

MECH 6327 - Homework 2

Jonas Wagner

2021, March 3

(Probably not necessary... but its long)

Contents

1	Problem Set 1: Convex Sets	3
1.1	Problem 2.5	3
1.2	Problem 2.7	3
1.3	Problem 2.12	4
1.3.1	(a) - Slab	4
1.3.2	(b) - Rectangle/Hyperrectangle	4
1.3.3	(c) - Wedge	4
1.3.4	(d) - Closer to a point then a set	5
1.3.5	(e) - Closer to a set then another set	5
1.3.6	(f) - Set of the sum being within a convex set	5
1.3.7	(g) - Set with weighted distances to two points	5
1.4	Problem 2.28	6
1.5	Problem 2.33	7
1.5.1	Part a	7
1.5.2	Part b	8
2	Problem Set 2: Convex Functions	9
2.1	Problem 3.6	9
2.1.1	(a) - Epigraph of a halfspace	9
2.1.2	(b) - Epigraph of a convex cone	9
2.1.3	(c) - Epigraph of a polyhedron	9
2.2	Problem 3.16	10
2.2.1	(a) - $e^x - 1$	10
2.2.2	(b) - x_1x_2	10
2.2.3	(c) - $1/(x_1x_2)$	10
2.2.4	(d) - x_1/x_2	10
2.3	(e) - x_1^2/x_2	11
2.4	(f) - $x_1^\alpha x_2^{1-\alpha}$	11
2.5	Problem 3.18a	12
2.6	Problem 3.22	14
2.6.1	(a) - double log functions	14
2.6.2	(b) - square root of some product sum	14
2.6.3	(c) - log of some product sum	15
2.6.4	(d) - complicated root of a powered sum and norm	15
2.6.5	(e) - complicated log of a powered sum and norm	16

1 Problem Set 1: Convex Sets

1.1 Problem 2.5

Problem:

What is the distance between two parallel hyperplanes: $\{x \in \mathbb{R}^n | a^T x = b_1\}$ and $\{x \in \mathbb{R}^n | a^T x = b_2\}$?

Solution:

Under the assumption that $a \in \mathbb{R}^n$ and $b_1, b_2 \in \mathbb{R}$, the quantity $a^T x_0$ represents the component of x_0 in the normal direction. Similarly, the quantities b_1 and b_2 represent the euclidean distance of the hyperplane from the origin (in the normal direction). Since the hyperplanes are parallel, the distance between them is the difference between their offsets:

$$\text{Distance between hyperplanes: } b_1 - b_2 \quad (1)$$

1.2 Problem 2.7

Problem:

Voronoi description of halfspace. Let a and b be distinct points in \mathbb{R}^n . Show that the set of all points that are closer to a than b via the euclidean norm is a halfspace. Describe it explicitly as an inequality and draw a picture.

Solution:

The set of all points closer to a than b can be defined as:

$$\{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \|x - b\|_2\} \quad (2)$$

The boundary defining this halfspace will be a plane defined by the normal vector c representing the distance between a and b , and the offset coefficient d describing intersection of the plane through the half-way point between a and b . The quantities c and d can therefore be defined by:

$$\begin{aligned} c &= b - a \\ d &= \frac{c^T a + c^T b}{2} \\ &= \frac{1}{2} c^T (a + b) \end{aligned} \quad (3)$$

The halfspace, that is equivalent to x , can be described by the following:

$$\{x \in \mathbb{R}^n \mid c^T x \leq d\} \quad (4)$$

This can be visualized in two dimensions for $a = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $b = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$. The boundary (the red line) is calculated in the standard form using

$$x_2 = \frac{-1}{c_2} (c_1 * x_1 - d)$$

and then plotted. The half-space itself is the region below the boundary.

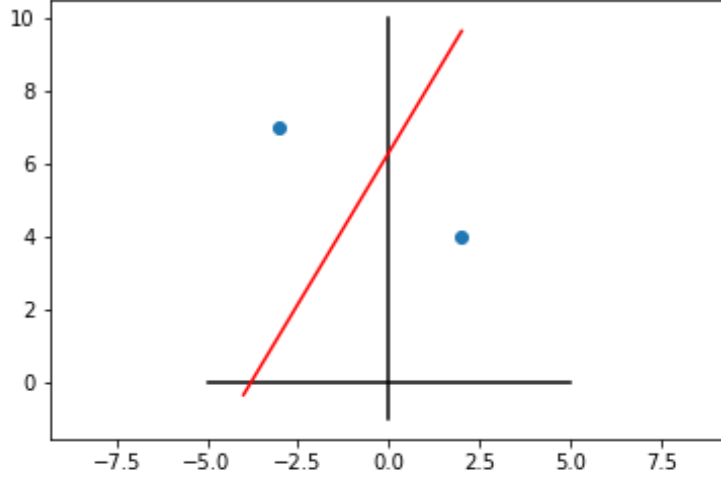


Figure 1: Visualization of the boundary for the halfspace.

1.3 Problem 2.12

Problem:

Which of the following sets are convex?

Solution:

1.3.1 (a) - Slab

A slab defined as

$$\{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq \beta\}$$

is convex as it consists of the intersection of two halfspaces which themselves are convex.

1.3.2 (b) - Rectangle/Hyperrectangle

A rectangle set defined as

$$\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, i = 1, \dots, n\}$$

is convex as it is composed of the intersections of half spaces which are themselves convex. This is similar to the polyhedrons/polytopes that by definition are also convex.

1.3.3 (c) - Wedge

A wedge set given as

$$\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1, a_2^T x \leq b_2\}$$

is convex as it is just an intersection of two halfspaces (a polyhedron).

1.3.4 (d) - Closer to a point than a set

A set of points closer to a given point than a given set is defined as

$$\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2 \ \forall y \in S\}$$

where $S \subset \mathbb{R}^n$ **is not convex** in general. This is because there is not enough information about y for a conclusion to be made whether it is convex or not. A counter example would be if y is a point in orbit around a convex shape S that would end up generating a concave x .

1.3.5 (e) - Closer to a set than another set

A set of points closer to a given set than another given set is defined as

$$\{x \mid \mathbf{dist}(x, S) \leq \mathbf{dist}(x, T)\}$$

where $S, T \subset \mathbb{R}^n$, and

$$\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}$$

is not convex in general. This is because there is not enough information about S and T for a conclusion to be made whether it is convex or not. A counter example includes if S or T themselves are a concave shape that causes the set x to also be concave and therefore not convex.

1.3.6 (f) - Set of the sum being within a convex set

The set defined as

$$\{x \mid x + S_2 \subseteq S_1\}$$

with S_1 being convex **is not convex** in general. This is because we do not know enough information about S_2 to conclude that x is convex purely due to the relationship with the convex set S_1 .

1.3.7 (g) - Set with weighted distances to two points

The set of all points that is closer to a than b by at least a factor of θ , defined as

$$\{x \in \mathbb{R}^n \mid \|x - a\|_2 \leq \theta \|x - b\|_2\}$$

with $a \neq b$ and $0 \leq \theta \leq 1$ **is convex**. This is known because, as proven in a previous problem, a hyperplane is formed for a similarity stated problem which itself is convex. When the distance to a must be less than a portion of the distance to b it will cause the pseudo-hyperplane to curve inwards and ultimately remain convex.

1.4 Problem 2.28

Problem:

Define the positive semi-definite cone (S_+^n) for $n = 1, 2, 3$ in terms of ordinary inequalities with the matrix coefficients themselves.

Solution:

The positive semi-definite cone is defined for size n as the set of all symmetric matrices that are positive semi-definite:

$$S_+^n \equiv \{x \in S^n \mid x \succeq 0\} \quad (5)$$

One method to ensure that a matrix is positive semi-definite is to ensure that its leading principle minors are all non-negative (strictly positive for positive definite).

For $n = 1$ the required inequalities are simple,

$$X = \begin{bmatrix} x_1 \end{bmatrix} \in S_+^1 \iff x_1 \geq 0 \quad (6)$$

For $n = 2$ the inequalities can be found by ensuring the leading principle minors are all non negative:

$$\begin{aligned} m_1 &= \det[x_1] \\ &= x_1 \\ m_2 &= \det \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \\ &= x_1 x_3 - x_2^2 \end{aligned} \quad (7)$$

These definitions of the minors can be then be used to construct inequalities such that all the minors are positive:

$$X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \in S_+^2 \iff \begin{aligned} x_1 &\geq 0 \\ x_1 x_3 &\geq x_2^2 \end{aligned} \quad (8)$$

For $n = 3$ the inequalities can be found by ensuring the leading principle minors are all non negative:

$$\begin{aligned} m_1 &= \det[x_1] \\ &= x_1 \\ m_2 &= \det \begin{bmatrix} x_1 & x_2 \\ x_2 & x_4 \end{bmatrix} \\ &= x_1 x_4 - x_2^2 \\ m_3 &= \det \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix} \\ &= x_1(x_1 x_4 - x_2^2) - x_2(x_2 x_6 - x_3 x_5) + x_3(x_2 x_5 - x_3 x_4) \\ &= x_1^2 x_4 - x_1 x_2^2 - x_2^2 x_6 + x_2 x_3 x_5 + x_2 x_3 x_5 - x_3^2 x_4 \end{aligned} \quad (9)$$

These definitions of the minors can be then be used to construct inequalities such that all the minors are positive:

$$X = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix} \in S_+^3 \iff \begin{aligned} x_1 &\geq 0 \\ x_1 x_4 &\geq x_2^2 \\ x_1 x_2^2 + x_2^2 x_6 + x_3^2 x_4 &\geq x_1^2 x_4 + 2x_2 x_3 x_5 \end{aligned} \quad (10)$$

1.5 Problem 2.33

The monotone non-negative cone is defined as all the nonnegative vectors with components sorted in non-increasing order:

$$K_{m+} = \{x \in \mathbb{R}^n \mid x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\} \quad (11)$$

1.5.1 Part a

Problem:

Show that K_{m+} is a proper cone.

Solution:

A set, $C \subseteq \mathbb{R}^n$, is considered a cone if

$$\theta x \in C \quad \forall x \in C \text{ and } \theta \geq 0 \quad (12)$$

It can be easily seen that the set K_{m+} satisfies this condition as scaling each element of the matrix $x \in K_{m+}$ will equally be scaled by the same amount and the conditions of non-increasing order will still apply. This guarantees that K_{m+} is in fact a cone.

To ensure convexity, the definition of convexity and of a cone can be incorporated into the following test:

The set $C \subseteq \mathbb{R}^n$ is a convex cone iff

$$\theta_1 x_1 + \theta_2 x_2 \in C \quad \forall x_1, x_2 \in C, \theta_1, \theta_2 \geq 0 \quad (13)$$

It is also clear that K_{m+} will satisfy as if each element of one of the matrices is scaled it maintains the nonincreasing order. The same is true for summing two K_{m+} matrices as the nonincreasing order will be maintained.

It is also clear that the cone is closed because the complimentary set K'_{m+} is clearly open.

Similarly, K_{m+} is solid because its definition includes all of the subspace above the cone's boundary.

The cone is also known to be pointed because the definition defines that all of the vectors contained within the set are nonnegative. It is therefore only ever possible for elements to be in the positive sector (quadrant/octant/etc.). This means the cone cannot contain a line.

We can finish stating that K_{m+} is a proper cone because it is a cone, convex, closed, solid, and pointed.

1.5.2 Part b

Problem:

Find the dual cone, K_{m+}^* .

Solution:

A dual cone for K is defined as:

$$K^* = \{y \in \mathbb{R}^n \mid x^T y \geq 0 \ \forall x \in K\} \quad (14)$$

The left side of the inequality that defines the dual cone can also be written in summation form as:

$$x^T y = \sum_{i=1}^n x_i y_i \geq 0$$

It is known that the following is equivalent:

$$\begin{aligned} \sum_{i=1}^n x_i y_i &= (x_1 - x_2)y_1 + (x_2 - x_3)(y_1 + y_2) + \cdots \\ &= +(x_{n-1} - x_n)(y_1 + \cdots + y_{n-1}) + x_n(y_1 + \cdots + y_n) \end{aligned} \quad (15)$$

By definition, $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$, thus it can be said that each $(x_i - x_j)$ term in the expanded form is positive. In order for the inequality to always hold true, each $(y_1 + \cdots + y_i)$, must be positive as well. This can be achieved by defining that the elements of y are never decreasing.

Thus, the dual cone of K_{k+} can be defined as:

$$K_{m+}^* = K_{m-} = \{x \in \mathbb{R}_+^n \mid x_1 \leq x_2 \leq \cdots \leq x_n\} \quad (16)$$

2 Problem Set 2: Convex Functions

2.1 Problem 3.6

(Note to self... remember to remove this time: Proofs don't really seem required... pretty simple question)

Problem:

What is the epigraph for the following convex functions?

Solution:

The epigraph is defined as

$$\mathbf{epi}f = \{(x, t) \mid x \in \mathbf{dom}f, f(x) \leq t\} \quad (17)$$

where $\mathbf{epi}f \subset R^{n+1}$. This is equivalent to saying the space 'above' the function.

2.1.1 (a) - Epigraph of a halfspace

The epigraph of a halfspace in n dimensions is a halfspace in $n + 1$ dimensions.

2.1.2 (b) - Epigraph of a convex cone

The epigraph of a convex cone is another convex cone of in a higher dimension. (It might be possible to generalize it to a triangular slab type thing)

2.1.3 (c) - Epigraph of a polyhedron

The epigraph of a polyhedron is another polyhedron.

2.2 Problem 3.16

Problem:

Determine if the following functions are convex, concave, quasiconvex, or quasiconcave.

Solution:

2.2.1 (a) - $e^x - 1$

Let

$$f(x) = e^x - 1$$

on \mathbb{R} .

$f(x)$ is **convex** and can be proven in multiple ways. For instance, visualizing the epigraph of the function, it is clear that a convex set is produced. Additionally, $f(x)$ can be constructed by putting the convex function e^x through the convex affine function $x - 1$.

2.2.2 (b) - x_1x_2

Let

$$f(x_1, x_2) = x_1x_2$$

over the domain \mathbb{R}_{++}^2 .

This function is **convex** and can be demonstrated using the definition:

$$f(\theta x + (1 - \theta)y) = \theta f(x) + (1 - \theta)f(y) \quad (18)$$

$$(\theta x_1 + (1 - \theta)y_1)(\theta x_2 + (1 - \theta)y_2) = \theta x_1x_2 + (1 - \theta)y_1y_2$$

$$\theta^2 x_1x_2 + \theta(1 - \theta)y_1x_2 + \theta(1 - \theta)x_1y_2 + (1 - \theta)^2 y_1y_2 = \theta x_1x_2 + (1 - \theta)y_1y_2$$

$$\theta^2 x_1x_2 + (\theta - \theta^2)(x_1y_2 + x_2y_1) + (1 - 2\theta + \theta^2)y_1y_2 = \theta x_1x_2 + (1 - \theta)y_1y_2 \quad (19)$$

If we analyze each set of terms it not explicitly clear by the result, but within the domain \mathbb{R}_{++}^2 the equality holds true.

2.2.3 (c) - $1/(x_1x_2)$

Let

$$f(x_1, x_2) = 1/(x_1x_2)$$

be defined over the domain \mathbb{R}_{++}^2 .

This function is **convex**.

2.2.4 (d) - x_1/x_2

Let

$$f(x_1, x_2) = x_1/x_2$$

be defined over the domain \mathbb{R}_{++}^2 .

This function is **convex**.

2.3 (e) - x_1^2/x_2

Let

$$f(x_1, x_2) = x_1^2/x_2$$

be defined over the domain $R \times R_+$. This function **is convex**.

2.4 (f) - $x_1^\alpha x_2^{1-\alpha}$

Let

$$f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$$

with $0 \leq \alpha \leq 1$ be defined over the domain \mathbb{R}_{++}^2 .

This function **is not convex**.

2.5 Problem 3.18a

Problem:

Using the proof of concavity of the log-determinant function to show that

$$f(X) = \text{tr}(X^{-1})$$

is convex over the domain S_{++}^n .

Solution:

A proof of the concavity for the

$$f(X) = \log \det X$$

is given as follows:

Given an arbitrary line, $X = Z + tV$, with $Z, V \in S^n$, the function

$$g(t) = f(Z + tV)$$

can be defined within $\{x \mid Z + tV \succ 0\}$. It can then be assumed (without loss of generalizty) that $t = 0$ is in the interval (so that $Z \succ 0$ is defined). This then allows the following:

$$g(t) = \log \det(Z + tV) \tag{20}$$

$$= \log \det \left(Z^{1/2} (I + tZ^{-1/2}VZ^{-1/2}) Z^{1/2} \right) \tag{21}$$

$$= \sum_{i=1}^n \log(1 + t\lambda_i) + \log \det Z \tag{22}$$

with $\lambda_1, \dots, \lambda_n$ being the eigenvalues of $Z^{-1/2}VZ^{-1/2}$.

This allows for the first and second derivatives to be computed as:

$$g'(t) = \sum_{i=1}^n \frac{\lambda_i}{1 + t\lambda_i} \tag{23}$$

$$g''(t) = - \sum_{i=1}^n \frac{\lambda_i^2}{(1 + t\lambda_i)^2} \tag{24}$$

Since $g''(t) \leq 0$, it can be concluded that f is concave.

This proof can then be applied to the trace(X-1) is also concave.

First, Let

$$f(X) = \text{tr}(X^{-1}) \tag{25}$$

be defined on the domain S_{++}^n .

An arbitrary line can then be defined as

$$g(t) = f(Z + tV) \tag{26}$$

over the domain $\{x \mid Z + tV \succ 0\}$. It can then be assumed (without loss of generality) that $t = 0$ is in the interval (so that $Z \succ 0$ is defined). This then allows the following:

$$g(t) = \text{tr} \left((Z + tV)^{-1} \right) \tag{27}$$

$$= \text{tr} \left(\left(Z^{1/2} (I + tZ^{-1/2} V Z^{-1/2}) Z^{1/2} \right)^{-1} \right) \quad (28)$$

from the fact that $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$, this can be manipulated to be

$$= \text{tr} \left(Z^{-1/2} \left(I + tZ^{-1/2} V Z^{-1/2} \right)^{-1} Z^{-1/2} \right) \quad (29)$$

since $\text{tr}(ABC) = \text{tr}(CAB)$,

$$= \text{tr} \left(Z^{-1/2} Z^{-1/2} \left(I + tZ^{-1/2} V Z^{-1/2} \right)^{-1} \right) \quad (30)$$

by taking the eigenfactor decomposition of $Z^{-1/2} V Z^{-1/2} = Q \Lambda Q^{-1}$, this can be rewritten as:

$$= \text{tr} \left(Z^{-1} (I + tQ \Lambda Q^{-1})^{-1} \right) \quad (31)$$

$$= \text{tr} \left(Z^{-1} (Q(I + t\Lambda)Q^{-1})^{-1} \right) \quad (32)$$

$$= \text{tr} \left(Z^{-1} Q(I + t\Lambda)^{-1} Q^{-1} \right) \quad (33)$$

since $\text{tr}(ABC) = \text{tr}(CAB)$,

$$= \text{tr} \left(Q Z^{-1} Q^{-1} (I + t\Lambda)^{-1} \right) \quad (34)$$

By the definition of the trace, we can rewrite this as

$$= \sum_{i=1}^n (Q^T Z^{-1} Q)_{ii} (1 + t\lambda_i)^{-1} \quad (35)$$

This allows for the first and second derivatives to be computed as:

$$g'(t) = - \sum_{i=1}^n (Q^T Z^{-1} Q)_{ii} \frac{\lambda_i}{(1 + t\lambda_i)^{-2}} \quad (36)$$

$$g''(t) = 2 \sum_{i=1}^n (Q^T Z^{-1} Q)_{ii} \frac{\lambda_i^2}{(1 + t\lambda_i)^{-3}} \quad (37)$$

Since $g''(t) \leq 0$, it can be concluded that f is concave.

2.6 Problem 3.22

Problem:

Use various composition rules to show that the following functions are convex.

Solution:

2.6.1 (a) - double log functions

Let

$$f(x) = -\log\left(-\log\left(\sum_{i=1}^m e^{a_i^T x + b_i}\right)\right)$$

be defined over the domain $\{x \mid \sum_{i=1}^m e^{a_i^T x + b_i} < 1\}$.

It is known that

$$\log\left(\sum_{i=1}^n e^{y_i}\right)$$

is convex. Since all compositions of convex functions with affine functions are also convex, it can be said that

$$\sum_{i=1}^m e^{a_i^T x + b_i}$$

is also convex.

Additionally, it is known that the $\log()$ function is concave, but when the sign changes it becomes convex, thus the composition of $-g(-g(h(x)))$ is convex for the concave function $g(x) = \log(x)$.

Therefore, the function $f(x)$ is convex.

2.6.2 (b) - square root of some product sum

Let

$$f(x, u, v) = -\sqrt{uv - x^T x}$$

be defined over the domain $\{(x, u, v) \mid uv > x^T x, u, v > 0\}$.

It is known that $g(x) = x^x/u$ is convex for $u > 0$ and that $h(x) = \sqrt{x_1 x_2}$ is convex on \mathbb{R}_{++}^2 .

The function $f(x, u, v)$ can be manipulated as follows:

$$\begin{aligned} f(x, u, v) &= -\sqrt{uv - x^T x} \\ &= -\sqrt{u\left(v - \frac{x^T x}{u}\right)} \end{aligned}$$

From what we know about the underling functions, it can be said that the convex function $g(x)$ is summed with v (a convex combination) and then used as an input to the convex function $h(x)$, resulting in an overall convex function.

2.6.3 (c) - log of some product sum

Let

$$f(x, u, v) = -\log(uv - x^T x)$$

be defined over the domain $\{(x, u, v) \mid uv > x^T x, u, v > 0\}$.

It is known that $g(x) = x^x/u$ is convex for $u > 0$ as well as that the $h(x) = \log(x)$ function is concave.

Performing the same manipulation as in the previous problem, $f(x, u, v)$ can be written as:

$$f(x, u, v) = -\log\left(u\left(v - \frac{x^T x}{u}\right)\right)$$

From this it can be derived that the convex function $g(x)$ is summed with v (a convex combination) and then used as an input to the concave function $h(x)$ but is then negated to result in an overall convex function.

2.6.4 (d) - complicated root of a powered sum and norm

Let

$$f(x, t) = -\left(t^p - \|x\|_p^p\right)^{1/p}$$

be defined with $p > 1$ over the domain $\{(x, t) \mid t \geq \|x\|_p\}$.

It is known that $g(x, u) = \|x\|_p^p/u^{p-1}$ is convex for $u > 0$ and that $h(x, y) = -x^{1/p}y^{1-1/p}$ is convex over \Re_{++}^2 .

$f(x, t)$ can be manipulated as follows:

$$f(x, t) = -\left(t^p - \|x\|_p^p\right)^{1/p} \tag{38}$$

$$= -\left(t^{p-1}\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)\right)^{1/p} \tag{39}$$

$$= -t^{1-1/p}\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)^{1/p} \tag{40}$$

$$\tag{41}$$

From this it is clear that the known convex function $g(x, t)$ is put through a weighted sum (which over the domain is never negative) and then is composed with the known convex function $h(x, t)$. This means that the original function $f(x, t)$ is a convex function.

2.6.5 (e) - complicated log of a powered sum and norm

Let

$$f(x, t) = -\log(t^p - \|x\|_p^p)$$

with $p > 1$ be defined over the domain $\{(x, t) \mid t \geq \|x\|_p\}$.

It is known that $g(x, u) = \|x\|_p^p / u^{p-1}$ is convex for $u > 0$ and that (with the negative sign) the function $h(x) = -\log(x)$ is convex. $f(x, t)$ can be manipulated as follows:

$$f(x, t) = -\log(t^p - \|x\|_p^p) \tag{42}$$

$$= -\log\left(t^{p-1}\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right)\right) \tag{43}$$

$$= -\log(t^{p-1}) - \log\left(t - \frac{\|x\|_p^p}{t^{p-1}}\right) \tag{44}$$

From this it is clear that the known convex function $g(x, t)$ is put through a weighted sum (which over the domain is never negative) and then is composed with the known convex function $h(x, t)$. This means that the original function $f(x, t)$ is a convex function.