## 1.2 Useful Properties of Convex Functions

We have already mentioned that convex functions are tractable in optimization (or minimization) problems and this is mainly because of the following properties:

1. Local optimality (or minimality) guarantees global optimality;
2. Duality such as min-max relation and separation theorem holds good.

This section is to give more specific descriptions of these properties, and to discuss their possible versions for discrete functions.

Let us first recall the definition of a convex function. A function $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is said to be *convex* if

$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y) \tag{1.2}$$

for all $x, y \in \mathbf{R}^n$ and for all $\lambda$ with $0 \leq \lambda \leq 1$, where it is understood that the inequality is satisfied if $f(x)$ or $f(y)$ is equal to $+\infty$. The inequality (1.2) implies that the set

$$S = \{x \in \mathbf{R}^n \mid f(x) < +\infty\},$$

called the *effective domain* of $f$, is a convex set. Hence the present definition of a convex function coincides with the one in (1.1) that makes an explicit reference to the effective domain $S$. A special case of inequality (1.2) for $\lambda = 1/2$ yields the *midpoint convexity*

$$\frac{f(x)+f(y)}{2} \geq f\left(\frac{x+y}{2}\right) \qquad (x, y \in \mathbf{R}^n), \tag{1.3}$$

and, conversely, this implies convexity, provided $f$ is continuous. We often assume (explicitly or implicitly) that $f(x) < +\infty$ for some $x \in \mathbf{R}^n$ whenever we talk about a convex function $f$. A function $h : \mathbf{R}^n \to \mathbf{R} \cup \{-\infty\}$ is said to be *concave* if $-h$ is convex.

A point (or vector) $x$ is said to be a *global optimum* of $f$ if the inequality

$$f(x) \leq f(y) \tag{1.4}$$

holds for every $y$, and $x$ is a *local optimum* if this inequality holds for every $y$ in some neighborhood of $x$. Obviously, global optimality implies local optimality. The converse is not true in general, but it is true for convex functions.

**Theorem 1.1.** *For a convex function, global optimality* (*or minimality*) *is guaranteed by local optimality.*

***Proof.*** Let $x$ be a local optimum of a convex function $f$. Then we have $f(z) \geq f(x)$ for any $z$ in some neighborhood $U$ of $x$. For any $y$, $z = \lambda x + (1-\lambda) y$ belongs to $U$ for $\lambda < 1$ sufficiently close to 1, and it follows from (1.2) that

$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda) y) = f(z) \geq f(x).$$

This implies $f(y) \geq f(x)$. □

The above theorem is significant and useful in that it reduces the global property to a local one. Still it refers to an infinite number of points or directions around $x$ for the local optimality. In considering discrete structures on top of convexity we may hope that a fixed and finite set of directions suffices to guarantee the local optimality. For example, in the simplest case of a *separable convex function*

$$f(x) = \sum_{i=1}^{n} f_i(x(i)), \tag{1.5}$$

which is the sum of univariate convex functions[5) ] $f_i(x(i))$ in each component of $x = (x(i) \mid i = 1, \ldots, n)$, it suffices to check for local optimality in $2n$ directions, positive and negative directions of the coordinate axes. Such phenomenon of "discreteness in direction", so to speak, is a reflection of the combinatorial structure of separable convex functions. Although the combinatorial structure of separable convex functions is too simple for further serious considerations, similar phenomena of "discreteness in direction" occur in nontrivial ways for L-convex or M-convex functions, as we see in §1.4.

We now go on to the second issue of duality and conjugacy. For a function $f$ (not necessarily convex), the *convex conjugate* $f^{\bullet} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ is defined by

$$f^{\bullet}(p) = \sup\{\langle p, x\rangle - f(x) \mid x \in \mathbf{R}^n\} \qquad (p \in \mathbf{R}^n), \tag{1.6}$$

where

$$\langle p, x\rangle = \sum_{i=1}^{n} p(i)x(i) \tag{1.7}$$

for $p = (p(i) \mid i = 1, \ldots, n)$ and $x = (x(i) \mid i = 1, \ldots, n)$. The function $f^{\bullet}$ is also referred to as the (convex) *Legendre–Fenchel transform* of $f$, and the mapping $f \mapsto f^{\bullet}$ as the (convex) *Legendre–Fenchel transformation*.

For example, for $f(x) = \exp(x)$, where $n = 1$, we see

$$f^{\bullet}(p) = \begin{cases} p \log p - p & (p > 0) \\ 0 & (p = 0) \\ +\infty & (p < 0) \end{cases}$$

by a simple calculation. See Fig. 1.3 for the geometric meaning in the case of $n = 1$.

The Legendre–Fenchel transformation gives a one-to-one correspondence in the class of well-behaved convex functions, called "closed proper convex functions," where the precise meaning of this technical terminology (not important here) will be explained later in §3.1. Notation $f^{\bullet\bullet}$ means $(f^{\bullet})^{\bullet}$, the conjugate of the conjugate function of $f$.

**Theorem 1.2 (Conjugacy).** *The Legendre–Fenchel transformation $f \mapsto f^{\bullet}$ gives a symmetric one-to-one correspondence in the class of all closed proper convex functions. That is, for a closed proper convex function $f$, $f^{\bullet}$ is a closed proper convex function and $f^{\bullet\bullet} = f$.*

5) A *univariate function* means a function in a single variable.

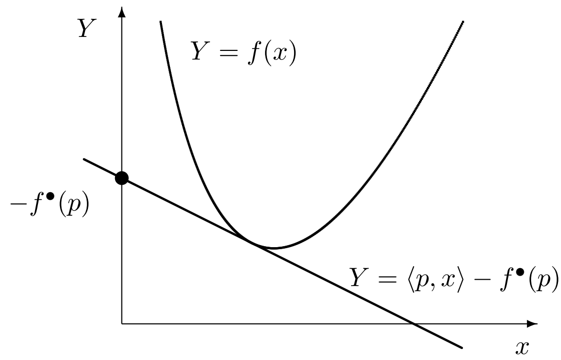


**Figure 1.3.** *Conjugate function* (*Legendre–Fenchel transform*)

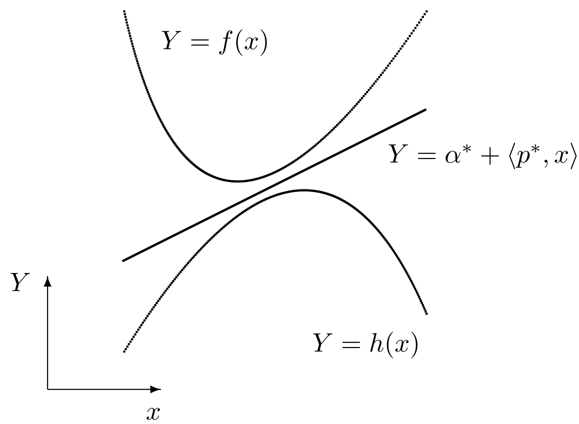


**Figure 1.4.** *Separation for convex and concave functions*

Similarly, for a function $h$ the *concave conjugate* $h^\circ : \mathbf{R}^n \to \mathbf{R} \cup \{-\infty\}$ is defined by

$$h^\circ(p) = \inf\{\langle p, x \rangle - h(x) \mid x \in \mathbf{R}^n\} \qquad (p \in \mathbf{R}^n). \tag{1.8}$$

The *duality* principle in convex analysis can be expressed in a number of different forms. One of the most appealing statements is in the form of the separation theorem, which asserts the existence of a separating affine function $Y = \alpha^* + \langle p^*, x \rangle$ for a pair of convex and concave functions (see Fig. 1.4).

**Theorem 1.3 (Separation theorem).** *Let $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ and $h : \mathbf{R}^n \to \mathbf{R} \cup \{-\infty\}$ be convex and concave functions, respectively* (*satisfying certain* *regularity conditions*). *If*[6)]

$$f(x) \geq h(x) \qquad (\forall x \in \mathbf{R}^n),$$

*there exist* $\alpha^* \in \mathbf{R}$ *and* $p^* \in \mathbf{R}^n$ *such that*

$$f(x) \geq \alpha^* + \langle p^*, x \rangle \geq h(x) \qquad (\forall x \in \mathbf{R}^n).$$

It is admitted that the statement above is mathematically incomplete, referring to "certain regularity conditions," which will be specified later in §3.1.

Another expression of the duality principle is in the form of the Fenchel duality. This is a min-max relation between a pair of convex and concave functions and their conjugate functions. The "certain regularity conditions" in the statement below will be specified later.

**Theorem 1.4 (Fenchel duality).** *Let* $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ *and* $h : \mathbf{R}^n \to \mathbf{R} \cup \{-\infty\}$ *be convex and concave functions, respectively* (*satisfying certain regularity conditions*). *Then*

$$\min\{f(x) - h(x) \mid x \in \mathbf{R}^n\} = \max\{h^\circ(p) - f^\bullet(p) \mid p \in \mathbf{R}^n\}.$$

Such a min-max theorem is computationally useful in that it affords a *certificate of optimality*. Suppose that we want to minimize $f(x) - h(x)$ and have $x = \hat{x}$ as a candidate for the minimizer. How can we verify or prove that $\hat{x}$ is indeed an optimal solution? One possible way is to demonstrate a vector $\hat{p}$ such that $f(\hat{x}) - h(\hat{x}) = h^\circ(\hat{p}) - f^\bullet(\hat{p})$. This implies the optimality of $\hat{x}$ by virtue of the min-max theorem. The vector $\hat{p}$, often called a dual optimal solution, serves as a certificate for the optimality of $\hat{x}$. It is emphasized that the min-max theorem guarantees the existence of such a certificate $\hat{p}$ for any optimal solution $\hat{x}$. It is also mentioned that the min-max theorem does not tell us how to find optimal solutions $\hat{x}$ and $\hat{p}$.

It is one of the recurrent themes in discrete convexity how the conjugacy and the duality above should be adapted in discrete settings. To be specific, let us consider integer-valued functions on integer lattice points, and discuss possible notions of conjugacy and duality for $f : \mathbf{Z}^n \to \mathbf{Z} \cup \{+\infty\}$ and $h : \mathbf{Z}^n \to \mathbf{Z} \cup \{-\infty\}$. Some ingredients of discreteness (integrality) are naturally expected in the formulation of conjugacy and duality. This amounts to discussing another kind of discreteness, "discreteness in value" so to speak, in contrast to "discreteness in direction" mentioned above.

Discrete versions of the Legendre–Fenchel transformations can be defined by

$$f^\bullet(p) = \sup\{\langle p, x \rangle - f(x) \mid x \in \mathbf{Z}^n\} \qquad (p \in \mathbf{Z}^n), \tag{1.9}$$

$$h^\circ(p) = \inf\{\langle p, x \rangle - h(x) \mid x \in \mathbf{Z}^n\} \qquad (p \in \mathbf{Z}^n). \tag{1.10}$$

6) Notation $\forall$ means "for all", "for any," or "for each."

They are meaningful as transformations of discrete functions, in that the resulting functions $f^{\bullet}$ and $h^{\circ}$ are also integer-valued on integer points. We call (1.9) and (1.10), respectively, convex and concave *discrete Legendre–Fenchel transformations*.

With these definitions, a discrete version of the Fenchel duality would read as follows.

> [Discrete Fenchel-type duality theorem] *Let* $f : \mathbf{Z}^n \to \mathbf{Z} \cup \{+\infty\}$ *and* $h : \mathbf{Z}^n \to \mathbf{Z} \cup \{-\infty\}$ *be "convex" and "concave" functions, respectively* (*in an appropriate sense*)*. Then*
>
> $$\min\{f(x) - h(x) \mid x \in \mathbf{Z}^n\} = \max\{h^{\circ}(p) - f^{\bullet}(p) \mid p \in \mathbf{Z}^n\}.$$

Such a theorem, if any, claims a min-max duality relation for integer-valued nonlinear functions, which is not likely to be true for an arbitrary class of discrete functions. It is emphasized that the definition of "convexity" itself is left open in the above generic statement, although $h$ should be called "concave" when $-h$ is "convex."

As for the separation theorem, a possible discrete version would read as follows, imposing integrality ($\alpha^* \in \mathbf{Z}$, $p^* \in \mathbf{Z}^n$) on the separating affine function. See Fig. 1.5.

> [Discrete separation theorem] *Let* $f : \mathbf{Z}^n \to \mathbf{Z} \cup \{+\infty\}$ *and* $h : \mathbf{Z}^n \to \mathbf{Z} \cup \{-\infty\}$ *be "convex" and "concave" functions, respectively* (*in an appropriate sense*)*. If*
>
> $$f(x) \geq h(x) \qquad (\forall x \in \mathbf{Z}^n),$$
>
> *there exist* $\alpha^* \in \mathbf{Z}$ *and* $p^* \in \mathbf{Z}^n$ *such that*
>
> $$f(x) \geq \alpha^* + \langle p^*, x \rangle \geq h(x) \qquad (\forall x \in \mathbf{Z}^n).$$

Again the precise definition of "convexity" remains unspecified here.

To motivate our framework to be introduced in the subsequent sections, let us try with a naive and natural candidate for the "convexity" concept, which turns out to be insufficient.

Let us (temporarily) define $f : \mathbf{Z}^n \to \mathbf{Z} \cup \{+\infty\}$ to be "convex" if it can be extended to a convex function on $\mathbf{R}^n$, i.e., if there exists a convex function $\overline{f} : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ such that

$$\overline{f}(x) = f(x) \qquad (x \in \mathbf{Z}^n). \tag{1.11}$$

This is illustrated in Fig. 1.6.

In the one-dimensional case (with $n = 1$) this is equivalent to defining $f : \mathbf{Z} \to \mathbf{Z} \cup \{+\infty\}$ to be "convex" if

$$f(x-1) + f(x+1) \geq 2f(x) \qquad (\forall x \in \mathbf{Z}). \tag{1.12}$$

As is easily verified, the discrete separation theorem as well as the discrete Fenchel duality holds with this definition in the case of $n = 1$.

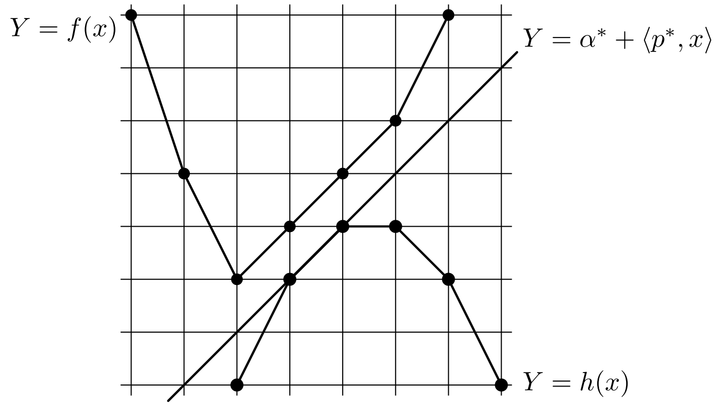


**Figure 1.5.** *Discrete separation*

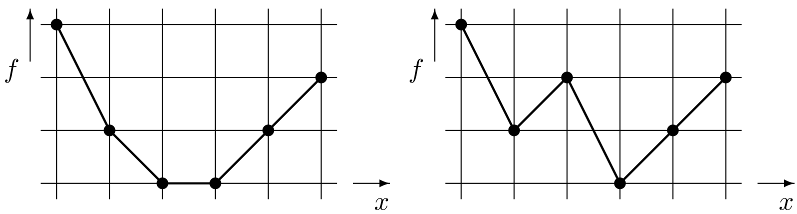


**Figure 1.6.** *"Convex" and non-"convex" discrete functions*

When it comes to higher dimensions, the situation is not that simple. The following examples demonstrate that the discrete separation fails with this naive definition of "convexity."

**Example 1.5.** [failure of discrete separation] Consider two discrete functions defined by

$$f(x) = \max(0, x(1) + x(2)), \qquad h(x) = \min(x(1), x(2)),$$

where $x = (x(1), x(2)) \in \mathbf{Z}^2$. They are integer-valued on the integer lattice $\mathbf{Z}^2$ with $f(\mathbf{0}) = h(\mathbf{0}) = 0$, and can be extended, respectively, to a convex function $\overline{f} : \mathbf{R}^2 \to \mathbf{R}$ and a concave function $\overline{h} : \mathbf{R}^2 \to \mathbf{R}$ given by

$$\overline{f}(x) = \max(0, x(1) + x(2)), \qquad \overline{h}(x) = \min(x(1), x(2)),$$

where $x = (x(1), x(2)) \in \mathbf{R}^2$. Since $\overline{f}(x) \geq \overline{h}(x)$ $(\forall x \in \mathbf{R}^2)$, the separation theorem in convex analysis (Theorem 1.3) applies to the pair $(\overline{f}, \overline{h})$, to yield a (unique) separating affine function $\langle \overline{p}^*, x\rangle$ with $\overline{p}^* = (1/2, 1/2)$. We have $\overline{f}(x) \geq \langle \overline{p}^*, x\rangle \geq \overline{h}(x)$ for all $x \in \mathbf{R}^2$, and a fortiori, $f(x) \geq \langle \overline{p}^*, x\rangle \geq h(x)$ for all $x \in \mathbf{Z}^2$. However, there exists no integral vector $p^* \in \mathbf{Z}^2$ such that $f(x) \geq \langle p^*, x \rangle \geq h(x)$ for all $x \in \mathbf{Z}^2$. This demonstrates the failure of the desired discreteness in the separating affine function. ■

**Example 1.6.** [failure of real-valued separation] This example shows that even the existence of a separating affine function can be denied. For the discrete functions

$$f(x) = |x(1) + x(2) - 1|, \qquad h(x) = 1 - |x(1) - x(2)|,$$

where $x = (x(1), x(2)) \in \mathbf{Z}^2$, we have $f(x) \geq h(x)$ $(\forall x \in \mathbf{Z}^2)$. There exists, however, no pair of real number $\alpha^* \in \mathbf{R}$ and a real vector $p^* \in \mathbf{R}^2$ for which $f(x) \geq \alpha^* + \langle p^*, x \rangle \geq h(x)$ for all $x \in \mathbf{Z}^2$. Note that the separation theorem in convex analysis (Theorem 1.3) does not apply to the pair of their convex/concave extensions $(\overline{f}, \overline{h})$, which are given by

$$\overline{f}(x) = |x(1) + x(2) - 1|, \qquad \overline{h}(x) = 1 - |x(1) - x(2)|$$

for $x = (x(1), x(2)) \in \mathbf{R}^2$, since $\overline{f}(1/2, 1/2) < \overline{h}(1/2, 1/2)$. This example shows also that $\overline{f} \geq \overline{h}$ on $\mathbf{R}^n$ does not follow from $f \geq h$ on $\mathbf{Z}^n$. ■

Similarly, the discrete Fenchel duality fails under the naive definition of "convexity." The above two examples serve to demonstrate this.

Thus the naive approach to discrete convexity does not work, and some deep combinatorial or discrete-mathematical considerations are needed. We are now motivated to look at some results in the area of matroids and submodular functions, which hopefully provide a clue for fruitful definitions of discrete convexity.