, for any $x, y \in S$:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} c \|y - x\|^2, \tag{1.3}$$

where $c > 0$ is a constant.

Proof. *Necessity*: Consider $f(x)$ a convex function, then for all $0 < \alpha < 1$ we have

$$f(\alpha y + (1-\alpha)x) \le \alpha f(y) + (1-\alpha) f(x).$$

Therefore,

$$\frac{f(x+\alpha(y-x)) - f(x)}{\alpha} \le f(y) - f(x).$$

Now, setting $\alpha \to 0$ yields

$$\nabla f(x)^T (y-x) \le f(y) - f(x).$$

Sufficiency: Assume that (1.1) holds. Select two arbitrary points $x_1, x_2 \in S$ and consider their convex combination $x = \alpha x_1 + (1-\alpha)x_2$, where $0 < \alpha < 1$. Then,

$$f(x_1) \ge f(x) + \nabla f(x)^T (x_1 - x),$$
$$f(x_2) \ge f(x) + \nabla f(x)^T (x_2 - x).$$

Therefore

$$\begin{aligned}\alpha f(x_1) + (1-\alpha) f(x_2) &\ge f(x) + \nabla f(x)^T (\alpha x_1 + (1-\alpha)x_2 - x) \\ &= f(\alpha x_1 - (1-\alpha)x_2),\end{aligned}$$

i.e. $f(x)$ is a convex function.

In a similar manner we can prove (1.2) and (1.3). For example, to get (1.3), it is enough to apply (1.1) to the function $f - 1/2(c\|.\|^2)$. ♦

The geometrical interpretation of theorem 4 is that in any point the linear approximation based on a local derivative is a lower estimate of the function, i.e. the convex function always lies above its tangent at any point. Such a tangent is called a *supporting hyperplane* of the convex function.

Theorem 8. *Assume that $f : S \subset R^n \to R$ is differentiable on the convex set S. Then*
1) f is convex on S if and only if is its gradient ∇f is monotone, i.e.

$$(\nabla f(x) - \nabla f(y))^T (x-y) \ge 0, \tag{1.4}$$

for any $x, y \in S$.
2) f is strictly convex on S if and only if is its gradient ∇f is strictly monotone, i.e.

$$(\nabla f(x) - \nabla f(y))^T (x-y) > 0, \tag{1.5}$$

for any $x, y \in S$, $x \ne y$.
3) f is uniformly convex (or strongly convex) on S if and only if is its gradient ∇f is uniformly monotone, i.e.

$$(\nabla f(x) - \nabla f(y))^T (x-y) \ge c\|x-y\|^2, \tag{1.6}$$

for any $x, y \in S$, and where $c > 0$ is the constant from definition 5.

Proof. *Necessity.* Assume that f is uniformly convex on S, then from theorem 4, for any $x, y \in S$, we have

$$f(y) \ge f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}c\|y-x\|^2, \tag{1.7}$$

$$f(x) \ge f(y) + \nabla f(y)^T (x-y) + \frac{1}{2}c\|x-y\|^2. \tag{1.8}$$

Adding these two inequalities we get (1.6). It is very simple to see that if f is convex, then (1.7) and (1.8) hold with $c = 0$. Therefore (1.4) holds. If f is strictly convex, then (1.7) and 91.8) hold with $c = 0$, but with strict inequality for any $x \neq y$. Hence, we get (1.5).

Sufficiency. Suppose now that ∇f is monotone. For any fixed $x, y \in S$, by the mean value thorem we have

$$f(y) - f(x) = \nabla f(z)^T (y - x), \tag{1.9}$$

where $z = x + t(y - x)$, $t \in (0,1)$. From (1.4) it follows that

$$(\nabla f(z) - \nabla f(x))^T (y - x) = \frac{1}{t}[\nabla f(z) - \nabla f(x)]^T (z - x) \geq 0,$$

which together with (1.9) gives

$$\begin{aligned} f(y) - f(x) &= (\nabla f(z) - \nabla f(x))^T (y - x) + \nabla f(x)^T (y - x) \\ &\geq \nabla f(x)^T (y - x). \end{aligned} \tag{1.10}$$

This inequality, by theorem 4, shows that f is convex.

Now, if (1.5) holds, than (1.10) is true with strict inequality and $x \neq y$. Therefore, it follows that f is strictly convex.

Finally, suppose that (1.6) is true and consider the function

$$h(t) = f(x + t(y - x)) = f(u),$$

where $u = x + t(y - x)$, $t \in (0,1)$. But, $h'(t) = \nabla f(u)^T (y - x)$ and $h'(0) = \nabla f(x)^T (y - x)$. Then (1.6) means

$$\begin{aligned} h'(t) - h'(0) &= (\nabla f(u) - \nabla f(x))^T (y - x) = \frac{1}{t}(\nabla f(u) - \nabla f(x))^T (u - x) \\ &\geq \frac{1}{t} c \|u - x\|^2 = tc \|y - x\|^2. \end{aligned}$$

Therefore,

$$h(1) - h(0) - h'(0) = \int_0^1 [h'(t) - h'(0)] dt \geq \frac{1}{2} c \|y - x\|^2,$$

which shows that

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} c \|y - x\|^2. \ ♦$$

4. Second order condition

For the twice continuously differentiable convex functions the following characterization can be given.

Theorem 9. *Let $S \subset R^n$ be a nonempty open convex set and $f : S \to R$ be a twice continuously differentiable function. Then*

1) f is convex if and only if its Hessian matrix is positive semidefinite at each point in S.

2) f is strictly convex if its Hessian matrix is positive definite at each point in S.

3) f is uniformly convex if and only if its Hessian matrix is uniformly positive definite at each point in S.

Proof. We will prove only 1), the other cases are proved in a similar manner.

Necessity. Suppose that f is a convex function, and consider a point $y \in S$. We must prove that for any $p \in R^n$, $p^T \nabla^2 f(y) p \geq 0$. But, by theorem 4 we have

$$f(y + \lambda p) \geq f(y) + \lambda \nabla f(y)^T p.$$

Since $f(x)$ is twice continuously differentiable at y, it follows that

$$f(y + \lambda p) = f(y) + \lambda \nabla f(y)^T p + \frac{1}{2}\lambda^2 p^T \nabla^2 f(y) p + o\left(\|\lambda p\|^2\right).$$

Now, introducing (1.4) in (1.3) we get

$$\frac{1}{2}\lambda^2 p^T \nabla^2 f(y) p + o\left(\|\lambda p\|^2\right) \geq 0.$$

Dividing by λ^2 and letting $\lambda \to 0$, it follows that

$$p^T \nabla^2 f(y) p \geq 0.$$

Sufficiency. Suppose that the Hessian matrix $\nabla^2 f(x)$ is positive semidefinite at each point $x \in S$. Consider the points $x, y \in S$. Using the mean value theorem, we can write

$$f(x) = f(y) + \nabla f(y)^T (x - y) + \frac{1}{2}(x - y)^T \nabla^2 f(z)(x - y),$$

where z is a point on the segment line connecting x and y, i.e. $z = y + \theta(x - y)$, $\theta \in (0,1)$. Since S is convex, $z \in S$. From the assumption it follows that

$$f(x) \geq f(y) + \nabla f(y)^T (x - y),$$

i.e. the function f is a convex function by theorem 4. ♦

Now, combining this theorem with theorem 5 we obtain the following result

Theorem 10. *Let $S \subset R^n$ be a nonempty open set and $f : S \to R$ be a twice continuously differentiable function on S. Then*

1) $\nabla f(x)$ is monotone on S if and only if $\nabla^2 f(x)$ is positive semidefinite for all $x \in S$.

2) If $\nabla^2 f(x)$ is positive definite for all $x \in S$, then $\nabla f(x)$ is strictly monotone on S.

3) $\nabla f(x)$ is uniformly monotone (or strongly monotone) on S if and only if $\nabla^2 f(x)$ is uniformly positive definite, i.e. there exists a number $c > 0$ so that for any $x \in S$ and $p \in R^n$,

$$p^T \nabla^2 f(x) p \geq c\|p\|^2. \; ♦$$

In theorem 3 we proved that the level sets of a convex function are convex sets. From theorem 2 it follows that if f is a continuously convex function, then the level set L_α is a closed convex set. Furthermore, we can prove the following theorem

Theorem 11. *Let $S \subset R^n$ be a nonempty convex set, and $f(x) : S \subset R^n \to R$ be a twice continuously differential function on S. Suppose that there exists a positive number $m > 0$ such that for any $x \in L(x_0)$ and $u \in R^n$*

$$u^T \nabla^2 f(x) u \geq m\|u\|^2. \tag{1.11}$$

Then the level set $L(x_0) = \{x \in S : f(x) \leq f(x_0)\}$ is a bounded closed convex set.

Proof. By theorem x we know that f is convex on $L(x_0)$. Therefore, by theorem 3 it follows that $L(x_0)$ is convex. Observe that if $f(x)$ is continuous, then $L(x_0)$ is a closed convex set for all $x_0 \in R^n$. Now, let us prove the boundedness of $L(x_0)$. Consider two arbitrary points $x, y \in L(x_0)$. Since $L(x_0)$ is convex, from (1.11) we have

$$m\|y-x\|^2 \le (y-x)^T \nabla^2 f(x+\alpha(y-x))(y-x).$$

By differentiability we have

$$\begin{aligned} f(y) &= f(x) + \nabla f(x)^T (y-x) \\ &\quad + \int_0^1\int_0^t (y-x)^T \nabla^2 f(x+\alpha(y-x))(y-x)\,d\alpha dt \\ &\ge f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}m\|y-x\|^2, \end{aligned}$$

where m is independent of x and y. Therefore, for arbitrary $y \in L(x_0)$ and $y \ne x_0$ we have

$$\begin{aligned} f(y) - f(x_0) &\ge \nabla f(x_0)^T (y-x_0) + \frac{1}{2}m\|y-x_0\|^2 \\ &\ge -\|\nabla f(x_0)\|\,\|y-x_0\| + \frac{1}{2}m\|y-x_0\|^2. \end{aligned}$$

Since $f(y) \le f(x_0)$, from the above inequality we get

$$\|y-x_0\| \le \frac{2}{m}\|\nabla f(x_0)\|,$$

which prove that the level set $L(x_0)$ is bounded. ♦

5. Examples of convex functions

Convexity or concavity of a function can be verified using different techniques, such as directly by means of definition 1, verifying that its Hessian is positive semidefinite, or restricting the function to an arbitrary line and then verifying convexity of the resulting function of one variable. In the follows we present some example of convex or concave functions.

Indicator function. Let $S \subset R^n$ be a nonempty set. The indicator function $I_S : R^n \to R \cup \{\infty\}$ is defined as

$$I_S(x) = \begin{cases} 0 & \text{if } x \in S, \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly, I_S is convex if and only if S is convex.

Support function of a set. Let $S \subset R^n$ be a nonempty set. The support function of S is defined as

$$\sigma_S(s) = \sup\{s^T x : x \in S\},$$

which is a convex function.

Linear functions. Any linear function $f(x) = a^T x + b$ is both convex and concave function on R^n, where $a, x \in R^n$ and $b \in R$.

Quadratic functions. Consider the quadratic function $f : R^n \to R$, given by

$$f(x) = x^T Qx + 2q^T x + p,$$

where $P \in R^{n\times n}$ is a symmetric matrix, $q \in R^n$ and $p \in R.$ Since $\nabla^2 f(x) = 2P$ for all x, f is convex if and only if P is positive semidefinite $(P \geq 0)$. f is concave if and only if P is negative semidefinite $(P \leq 0)$. For quadratic functions, strict convexity is easily characterized as: f is strictly convex if and only if P is positive definite $(P > 0)$. f is strictly concave if and only if P is negative definite $(P < 0)$.

Least-squares functions. Function $f(x) = \|Ax - b\|_2^2$ is convex for any A. Indeed, we have $\nabla f(x) = 2A^T(Ax - b)$ and $\nabla^2 f(x) = 2A^T A.$

Quadratic over linear. Function $f(x) = x^2 / y$ is convex for $y > 0$. We have

$$\nabla^2 f(x,y) = \frac{2}{y^3}\begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix} = \frac{2}{y^3}\begin{bmatrix} y \\ -x \end{bmatrix}\begin{bmatrix} y \\ -x \end{bmatrix}^T \geq 0.$$

Exponential. Function e^{ax} is convex on $R,$ for any $a \in R.$

Power. Function x^a is convex on R_{++} when $a \geq 1$ or $a \leq 0,$ and concave for $0 \leq a \leq 1.$

Logarithm. Function $\log x$ is concave on R_{++}.

Negative entropy $x \log x$**.** If $f(x) = x\log x$, defined on R_{++}, then $f'(x) = \log x + 1$, $f''(x) = 1/x,$ and $f''(x) > 0$ for $x > 0$. Therefore, the negative entropy is a strictly convex function for all $x > 0$.

Norms. Every norm on R^n is a convex function. If $f : R^n \to R$ is a norm and $0 \leq \alpha \leq 1$, then $f(\alpha x_1 + (1-\alpha)x_2) \leq f(\alpha x_1) + f((1-\alpha)x_2) = \alpha f(x_1) + (1-\alpha) f(x_2),$ since by definition a norm is a homogeneous function and satisfies the triangle inequality.

Geometric mean. The geometric mean $f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ is a concave function on $R^n_{++}.$ Indeed, its Hessian $\nabla^2 f(x)$ is given by

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k^2}, \quad \frac{\partial^2 f(x)}{\partial x_k \partial x_j} = \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k x_j} \quad (k \neq j),$$

and

$$\nabla^2 f(x) = -\frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2}\left(n diag\left(\frac{1}{x_1^2},\cdots,\frac{1}{x_n^2}\right) - pp^T \right),$$

where $p = [1/x_1,\ldots,1/x_n]^T$. Observe that, for all v

$$v^T \nabla^2 f(x) v = -\frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2}\left(n\sum_{i=1}^n v_i^2 / x_i^2 - \left(\sum_{i=1}^n v_i / x_i\right)^2 \right) \leq 0,$$

which follows from the Cauchy-Schwarz inequality.

Log-Sum-Exp. Function $f(x) = \log \sum_{i=1}^{n} \exp(x_i)$ is convex. The Hessian of log-sum-exp function is

$$\nabla^2 f(x) = \frac{1}{(e^T z)^2}\left((e^T z) diag(z) - zz^T\right),$$

where $e = [1,\ldots,1]^T$ and $z = [\exp(x_1),\ldots,\exp(x_n)]$. To verify that $\nabla^2 f(x) \geq 0$ we must show that $v^T \nabla^2 f(x) v \geq 0$ for all v. But

$$v^T \nabla^2 f(x) v = \frac{1}{(e^T z)^2}\left(\left(\sum_{i=1}^{n} z_i\right)\left(\sum_{i=1}^{n} v_i^2 z_i\right) - \left(\sum_{i=1}^{n} v_i z_i\right)^2\right) \geq 0$$

by Cauchy-Schwarz inequality.

References

Andrei, N., Programarea Matematică Avansată. Teorie, Metode Computaționale, Aplicații. Editura Tehnică, Bucureşti, 1999.

Bazaraa, M.S., Sherali, H.D. and Shety, C.M., *Nonlinear Programming. Theory and Algorithms.* John Wiley & Sons, I., 1993.

Boyd. S. and Vandenberghe, L., *Convex Optimization.* Cambridge University Press, 2003.

Wenyu Sun and Ya-xiang Yuan, *Optimization Theory and Methods.* Electronic book. 2003.

June 15, 2005