John Nachbar
Washington University
March 27, 2018

Concave and Convex Functions[1]

1 Basic Definitions.

Definition 1. *Let $C \subseteq \mathbb{R}^N$ be non-empty and convex and let $f : C \to \mathbb{R}$.*

1. (a) *f is* concave *iff for any $a, b \in C$ and any $\theta \in [0, 1]$,*

$$f(\theta a + (1 - \theta)b) \geq \theta f(a) + (1 - \theta)f(b);$$

   (b) *f is* strictly concave *iff for any $a, b \in C$ and any $\theta \in (0, 1)$, the above inequality is strict.*

2. (a) *f is* convex *iff for any $a, b \in C$ and any $\theta \in [0, 1]$,*

$$f(\theta a + (1 - \theta)b) \leq \theta f(a) + (1 - \theta)f(b);$$

   (b) *f is* strictly convex *iff for any $a, b \in C$ and any $\theta \in (0, 1)$, the above inequality is strict.*

The following equivalence is immediate from the definitions.

Theorem 1. *Let $C \subseteq \mathbb{R}^N$ be non-empty and convex and let $f : C \to \mathbb{R}$. f is convex iff $-f$ is concave. f is strictly convex iff $-f$ is strictly concave.*

f is both concave and convex iff for any $a, b \in \mathbb{R}^N$ and any $\theta \in (0, 1)$, $f(\theta a + (1 - \theta)b) = \theta f(a) + (1 - \theta)f(b)$. A function f is *affine* iff there is a $1 \times N$ matrix A and a number $y^* \in \mathbb{R}$ such that for all $x \in C$, $f(x) = Ax + y^*$. f is *linear* if it is affine with $y^* = 0$.

Theorem 2. *$f : \mathbb{R}^N \to \mathbb{R}$ is affine iff it is both concave and convex.*

Proof.

1. $\Rightarrow$. For any $a, b \in C$ and any $\theta \in (0, 1)$, $f(\theta a + (1-\theta)b) = A(\theta a + (1-\theta)b) + y^* = \theta(Aa + y^*) + (1 - \theta)(Ab + y^*) = \theta f(a) + (1 - \theta)f(b)$, so that f is both concave and convex.

2. $\Leftarrow$. Let $y^* = f(0)$ and let $g(x) = f(x) - y^*$, so that $g(0) = 0$. Since f is both concave and convex, so is g.

- **Claim**: for any $a \in \mathbb{R}^N$, for any $\gamma \geq 0$, $g(\gamma a) = \gamma g(a)$.

  The claim is trivially true for γ equal to either 0 or 1. Suppose $\gamma \in (0, 1)$. Then $g(\gamma a) = g(\gamma a + (1 - \gamma)0) = \gamma g(a) + (1 - \gamma)g(0) = \gamma g(a)$.

  On the other hand, if $\gamma > 1$, then $1/\gamma \in (0, 1)$ and hence $g(a) = g((1/\gamma)\gamma a + (1 - 1/\gamma)0) = (1/\gamma)g(\gamma a) + (1 - 1/\gamma)g(0) = (1/\gamma)g(\gamma a)$. Multiplying through by γ gives $\gamma g(a) = g(\gamma a)$.

- **Claim:** for any $a, b \in \mathbb{R}^N$, $g(a + b) = g(a) + g(b)$.

  $g(a + b) = g((1/2)(2a) + (1/2)(2b)) = (1/2)g(2a) + (1/2)g(2b) = g(a) + g(b)$, where the last equality comes from the previous claim.

Construct A by setting $a_n = g(e^n)$, where $e^n = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the coordinate n unit vector. Since $x = \sum_n x_n e^n$, induction on the second claim above gives $g(x) = g(\sum_n x_n e^n) = \sum_n g(x_n e^n) = \sum_n x_n g(e^n) = \sum_n x_n a_n = Ax$. Finally, $f(x) = g(x) + y^* = Ax + y^*$.

∎

For $N = 1$, the next result says that a function is concave iff, informally, its slope is weakly decreasing. If the function is differentiable then the implication is that the derivative is weakly decreasing.

Theorem 3. *Let $C \subseteq \mathbb{R}$ be an open interval.*

1. *$f : C \to \mathbb{R}$ is concave iff for any $a, b, c \in C$, with $a < b < c$,*

$$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(b)}{c - b},$$

   and,

$$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(a)}{c - a}.$$

   For strict concavity, the inequalities are strict.

2. *$f : C \to \mathbb{R}$ is convex iff for any $a, b, c \in C$, with $a < b < c$,*

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(b)}{c - b},$$

   and,

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a}.$$

   For strictly convexity, the inequalities are strict.

Proof. Take any $a, b, c \in C$, $a < b < c$. Since $b - a$ and $c - b > 0$, the first inequality under (1), holds iff

$$[f(b) - f(a)](c - b) \geq [f(c) - f(b)](b - a),$$

which holds iff (collecting terms in $f(b)$),

$$f(b)(c - a) \geq f(a)(c - b) + f(c)(b - a),$$

which (since $c - a > 0$) holds iff

$$f(b) \geq \left[\frac{c - b}{c - a}\right] f(a) + \left[\frac{b - a}{c - a}\right] f(c).$$

Take $\theta = (c - b)/(c - a) \in (0, 1)$ and verify that, indeed, $b = \theta a + (1 - \theta)c$. Then the last inequality holds since f is concave. Conversely, the preceding argument shows that if the first inequality in (1) holds then f is concave (take any $a < c$, any $\theta \in (0, 1)$, and let $b = \theta a + (1 - \theta)c$). The proofs of the other claims are similar. ∎

It is also possible to characterize concavity or convexity of functions in terms of the convexity of particular sets. Given the graph of a function, the *hypograph* of f, written $\mathrm{hyp} f$, is the set of points that lies on or below the graph of f, while the *epigraph* of f, written $\mathrm{epi} f$, is the set of points that lies on or above the graph of f.[2] Formally,

$$\mathrm{epi} f = \{(x, y) \in \mathbb{R}^{N+1} : y \geq f(x)\},$$
$$\mathrm{hyp} f = \{(x, y) \in \mathbb{R}^{N+1} : y \leq f(x)\}.$$

Theorem 4. *Let $C \subseteq \mathbb{R}^N$ be convex and let $f : C \to \mathbb{R}$.*

  1. *f is concave iff $\mathrm{hyp} f$ is convex.*

  2. *f is convex iff $\mathrm{epi} f$ is convex.*

Proof. Suppose that f is concave. I will show that $\mathrm{hyp} f$ is convex. Take any $z_1, z_2 \in \mathrm{hyp} f$ and any $\theta \in [0, 1]$. Then there is an $a, b \in C$ and $y_1, y_2 \in \mathbb{R}$, such that $z_1 = (a, y_1)$, $z_2 = (b, y_2)$, with $f(a) \geq y_1$, $f(b) \geq y_2$. By concavity of f, $f(\theta a + (1 - \theta)b) \geq \theta f(a) + (1 - \theta)f(b)$. Hence $f(\theta a + (1 - \theta)b) \geq \theta y_1 + (1 - \theta)y_2$. The latter says that the point $\theta z_1 + (1 - \theta)z_2 = (\theta a + (1 - \theta)b, \theta y_1 + (1 - \theta)y_2) \in \mathrm{hyp} f$, as was to be shown. The other directions are similar. ∎

---

[2] "Hypo" means "under" and "epi" means "over." A hypodermic needle goes under your skin, the top layer of which is your epidermis.

2 Concavity, Convexity, and Continuity.

Theorem 5. *Let $C \subseteq \mathbb{R}^N$ be non-empty, open and convex and let $f : C \to \mathbb{R}$ be either convex or concave. Then f is continuous.*

Proof. Let f be concave. Consider first the case $N = 1$. Theorem 3 implies that for any $a, b, c \in C$, with $a < b < c$, the graph of f is sandwiched between the graphs of two lines through the point $(b, f(b))$, one line through the points $(a, f(a))$ and $(b, f(b))$ and the other through the points $(b, f(b))$ and $(c, f(c))$. Explicitly, Theorem 3 implies that for all $x \in [a, b]$,

$$f(b) - \frac{f(b) - f(a)}{b - a}(b - x) \le f(x) \le f(b) - \frac{f(c) - f(b)}{c - b}(b - x), \tag{1}$$

and for all $x \in [b, c]$,

$$f(b) + \frac{f(b) - f(a)}{b - a}(x - b) \ge f(x) \ge f(b) + \frac{f(c) - f(b)}{c - b}(x - b). \tag{2}$$

These inequalities imply that f is continuous at b.

The argument for general N uses the following additional fact from the 1-dimensional case. Assume that b lies in the center of the $[a, c]$ line segment, so that $b - a = c - b$. If $f(a) \le f(c)$ then for any $x \in [a, c]$,

$$|f(b) - f(x)| \le f(b) - f(a). \tag{3}$$

If $f(c) \le f(a)$, then the analog of (3) holds with $f(b) - f(c)$ on the right-hand side.

Inequality (3) may be obvious from the "sandwich" characterization above, but for the sake of completeness, here is a detailed argument. First, note that concavity implies that $f(b) \ge f(a)$. Next, the first inequality in (1) implies that for all $x \in [a, b]$,

$$f(b) - f(x) \le \frac{f(b) - f(a)}{b - a}(b - x) \le f(b) - f(a), \tag{4}$$

where the second inequality in (4) follows since $f(b) - f(a) \ge 0$ and since $(b-x)/(b-a)$ is non-negative and has a maximum value of 1 on $[a, b]$. Similarly, the second inequality in (1) implies that

$$-(f(b) - f(x)) \le -\frac{f(c) - f(b)}{c - b}(b - x) \le f(b) - f(a), \tag{5}$$

where the second inequality in (5) follows trivially if $-(f(c) - f(b)) \le 0$, since $f(b) - f(a) \ge 0$; if $-(f(c) - f(b)) > 0$, then the inequality follows since $-f(c) \le -f(a)$ and since $(b - x)/(b - a)$ is non-negative and has a maximum value of 1 on $[a, b]$). And similar arguments obtain if $x \in [b, c]$. Combining all this gives inequality (3).

4

To complete the proof of continuity, take any $x^* \in C$ and consider the (hyper) cube formed by the $2N$ vertices of the form $x^* + (1/t)e^n$ and $x^* - (1/t)e^n$, where e^n is the unit vector for coordinate n and where $t \in \{1, 2, \dots\}$ is large enough that this cube lies in C. Let v_t be the vertex that minimizes f across the $2N$ vertices. For any x in the cube, concavity implies $f(v_t) \leq f(x)$. Take any x in the cube. Then for the line segment given by the intersection of the cube with the line through x and x^*, inequality (3) implies, since $f(v_t)$ is less than or equal to the minimum value of f along this line segment, and since x^* lies in the center of this line segment.

$$|f(x) - f(x^*)| \leq f(x^*) - f(v_t). \tag{6}$$

Since there are only a finite number of coordinates, it is possible to find a subsequence $\{v_{t_k}\}$ lying on a single axis. Continuity in the 1-dimensional case, established above, implies $f(v_{t_k}) \to f(x^*)$. Inequality (6) then implies continuity at x^*.

Finally, for convex f, $-f$ is concave, hence $-f$ is continuous, and f is continuous iff $-f$ is continuous. ■

For functions defined on non-open sets, continuity can fail at the boundary. In particular, if the domain is a closed interval in $\mathbb{R}$, then concave functions can jump down at end points and convex functions can jump up.

Example 1. Let $C = [0, 1]$ and define

$$f(x) = \begin{cases} -x^2 & \text{if } x > 0, \\ -1 & \text{if } x = 0. \end{cases}$$

Then f is concave. It is lower semi-continuous on $[0, 1]$ and continuous on $(0, 1]$. □

Remark 1. The proof of Theorem 5 makes explicit use of the fact that the domain is finite dimensional. The theorem does not generalize to domains that are arbitrary vector metric spaces. In particular, there are infinite dimensional vector space domains for which even some linear functions (which are both concave and convex) are not continuous. □

3 Concavity, Convexity, and Differentiability.

A differentiable function is concave iff it lies on or below the tangent line (or plane, for $N > 1$) at any point.

Theorem 6. *Let $C \subseteq \mathbb{R}^N$ be non-empty, open and convex and let $f : C \to \mathbb{R}$ be differentiable.*

1. *f is concave iff for any $x^*, x \in C$*

$$f(x) \leq \nabla f(x^*) \cdot (x - x^*) + f(x^*). \tag{7}$$

2. *f is convex iff for any $x^*, x \in C$*

$$f(x) \geq \nabla f(x^*) \cdot (x - x^*) + f(x^*). \tag{8}$$

Proof. If f is concave then for any $x, x^* \in C$, $x \neq x^*$, and any $\theta \in (0,1)$, $f(\theta x + (1-\theta)x^*) \geq \theta f(x) + (1-\theta)f(x^*)$, or, dividing by θ and rearranging,

$$f(x) - f(x^*) \leq \frac{f(x^* + \theta(x - x^*)) - f(x^*)}{\theta}.$$

Taking the limit of the right-hand side as $\theta \downarrow 0$ and rearranging yields inequality (7).

Conversely, consider any $a, b \in C$, take any $\theta \in (0,1)$, and let $x^* = \theta a + (1-\theta)b$. Note that $a - x^* = -(1-\theta)(b-a)$ and $b - x^* = \theta(b-a)$. Therefore, by inequality (7),

$$f(a) \leq \nabla f(x^*) \cdot [-(1-\theta)(b-a)] + f(x^*)$$
$$f(b) \leq \nabla f(x^*) \cdot [\theta(b-a)] + f(x^*)$$

Multiplying the first by $\theta > 0$ and the second by $1 - \theta > 0$, and adding, yields $\theta f(a) + (1-\theta)f(b) \leq f(x^*)$, as was to be shown.

The proofs for inequality (8) are analogous. ■

Remark 2. If f is concave then a version of inequality (7) holds even if f is not differentiable (e.g., $f(x) = -|x|$), and analogously for inequality (8) if f is convex.

Explicitly, recall from Theorem 4 that f is concave iff hyp f, the set lying on or below the graph of f, is convex. Therefore, by the Supporting Hyperplane Theorem for $\mathbb{R}^{N+1}$, for any $x^* \in C$, there is a vector $(v, w) \in \mathbb{R}^{n+1}$, $v \in \mathbb{R}^N$, $w \in \mathbb{R}$, $(v, w) \neq 0$, such that for all $(x, y) \in \text{hyp} f$,

$$(v, w) \cdot (x, y) \geq (v, w) \cdot (x^*, f(x^*)). \tag{9}$$

If $w > 0$, this inequality will be violated for (x^*, y) for any $y < 0$ of sufficiently large magnitude. By definition of hyp$f$, there will be such $(x^*, y) \in \text{hyp} f$. Therefore, $w \leq 0$. Moreover, if $w = 0$ (and $v \neq 0$ since $(v, w) \neq 0$), then inequality (9) will be violated at any $(x, f(x))$ with $x = x^* - \gamma v$, with $\gamma > 0$ small enough that $x \in C$. Therefore, $w < 0$.

Since $w < 0$, I can assume $w = -1$: inequality (9) holds for (v, w) iff it holds for $\gamma(v, w)$, for any $\gamma > 0$; take $\gamma = 1/|w|$. With $(v, -1)$ as the supporting vector, inequality (9) then implies, taking $(x, y) = (x, f(x))$ and rearranging,

$$f(x) \leq v \cdot (x - x^*) + f(x^*). \tag{10}$$

If f is differentiable at x^*, then $\nabla f(x^*)$ is the unique v for which (10) holds for all x. If f is not differentiable at x^* then there will be a continuum of vectors, called *subgradients*, for which inequality (10) holds.

It is easy to verify that the set of subgradients is closed and convex. In the case $N = 1$, a subgradient is just a number and the set of subgradients is particularly easy to characterize. Explicitly, for any $x^* \in C$, Theorem 3 implies that the left-hand and right-hand derivatives at x^*,

$$\underline{m} = \lim_{x \uparrow x^*} \frac{f(x) - f(x^*)}{x - x^*},$$

$$\overline{m} = \lim_{x \downarrow x^*} \frac{f(x) - f(x^*)}{x - x^*},$$

are well defined even if f is not differentiable at x^*. The set of subgradients at x^* is $[\underline{m}, \overline{m}]$; if f is differentiable at x^* then $\underline{m} = \overline{m} = Df(x^*)$.

Subgradients play an important role in some parts of economic theory, but I will not be pursuing them here. $\square$

For differentiable functions on $\mathbb{R}$ ($N = 1$), Theorem 3 says implies that a function is concave iff its derivative is weakly decreasing. For twice differentiable functions, the derivative is weakly decreasing iff the second derivative is weakly negative everywhere. The following result, Theorem 7, records this fact and generalizes it to $N > 1$. Recall that a symmetric $N \times N$ matrix A is *negative semi-definite* iff, for any $v \in \mathbb{R}^N$, $v'Av \leq 0$. The matrix is *negative definite* iff, for any $v \in \mathbb{R}^N$, $v \neq 0$, $v'Av < 0$. The definitions of *positive semi-definite* and *positive definite* are analogous. As is standard practice, I write $D^2f(x)$ for the Hessian, $D^2f(x) = D(\nabla f)(x)$, which is the $N \times N$ matrix of second order partial derivatives. By Young's Theorem, if f is $\mathcal{C}^2$ then $D^2f(x)$ is symmetric.

Theorem 7. *Let $C \subseteq \mathbb{R}^N$ be non-empty, open and convex and let $f : C \to \mathbb{R}$ be $\mathcal{C}^2$.*

1. (a) *$D^2f(x)$ is negative semi-definite for every $x \in C$ iff f is concave.*

   (b) *If $D^2f(x)$ is negative definite for every $x \in C$ then f is strictly concave.*

2. (a) *$D^2f(x)$ is positive semi-definite for every $x \in C$ iff f is convex.*

   (b) *If $D^2f(x)$ is positive definite for every $x \in C$ then f is strictly convex.*

Proof.

1. $N = 1$. In this case, $D^2f(x) \in \mathbb{R}$, hence $D^2f(x)$ is negative semi-definite iff $D^2f(x) \leq 0$.

   Consider first the $\Rightarrow$ direction of 1(a). If $D^2f(x) \leq 0$ for all x then $Df(x)$ is weakly decreasing for all x, which implies that for any $a, b, c \in C$, $a < b < c$,

   $$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(b)}{c - b},$$

   which, by Theorem 3, implies that f is concave. The proof of 1(b) is almost identical, as is the proof of 2(b) and the $\Rightarrow$ direction of 2(a).

It remains to prove the $\Leftarrow$ direction of 1(a) and 2(a). Consider the $\Leftarrow$ direction of 1(a). I argue by contraposition. Suppose that $D^2 f(x^*) > 0$ for some $x^* \in C$. Since f is $\mathcal{C}^2$, $D^2 f(x) > 0$ for every x in some open interval containing x^*. Then Theorem 3 implies that f is not concave (it is, in fact, strictly convex) for x in this interval. The proof of the $\Leftarrow$ direction of 2(a) is similar.

2. $N > 1$. Then $D^2 f(x)$ is a symmetric matrix. I will show 1(b). The other cases are similar.

   Suppose, therefore, that $D^2 f(x)$ is negative definite for all $x \in C$. Consider any $a, b \in C$, $b \neq a$, and any $\theta \in (0, 1)$. Let $x_\theta = \theta a + (1 - \theta)b$. To show 1(b), I need to show that $f(x_\theta) > \theta f(a) + (1 - \theta)f(b)$.

   Let $g(\theta) = b + \theta(a - b)$, let $h(\theta) = f(g(\theta)) = f(b + \theta(a - b))$, and let $v = a - b$. By the $N = 1$ step above, a sufficient condition for the strict concavity of h is that $D^2 h(\theta) < 0$. The interpretation is that $D^2 h(\theta)$ is the second derivative of f, evaluated at x_θ, in the direction $v = a - b$.

   By the Chain Rule, for any $\theta \in (0, 1)$, $Dh(\theta) = Df(g(\theta))Dg(\theta) = Df(g(\theta))v = \nabla f(g(\theta)) \cdot v$. Also by the Chain Rule (and the symmetry of $D^2 f$),

   $$D[\nabla f(g(\theta)) \cdot v] = [D^2 f(x_\theta)v] \cdot v = v' D^2 f(x_\theta)v.$$

   Hence,
   $$D^2 h(\theta) = v' D^2 f(x_\theta)v.$$

   Therefore, if $v' D^2 f(x_\theta)v < 0$ for every $\theta \in (0, 1)$ then h is strictly concave on $(0, 1)$, which implies that $h(\theta) > \theta h(1) + (1 - \theta)h(0)$, which implies $f(x_\theta) > \theta f(a) + (1 - \theta)f(b)$, as was to be shown.

■

Say that f is *differentiably strictly concave* iff $D^2 f(x)$ is negative definite for every x. If $N = 1$ then f is differentiably strictly concave iff $D^2 f(x) < 0$ for every x. The definition of *differentiable strict convexity* is analogous.

Example 2. Consider $f : \mathbb{R} \to \mathbb{R}$, $f(x) = -x^4$. f is strictly concave but fails differentiable strict concavity since $D^2 f(0) = 0$. □

Remark 3. To reiterate a point made in the proof of Theorem 7, if f is $\mathcal{C}^2$ and $D^2 f$ is negative definite at the point x^* then $D^2 f(x)$ is negative definite for every x in some open ball around x^*, and hence f is strictly concave in some open ball around x^*. In words, if f is differentiable strictly concave at a point then it is differentiably strictly concave near that point.

If, on the other hand, $D^2 f$ is merely negative semi-definite at x^* then we cannot infer anything about the concavity or convexity of f near x^*. For example, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = x^4$ then $Df(0) = 0$, which is negative semi-definite, but f is not concave; it is, in fact, strictly convex. □

4 Facts about Concave and Convex Functions.

Recall that $f : \mathbb{R}^N \to \mathbb{R}$ is affine iff it is of the form $f(x) = Ax + b$ for some $1 \times N$ matrix A and some point $b \in \mathbb{R}$. Geometrically, the graph of a real-valued affine function is a plane (a line, if the domain is $\mathbb{R}$). An important elementary fact is that real-valued affine functions are both concave and convex. This is consistent with the fact that the second derivative of any affine function is the zero matrix.

Showing that other functions are concave or convex typically requires work. For $N = 1$, Theorem 7 can be used to show that many standard functions are concave, strictly concave, and so on.

Example 3. All of the following claims can be verified with a simple calculation.

1. e^x is strictly convex on $\mathbb{R}$,

2. $\ln(x)$ is strictly concave on $\mathbb{R}_+$,

3. $1/x$ is strictly convex on $\mathbb{R}_{++}$ and strictly concave on $\mathbb{R}_{--}$,

4. x^t, where t is an integer greater than 1, is strictly convex on $\mathbb{R}_+$. On $\mathbb{R}_-$, x^t is strictly convex for t even and strictly concave for t odd.

5. x^α, where α is a real number in $(0, 1)$ is strictly concave on $\mathbb{R}_+$.

□

One can often verify the concavity of other, more complicated functions by decomposing the functions into simpler pieces. The following results help do this.

Theorem 8. *Let $C \subseteq \mathbb{R}^N$ be non-empty and convex. Let $f : C \to \mathbb{R}$ be concave. Let D be any interval containing $f(C)$ and let $g : D \to \mathbb{R}$ be concave and weakly increasing. Then $h : C \to \mathbb{R}$ defined by $h(x) = g(f(x))$ is concave. Moreover, if f is strictly concave and g is strictly increasing then h is strictly concave. Analogous claims hold for f convex (again with g increasing).*

Proof. Consider any $a, b \in C$ and any $\theta \in [0, 1]$. Let $x_\theta = \theta a + (1 - \theta)b$. Since f is concave,
$$f(x_\theta) \geq \theta f(a) + (1 - \theta)f(b).$$
Then
$$
\begin{aligned}
h(x_\theta) = g(f(x_\theta)) &\geq g(\theta f(a) + (1 - \theta)f(b)) \\
&\geq \theta g(f(a)) + (1 - \theta)g(f(b)) \\
&= \theta h(a) + (1 - \theta)h(b),
\end{aligned}
$$
where the first inequality does from the fact that f is concave and g is weakly increasing and the second inequality comes from the fact that g is concave. The other parts of the proof are essentially identical. ■

Example 4. Let the domain be $\mathbb{R}$. Consider $h(x) = e^{1/x}$. Let $f(x) = 1/x$ and let $g(y) = e^y$. Then $h(x) = g(f(x))$. Function f is strictly convex and g is (strictly) convex and strictly increasing. Therefore, by Theorem 8, h is strictly convex. $\square$

It is important in Theorem 8 that g be increasing.

Example 5. Let the domain by $\mathbb{R}$. Consider $h(x) = e^{-x^2}$. This is just the standard normal density except that it is off by a factor of $1/\sqrt{2\pi}$. Let $f(x) = e^{x^2}$ and let $g(y) = 1/y$. Then $h(x) = g(f(x))$. Now, f is convex on $\mathbb{R}_{++}$ (indeed, on $\mathbb{R}$) and g is also convex on $\mathbb{R}_{++}$. The function h is not, however, convex. While it is strictly convex for $|x|$ sufficiently large, for x near zero it is strictly concave. This does not contradict Theorem 8 because g here is decreasing. $\square$

Theorem 9. *Let $C \subseteq \mathbb{R}$ be a interval and let $f : C \to \mathbb{R}$ be concave and strictly positive for all $x \in C$. Then $h : C \to \mathbb{R}$ defined by $h(x) = 1/f(x)$ is convex.*

Proof. Follows from Theorem 8, the fact that if f is concave then $-f$ is convex, and the fact that the function $g(x) = -1/x$ is convex and increasing on $\mathbb{R}_-$. ∎

Theorem 10. *Let $C \subseteq \mathbb{R}$ be an open interval. If $f : C \to \mathbb{R}$ is strictly increasing or decreasing then $f^{-1} : f(C) \to C$ is well defined. If, in addition, f is concave or convex, then $f(C)$ is convex and the following hold.*

1. *If f is concave and strictly increasing then f^{-1} is convex.*

2. *If f is concave and strictly decreasing then f^{-1} is concave.*

3. *If f is convex and strictly increasing then f^{-1} is concave.*

4. *If f is convex and strictly decreasing then f^{-1} is convex.*

Proof. For 1, by Theorem 5, f is continuous. It follows (see the notes on connected sets) that $f(C)$ is an interval, and hence is convex. Let y, $\hat{y}$ be any two points in $f(C)$ and let $x = f^{-1}(y)$, $\hat{x} = f^{-1}(\hat{y})$. Take any $\theta \in [0, 1]$. Then, since f is concave,

$$f(\theta x + (1-\theta)\hat{x}) \geq \theta f(x) + (1-\theta)f(\hat{x})$$
$$= \theta y + (1-\theta)\hat{y}.$$

Taking the inverse of both sides yields, since f is strictly increasing, and since $x = f^{-1}(y)$, $\hat{x} = f^{-1}(\hat{y})$,

$$\theta f^{-1}(y) + (1-\theta)f^{-1}(\hat{y}) = \theta x + (1-\theta)\hat{x}$$
$$\geq f^{-1}(\theta y + (1-\theta)\hat{y}).$$

Since y, $\hat{h}$, and θ were arbitrary, this implies that f^{-1} is convex. The other results are similar. Note in particular that if f is concave but strictly decreasing then the last inequality above flips, and f^{-1} is concave. ∎

Example 6. Let $C = (0, \infty)$ and let $f(x) = \ln(x)$. This function is (strictly) concave and strictly increasing. Its inverse is $f^{-1}(y) = e^y$, which is (strictly) convex and strictly increasing. $\square$

Example 7. Let $C = (-\infty, 0)$ and let $f(x) = \ln(-x)$. This function is (strictly) concave and strictly decreasing. Its inverse is $f^{-1}(y) = -e^y$, which is (strictly) concave and strictly decreasing. $\square$

Theorem 11. *Let $C \subseteq \mathbb{R}^N$ be non-empty and convex. Let $f_1 : C \to \mathbb{R}$ and $f_2 : C \to \mathbb{R}$ be concave.*

1. *The function $f_1 + f_2$ is concave. Moreover, if either f_1 or f_2 is strictly concave then $f_1 + f_2$ is strictly concave.*

2. *For any $r \geq 0$, the function rf_1 is concave. Moreover, if f_1 is strictly concave then for any $r > 0$, rf_1 is strictly concave.*

Analogous claims hold if f_1, f_2 are convex.

Proof. Omitted. Almost immediate from the definition of concavity. $\blacksquare$

A special case in which $N > 1$ is effectively as easy to analyze as $N = 1$ is when the function is separable in the sense that it is the sum of univariate functions. For example, the function $h(x_1, x_2) = e^{x_1} + e^{x_2}$ is separable. More generally, suppose $C_n \subseteq \mathbb{R}$ are intervals and let $C = \prod_n C_n$. Then $h : C \to \mathbb{R}$ is *separable* iff there are functions $f_n : C_n \to \mathbb{R}$ such that

$$h(x) = \sum_{n=1}^{N} f_n(x_n).$$

Theorem 12. *Let $C_n \subseteq \mathbb{R}$ be intervals, let $C = \prod_n C_n$, and for each n let $f_n : C_n \to \mathbb{R}$. Define the separable function $h : C \to \mathbb{R}$ by $h(x) = \sum_n f_n(x)$. Then h is concave iff every f_n is concave and h is strictly concave iff every f_n is strictly concave. If every f_n is $\mathcal{C}^2$ then h is differentiably strictly concave iff every f_n is differentiably strictly concave. And analogous statements hold for convexity.*

Proof. The claims for concavity and strict concavity are almost immediate. For differentiable strict concavity, note that

$$D^2 h(x^*) = \begin{bmatrix} D^2 f_1(x^*) & 0 & \cdots & 0 & 0 \\ 0 & D^2 f_2(x^*) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & D^2 f_{N-1}(x^*) & 0 \\ 0 & 0 & \cdots & 0 & D^2 f_N(x^*) \end{bmatrix}.$$

This matrix is negative definite iff the diagonal terms are all strictly negative, which is equivalent to saying that the f_n are all differentiably strictly concave. ∎

Theorem 12 is superficially similar to Theorem 11 but there are important differences, as illustrated in the following example.

Example 8. Consider $h(x_1, x_2) = e^{x_1} + e^{x_2}$. Since h is separable, Theorem 12, and the fact that the second derivative of e^x is always strictly positive, implies that h is strictly convex.

I can also employ Theorem 11, but I reach the weaker conclusion that h is convex, rather than strictly convex. Explicitly, I can view e^{x_1} not as a function on $\mathbb{R}$ but as a function on $\mathbb{R}^2$ where the second coordinate is simply ignored. Viewed as a function on $\mathbb{R}^2$, e^{x_1} is convex (one can use Theorem 8 to show this) but not strictly convex. Because the e^{x_n} are convex but not strictly convex, Theorem 11 implies only that h is convex, not strictly convex. □