EECS 127/227AT **Fall 2020**
Solutions for Discussion 4

1 Convexity of Sets

A set $C \subseteq \mathbb{R}^n$ is called *convex* if the line segment between any two points in C lies in C, i.e., for any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$, we have,

$$\theta x_1 + (1-\theta)x_2 \in C.$$

(a) Show that the following sets are convex:

i A vector subspace $\mathcal{V}$ of $\mathbb{R}^n$.

Solution:

Let $x_1, x_2 \in \mathcal{V}$ and $0 \leq \theta \leq 1$. Since $\mathcal{V}$ is closed under scalar multiplication and vector addition, we have $\theta x_1 + (1-\theta)x_2 \in \mathcal{V}$. Since this holds for all $x_1, x_2 \in \mathcal{V}$ and $0 \leq \theta \leq 1$, this establishes that $\mathcal{V}$ is convex.

ii A hyperplane $\mathcal{L}$ in $\mathbb{R}^n$, given by $\mathcal{L} := \{x \ : \ a^\top x = b\}$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

Solution:

If $b = 0$ the hyperplane $\mathcal{L}$ would be a subspace and this would be a special case of part (a)i of the problem. In general, let $x_1, x_2 \in \mathcal{L}$ and $0 \leq \theta \leq 1$. Then we have

$$a^T(\theta x_1 + (1-\theta)x_2) = \theta a^T x_1 + (1-\theta)a^T x_2 = \theta b + (1-\theta)b = b.$$

This shows that $\theta x_1 + (1-\theta)x_2 \in \mathcal{L}$. Since this holds for all $x_1, x_2 \in \mathcal{L}$ and $0 \leq \theta \leq 1$, this establishes that $\mathcal{L}$ is convex.

iii A halfspace $\mathcal{H}$ in $\mathbb{R}^n$, given by $\mathcal{H} := \{x \ : \ a^\top x \leq b\}$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

Solution:

Let $x_1, x_2 \in \mathcal{H}$ and $0 \leq \theta \leq 1$. Then we have

$$a^T(\theta x_1 + (1-\theta)x_2) = \theta a^T x_1 + (1-\theta)a^T x_2 \leq \theta b + (1-\theta)b = b.$$

This shows that $\theta x_1 + (1-\theta)x_2 \in \mathcal{H}$. Since this holds for all $x_1, x_2 \in \mathcal{H}$ and $0 \leq \theta \leq 1$, this establishes that $\mathcal{H}$ is convex.

iv A norm-ball $\mathcal{B}$ in $\mathbb{R}^n$, given by $\mathcal{B} := \{x \ : \ \|x - x_c\| \leq r\}$, where $x_c \in \mathbb{R}^n$, $r \geq 0$ and $\|\cdot\|$ is a norm on $\mathbb{R}^n$.

Solution:

Suppose $x_1 \in \mathcal{B}$, $x_2 \in \mathcal{B}$ and $\theta \in [0,1]$. Then we have

$$\begin{aligned}
\|\theta x_1 + (1-\theta)x_2 - x_c\| &\overset{(a)}{\leq} \|\theta(x_1 - x_c)\| + \|(1-\theta)(x_2 - x_c)\| \\
&\overset{(b)}{=} \theta\|x_1 - x_c\| + (1-\theta)\|x_2 - x_c\| \\
&\leq \theta r + (1-\theta)r \\
&= r,
\end{aligned}$$

where step (a) comes from the triangle inequality for norms and step (b) comes from the homogeneity of norms under scaling (and because θ and $1-\theta$ are nonnegative). This shows that $\theta x_1 + (1-\theta)x_2 \in \mathcal{B}$. Since this holds for all $x_1, x_2 \in \mathcal{B}$ and $0 \leq \theta \leq 1$, this establishes that $\mathcal{B}$ is convex.

(b) Operations that preserve convexity of sets:

i **Intersection**

1. Show that convexity is preserved under intersection, i.e. if S_1 and S_2 are convex subsets of $\mathbb{R}^n$, then $S := S_1 \cap S_2$ is convex.
   Solution:
   Let $x_1, x_2 \in S$. Then $x_1, x_2 \in S_1$ and $x_1, x_2 \in S_2$. Since S_1 and S_2 are convex we have, $\theta x_1 + (1-\theta)x_2 \in S_1$ and $\theta x_1 + (1-\theta)x_2 \in S_2$, which implies $\theta x_1 + (1-\theta)x_2 \in S$.
2. Show that a polyhedron is convex. A *polyhedron* is the solution set of a finite number of linear inequalities. For example $P := \{x \in \mathbb{R}^2 \mid x_2 \leq x_1, x_1 \geq 0, x_1 \leq 1\}$ is a polyhedron.
   Solution:
   A polyhedron is the intersection of a finite number of halfspaces/hyperplanes. Hence it is an intersection of convex sets and is thus convex.

ii **Mapping under an affine function**

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called *affine* if it is a sum of a linear function and a constant, i.e.

$$f(x) = Ax + b,$$

for $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

1. Prove that if $S \subseteq \mathbb{R}^n$ is convex, then the image of S under f,

$$f(S) := \{f(x) \ : \ x \in S\},$$

   which is a subset of $\mathbb{R}^m$, is convex.
   Solution:
   Let $y_1, y_2 \in f(S)$. This implies there exist $x_1, x_2 \in S$ such that $y_1 = Ax_1 + b$ and $y_2 = Ax_2 + b$. We want to show that $\lambda y_1 + (1-\lambda)y_2 \in f(S)$ for $0 \leq \lambda \leq 1$. We have

$$\begin{aligned}
\lambda y_1 + (1-\lambda)y_2 &= \lambda f(x_1) + (1-\lambda) f(x_2) \\
&= \lambda(A(x_1) + b) + (1-\lambda)(A(x_2) + b) \\
&= \lambda A x_1 + \lambda b + (1-\lambda)Ax_2 + (1-\lambda)b \\
&= A(\lambda x_1 + (1-\lambda)x_2) + b.
\end{aligned}$$

   Since S is convex we have $\lambda x_1 + (1-\lambda)x_2 \in S$ and therefore $\lambda y_1 + (1-\lambda)y_2 \in f(S)$, as desired.
   Example: Let $e_i, 1 \leq i \leq n$ denote the standard basis vectors in $\mathbb{R}^n$. As an example of what we have just shown, the projection of a convex set in $\mathbb{R}^n$ onto the span of any subset of the standard basis vectors is convex. E.g. if $S \subseteq \mathbb{R}^3$ is convex, then $T := \{(x_1, x_2) \in \mathbb{R}^2 \ : \ (x_1, x_2, x_3) \in S$ for some $x_3 \in \mathbb{R}\}$ is convex.
2. Show that the inverse image of a convex set $C \subseteq \mathbb{R}^m$ under the affine function $f(x) = Ax + b$ is convex. Here, the inverse image of C is defined to be the subset of $\mathbb{R}^n$ given by

$$f^{-1}(C) := \{x \ : \ f(x) \in C\}.$$

   Solution:

Suppose $x_1, x_2 \in f^{-1}(C)$. Since $x_1, x_2 \in f^{-1}(C)$ there exist $y_1, y_2 \in C$ such that $f(x_1) = y_1, f(x_2) = y_2$. To show the desired result we need to show that if $\lambda \in [0, 1]$ then we have $\lambda x_1 + (1 - \lambda)x_2 \in f^{-1}(C)$. Since C is convex, we have $\lambda y_1 + (1 - \lambda)y_2 \in C$. We will show that $f(\lambda x_1 + (1 - \lambda)x_2) = \lambda y_1 + (1 - \lambda)y_2$, which would establish that $\lambda x_1 + (1 - \lambda)x_2 \in f^{-1}(C)$, as desired. To show this, we write

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &= A(\lambda x_1 + (1 - \lambda)x_2) + b \\ &= \lambda(Ax_1 + b) + (1 - \lambda)(Ax_2 + b) \\ &= \lambda f(x_1) + (1 - \lambda)f(x_2) \\ &= \lambda y_1 + (1 - \lambda)y_2, \end{aligned}$$

thus completing the argument.

2 Convexity of Functions

Recall the following definitions and facts involving convex functions:

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *convex* if $\mathrm{dom}(f)$ is a convex set and if, for all $x, y \in \mathrm{dom}(f)$ and $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y). \tag{1}$$

  The function f is called *strictly convex* if the inequality in (1) is strict whenever $\theta \in (0,1)$ and $x \neq y$.

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *concave* if $\mathrm{dom}(f)$ is a convex set and if, for all $x, y \in \mathrm{dom}(f)$ and $0 \leq \theta \leq 1$, we have

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta) f(y).$$

  The function f is called *strictly concave* if the inequality is strict when $\theta \in (0,1)$ and $x \neq y$.

- A function f is concave if and only if $-f$ is convex and strictly concave iff $-f$ is strictly convex. An affine function is both convex and concave.

It is useful to keep in mind several alternative characterizations of convexity:

- *First order condition*
  Suppose f is differentiable at every point in its domain (in particular, for this to make sense, $\mathrm{dom}(f)$ would need to be an open set). Then f is convex if and only if we have

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

  for all $x, y \in \mathrm{dom}(f)$.

- *Second order condition*
  Suppose f is twice differentiable at every point in its domain (so, in particular, $\mathrm{dom}(f)$ will need to be an open set when one applies this criterion). Then f is convex if and only if its Hessian $\nabla^2 f(x)$ is positive semidefinite at every $x \in \mathrm{dom}(f)$.

- *Restriction to a line*
  A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex iff for all $x \in \mathrm{dom}(f)$ and all $v \in \mathbb{R}^n$, the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(t) = f(x + tv)$, with its domain defined to be $\mathrm{dom}(g) := \{t \ : \ x + tv \in \mathrm{dom}(f)\}$, is convex.

- *Epigraph is convex*
  A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its epigraph, defined as epi$(f) := \{(x,t) \ : \ x \in \mathrm{dom}(f), f(x) \leq t\}$ is a convex subset of $\mathbb{R}^{n+1}$.

(a) Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex, $x_1, x_2, \ldots, x_k \in \operatorname{dom}(f)$, and $\theta_1, \theta_2, \ldots, \theta_k \geq 0$ with $\sum_{i=1}^{k} \theta_i = 1$, then we have

$$f(\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \theta_2 f(x_2) + \cdots + \theta_k f(x_k).$$

This result is known as *Jensen's inequality* for discrete probability distributions.

Solution:

We will prove this by induction on k. The base case $k = 1$ is obvious. Suppose that the statement is true for all cases up to $k - 1$; then,

$$\begin{aligned} f(\theta_1 x_1 + \ldots, \theta_k x_k) &= f\left((1-\theta_k)\left(\sum_{i=1}^{k-1} \theta_i/(1-\theta_k) x_i\right) + \theta_k x_k\right) \\ &\leq (1-\theta_k) f\left(\left(\sum_{i=1}^{k-1} \theta_i/(1-\theta_k)\right) x_i\right) + \theta_k f(x_k) \\ &\leq (1-\theta_k) \sum_{i=1}^{k-1} (\theta_i/(1-\theta_k)) f(x_i) + \theta_k f(x_k) \\ &= \sum_{i=1}^{k} \theta_i f(x_i), \end{aligned}$$

where in the second inequality, in order to apply the inductive hypothesis, we used the fact that

$$\sum_{i=1}^{k-1} \theta_i/(1-\theta_k) = 1.$$

(b) Find a condition on the symmetric matrix $A \in \mathbb{R}^{n \times n}$ in order to ensure that $f(x) = x^\top A x$ is a convex function from $\mathbb{R}^n$ to $\mathbb{R}$.

Solution:

We have $\operatorname{dom}(f) = \mathbb{R}^n$ and f is twice differentiable at every point in its domain, with its Hessian given by $\nabla^2 f(x) = 2A$. Hence f is convex iff A is positive semidefinite.

(c) Show that the following operations preserve convexity:

i. **Nonnegative weighted sum:** If f and g are convex functions on $\mathbb{R}^n$ with domains $\operatorname{dom}(f)$ and $\operatorname{dom}(g)$ respectively, then $af + bg$ for $a, b \geq 0$ is a convex function on $\mathbb{R}^n$ with domain $\operatorname{dom}(f) \cap \operatorname{dom}(g)$, assuming $\operatorname{dom}(f) \cap \operatorname{dom}(g) \neq \emptyset$.

Solution:

For any $x, y \in \operatorname{dom}(f) \cap \operatorname{dom}(g)$ and $0 \leq \theta \leq 1$, we have

$$\begin{aligned} (af + bg)(\theta x + (1-\theta) y) &= af(\theta x + (1-\theta) y) + bg(\theta x + (1-\theta) y) \\ &\leq a(\theta f(x) + (1-\theta) f(y)) + b(\theta g(x) + (1-\theta) g(y)) \\ &= \theta(af + bg)(x) + (1-\theta)(af + bg)(y). \end{aligned}$$

Note that the inequality follows from the of convexity f and g; the nonnegativity of a and b; and the fact that x and y belong to both $\operatorname{dom}(f)$ and $\operatorname{dom}(g)$.

The reason for assuming $\operatorname{dom}(f) \cap \operatorname{dom}(g)$ is nonempty is because we require all convex functions to have nonempty domain.

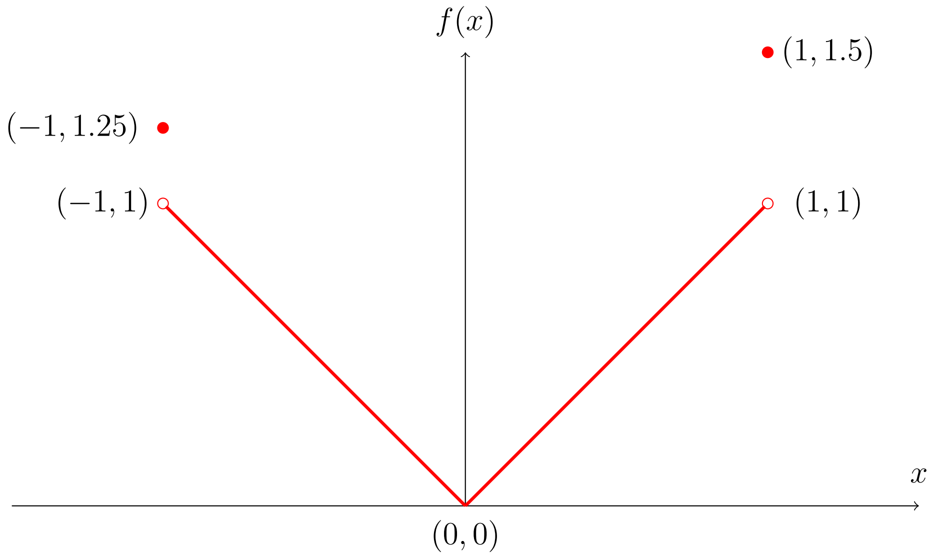

Figure 1: A graph of the function f in Problem 2(e). The function can be thought of as taking the value ∞ outside its domain, $\mathrm{dom}(f) = [-1, 1]$. Note that the function jumps at -1 and at 1 in the sense that $\lim_{x \uparrow 1} f(x) = 1 \neq f(1) = 1.5$ and $\lim_{x \downarrow -1} f(x) = 1 \neq f(-1) = 1.25$. The existence of such jumps is emphasized by the open circles and the closed circles in the diagram. With the interpretation that the function equals ∞ outside the closed interval $[-1, 1]$ we would also have $\lim_{x \downarrow 1} f(x) = \infty \neq f(1) = 1.5$ and $\lim_{x \uparrow -1} f(x) = \infty \neq f(-1) = 1.25$.

ii. **Pointwise maximum:** If f and g are convex functions on $\mathbb{R}^n$ with domains $\mathrm{dom}(f)$ and $\mathrm{dom}(g)$ respectively, then so is their pointwise maximum h, defined as

$$h(x) := \max\{f(x), g(x)\},$$

with domain $\mathrm{dom}(f) \cap \mathrm{dom}(g)$, assuming $\mathrm{dom}(f) \cap \mathrm{dom}(g) \neq \emptyset$.

Solution:

Note that the epigraph of h is just the intersection of the epigraphs of f and g. That is, the pair $(x, t) \in \mathbb{R}^n \times \mathbb{R}$ satisfies "$x \in \mathrm{dom}(h)$ and $t \geq h(x)$" if and only if it satisfies both "$x \in \mathrm{dom}(f)$ and $t \geq f(x)$" and "$x \in \mathrm{dom}(g)$ and $t \geq g(x)$". Since the intersection of convex sets is convex, we can conclude that h is a convex function by applying the epigraph condition. In more detail, because f and g are convex we have respectively that epi(f) and epi(g) are convex subsets of $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$, and since epi$(h) =$ epi$(f) \cap$ epi(g) we have that epi(h) is a convex subset of $\mathbb{R}^{n+1}$, from which we conclude that h is a convex function.

(d) Show that $f : \mathbb{S}^n_{++} \to \mathbb{R}$ defined by $f(X) = \log\det(X)$ is concave. Note that we can think of f as a function on $\mathbb{S}^n$ (which is a vector space that contains $\mathbb{S}^n_{++}$), with $\mathrm{dom}(f) = \mathbb{S}^n_{++}$.

Hint: Use the "restriction to a line" condition to check for convexity.

Solution:

Following the hint, we attempt to use the "restriction to a line" condition to check for convexity, thinking of f as a function on $\mathbb{S}^n$ with $\mathrm{dom}(f) = \mathbb{S}^n_{++}$, let $V \in \mathbb{S}^n$ and $X \in \mathbb{S}^n_{++}$, and consider $g(t) := \log\det(X + tV)$, which we think of as a function on $\mathbb{R}$ with

$$\mathrm{dom}(g) = \{t \in \mathbb{R} \ : \ X + tV \in \mathbb{S}^n_{++}\}.$$

We have

$$\begin{aligned} g(t) &:= \log\det(X+tV) \\ &= \log\det(X^{1/2}(I+X^{-1/2}tVX^{-1/2})X^{1/2}) \\ &= \log\det(X) + \log\det(I+tX^{-1/2}VX^{-1/2}). \end{aligned}$$

The last step here is actually subtle. Note that, while we can always write $\det(AB) = \det(A)\det(B)$ where AB is a matrix product of square matrices, when we write $\log\det(AB) = \log\det A + \log\det B$ for a product AB of nonsingular matrices, what we get in general is the sum of two complex numbers. However, here we are given that X is positive definite, so this step results in the sum of real numbers. Further, we need to observe that $I + tX^{-1/2}VX^{-1/2}$ is a similarity transform of $X + tV$ by a nonsingular matrix and is therefore positive definite iff $X + tV$ is positive definite. Therefore $\mathrm{dom}(g)$ can also be written as $\{t \in \mathbb{R} \,:\, I + tX^{-1/2}VX^{-1/2} \in \mathbb{S}^n_{++}\}$.

Now, note that $X^{-1/2}VX^{-1/2}$ is a symmetric matrix. Let λ_i, $1 \le i \le n$, denote the eigenvalues of $X^{-1/2}VX^{-1/2}$, which are real numbers (possibly negative). Then we have

$$g(t) = \log\det(X) + \sum_{i=1}^{n} \log(1+t\lambda_i),$$

with

$$\mathrm{dom}(g) = \{t \,:\, (1+t\lambda_i) > 0 \text{ for all } i\},$$

because

$$\{t \in \mathbb{R} \,:\, I + tX^{-1/2}VX^{-1/2} \in \mathbb{S}^n_{++}\} = \{t \,:\, (1+t\lambda_i) > 0 \text{ for all } i\}.$$

Note that $\mathrm{dom}(g)$ is quite interesting - it is the intersection of n open subsets of the real line of different types, depending on whether the corresponding eigenvalue of V is 0, strictly positive or strictly negative; indeed, if $\lambda = 0$ then $\{t \,:\, (1+t\lambda) > 0\}$ equals $\mathbb{R}$, if $\lambda > 0$ then $\{t \,:\, (1+t\lambda) > 0\}$ equals $(-1/\lambda, \infty)$, and if $\lambda < 0$ then $\{t \,:\, (1+t\lambda) > 0\}$ equals $(-\infty, -1/\lambda)$.

The function g is concave in t because it is the sum of a constant and n functions, each of which is concave on its own domain and therefore on $\mathrm{dom}(g)$ (which is the intersection of the individual domains of the component functions).

(e) Consider the function $f : \mathbb{R} \to \mathbb{R}$ with $\mathrm{dom}(f) = [-1, 1]$, given by

$$f(x) = \begin{cases} 1.5 & \text{if } x = 1, \\ |x| & \text{if } -1 < x < 1, \\ 1.25 & \text{if } x = -1. \end{cases}$$

This is illustrated in Figure 1.

Show that f is a convex function.

Solution:

Here it is not easy to apply any of the principles we have enunciated above for testing if a function is convex, so we proceed from first principles. Suppose $x_1, x_2 \in \mathrm{dom}(f)$, assuming without loss of generality that $x_1 \neq x_2$, and, without loss of generality, assuming that $\theta \in (0, 1)$. We need to check if

$$f(\theta x_1 + (1-\theta)x_2) \overset{?}{\le} \theta f(x_1) + (1-\theta) f(x_2).$$

This can be checked directly when $x_1 = -1$ and $x_2 = 1$ or vice versa. It can also be checked directly when one of x_1 and x_2 is either -1 or 1 and the other one lies in $(-1, 1)$. It therefore remains to check it when both x_1 and x_2 lie in $(-1, 1)$. But in this case we are dealing with the restriction to $(-1, 1)$ of the function $|x| = \max(x, -x)$, which is convex (as the maximum of two convex functions) when considered as having the domain $\mathbb{R}$ and will therefore continue to be convex when considered as having the domain $(-1, 1)$. This completes the proof.

One of the important take-away messages from this example is that a convex function should always be considered as being given with its domain. For example the function $|x|$ with domain $\mathbb{R}$ should, strictly speaking, be thought of as different from the function $|x|$ with domain $(-1, 1)$. Both of them are convex functions, but they are different functions. One way to keep this distinction in mind is to think of every convex function as being defined to be ∞ outside it domain of definition. It will then be obvious, for example, that $|x|$ with domain of definition $\mathbb{R}$ is a different function from $|x|$ with domain of definition $(-1, 1)$, because the former is finite outside $(-1, 1)$ while the latter equals ∞ outside $(-1, 1)$.

Another important take-away message from this example is that convex functions can be discontinuous on their domain. This has the possibility of happening only when the domain of the function has boundary points (i.e. when the interior of the domain is different from the domain), as in this example. Nevertheless, you will encounter such convex functions often enough that you should be aware that this kind of example is possible.

3 Square-to-linear function

Consider the *square-to-linear* function

$$f(x,y) := \begin{cases} \frac{x^T x}{y} & \text{if } y > 0, \\ \infty & \text{otherwise,} \end{cases}$$

defined for $(x,y) \in \mathbb{R}^n \times \mathbb{R}$, with $\text{dom}(f) = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} \;:\; y > 0\}$.

(a) Show that $\text{dom}(f)$ is a convex subset of $\mathbb{R}^{n+1}$.
Solution:
Given $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n \times \mathbb{R}$ with $y_1, y_2, > 0$ and given $\theta \in [0,1]$, we have $(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \in \mathbb{R}^n \times \mathbb{R}$ with $\theta y_1 + (1-\theta)y_2 > 0$. This proves that $\text{dom}(f)$ is a convex subset of $\mathbb{R}^{n+1}$.

(b) Show that the Hessian of f at $(x,y) \in \text{dom}(f)$ is given by

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

Solution:
We first compute the gradient $\nabla f(x,y)$ as

$$\nabla f(x,y) = \begin{bmatrix} \frac{2}{y} x \\ -\frac{x^T x}{y^2} \end{bmatrix} \in \mathbb{R}^{n+1}.$$

Viewing the gradient as a function with domain $\text{dom}(f)$ and range $\mathbb{R}^{n+1}$, we can compute its Jacobian to get the Hessian $\nabla^2 f(x,y)$. This gives

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

which is exactly the proposed formula.

(c) Show that $\nabla^2 f(x,y)$ is PSD at every $(x,y) \in \text{dom}(f)$ and thereby conclude that f is a convex function.
Solution:
One way to show this is by using Schur complements, which we have encountered in problem 4 of Homework 3, and which are discussed in Sec. 4.4.7 of the textbook of Calafiore and El Ghaoui. Since $y > 0$ the matrix $y^2 I_n$ is positive definite and so the theory of Schur complements tells us that $\begin{bmatrix} y^2 I_n & -yx \\ -yx^T & x^T x \end{bmatrix}$ is positive semidefinite iff

$$x^T x - (yx^T) \left(y^2 I_n\right)^{-1} (yx)$$

is positive semidefinite. Here, this quantity is a number (i.e. a 1×1 matrix) and equals 0, so it is positive semidefinite, and so we can conclude that $\begin{bmatrix} y^2 I_n & -yx \\ -yx^T & x^T x \end{bmatrix}$ is positive semidefinite.

We can also proceed more directly by verifying that $\begin{bmatrix} y^2 I_n & -yx \\ -yx^T & x^T x \end{bmatrix}$ satisfies the conditions required in definition of a positive semidefinite matrix. We need to show that for every $(z, w) \in \mathbb{R}^n \times \mathbb{R}$ we have

$$\frac{2}{y^3} \begin{bmatrix} z^T & w \end{bmatrix} \begin{bmatrix} y^2 I_n & -yx \\ -yx^T & x^T x \end{bmatrix} \begin{bmatrix} z \\ w \end{bmatrix} \overset{?}{\geq} 0.$$

The expression on the left hand side is

$$\frac{2}{y^3} \left(y^2 z^T z - 2ywx^T z + w^2 x^T x \right),$$

which equals $\frac{2}{y^3} v^T v$, where $v := yz - wx$, and is therefore nonnegative, as desired.