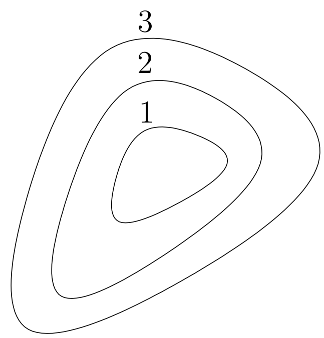
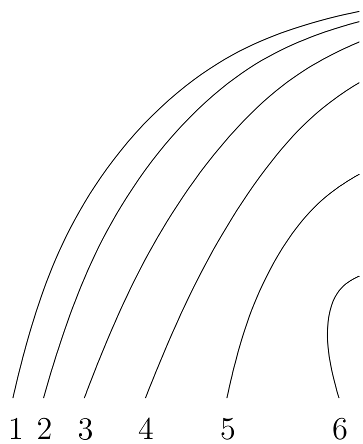
EE364a Homework 3 solutions

3.2 *Level sets of convex, concave, quasiconvex, and quasiconcave functions.* Some level sets of a function f are shown below. The curve labeled 1 shows $\{x \mid f(x) = 1\}$, etc.

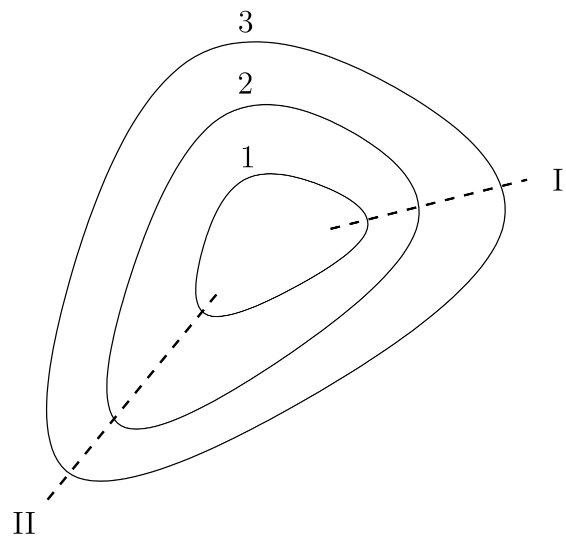


Could f be convex (concave, quasiconvex, quasiconcave)? Explain your answer. Repeat for the level curves shown below.

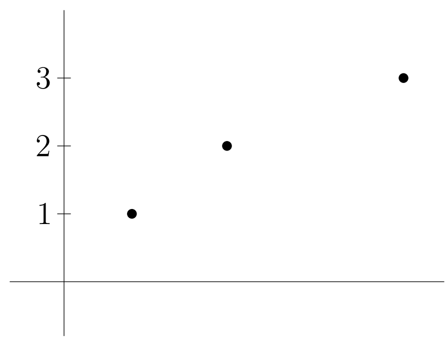


Solution. The first function could be quasiconvex because the sublevel sets appear to be convex. It is definitely not concave or quasiconcave because the superlevel sets are not convex.

It is also not convex, for the following reason. We plot the function values along the dashed line labeled I.



Along this line the function passes through the points marked as black dots in the figure below. Clearly along this line segment, the function is not convex.



If we repeat the same analysis for the second function, we see that it could be concave (and therefore it could be quasiconcave). It cannot be convex or quasiconvex, because the sublevel sets are not convex.

3.5 *Running average of a convex function.* Suppose $f : \mathbf{R} \to \mathbf{R}$ is convex, with $\mathbf{R}_+ \subseteq \mathbf{dom}\, f$. Show that its *running average* F, defined as

$$F(x) = \frac{1}{x} \int_0^x f(t)\, dt, \qquad \mathbf{dom}\, F = \mathbf{R}_{++},$$

is convex. *Hint.* For each s, $f(sx)$ is convex in x, so $\int_0^1 f(sx)\, ds$ is convex.

Solution. F is differentiable with

$$F'(x) \;=\; -(1/x^2) \int_0^x f(t)\, dt + f(x)/x$$

$$
\begin{aligned}
F''(x) &= (2/x^3)\int_0^x f(t)\,dt - 2f(x)/x^2 + f'(x)/x \\
&= (2/x^3)\int_0^x (f(t) - f(x) - f'(x)(t-x))\,dt.
\end{aligned}
$$

Convexity now follows from the fact that

$$
f(t) \geq f(x) + f'(x)(t-x)
$$

for all $x, t \in \mathbf{dom}\, f$, which implies $F''(x) \geq 0$.

Here's another proof. For each s, the function $f(sx)$ is convex in x. Therefore

$$
\int_0^1 f(sx)\; ds
$$

is a convex function of x. Now we use the variable substitution $t = sx$ to get

$$
\int_0^1 f(sx)\; ds = \frac{1}{x}\int_0^x f(t)\; dt.
$$

3.6 *Functions and epigraphs.* When is the epigraph of a function a halfspace? When is the epigraph of a function a convex cone? When is the epigraph of a function a polyhedron?

Solution. The epigraph of f is a halfspace if and only if f is affine.

The epigraph of f is a convex cone if and only if f is convex and positively homogeneous, *i.e.*, $f(\alpha x) = \alpha f(x)$ for any x and any $\alpha \geq 0$.

The epigraph of f is a polyhedron if and only if f is convex and piecewise affine.

3.16 For each of the following functions determine whether it is convex, concave, quasiconvex, or quasiconcave.

(a) $f(x) = e^x - 1$ on $\mathbf{R}$.

Solution. Strictly convex, and therefore quasiconvex. Also quasiconcave but not concave.

(b) $f(x_1, x_2) = x_1 x_2$ on $\mathbf{R}^2_{++}$.

Solution. The Hessian of f is

$$
\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},
$$

which is neither positive semidefinite nor negative semidefinite. Therefore, f is neither convex nor concave. It is quasiconcave, since its superlevel sets

$$
\{(x_1, x_2) \in \mathbf{R}^2_{++} \mid x_1 x_2 \geq \alpha\}
$$

are convex. It is not quasiconvex.

(c) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbf{R}^2_{++}$.

Solution. The Hessian of f is

$$\nabla^2 f(x) = \frac{1}{x_1 x_2} \begin{bmatrix} 2/(x_1^2) & 1/(x_1 x_2) \\ 1/(x_1 x_2) & 2/x_2^2 \end{bmatrix} \succeq 0$$

Therefore, f is convex and quasiconvex. It is not quasiconcave or concave.

(d) $f(x_1, x_2) = x_1/x_2$ on $\mathbf{R}^2_{++}$.

Solution. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{bmatrix}$$

which is not positive or negative semidefinite. Therefore, f is not convex or concave.

It is quasiconvex and quasiconcave (*i.e.*, quasilinear), since the sublevel and superlevel sets are halfspaces.

(e) $f(x_1, x_2) = x_1^2/x_2$ on $\mathbf{R} \times \mathbf{R}_{++}$.

Solution. f is convex, as mentioned on page 72. (See also figure 3.3). This is easily verified by working out the Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} 2/x_2 & -2x_1/x_2^2 \\ -2x_1/x_2^2 & 2x_1^2/x_2^3 \end{bmatrix} = (2/x_2) \begin{bmatrix} 1 \\ -x_1/x_2 \end{bmatrix} \begin{bmatrix} 1 & -x_1/x_2 \end{bmatrix} \succeq 0.$$

Therefore, f is convex and quasiconvex. It is not concave or quasiconcave (see the figure).

(f) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, on $\mathbf{R}^2_{++}$.

Solution. Concave and quasiconcave. The Hessian is

$$\begin{aligned}
\nabla^2 f(x) &= \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^\alpha x_2^{-\alpha-1} \end{bmatrix} \\
&= \alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \begin{bmatrix} -1/x_1^2 & 1/x_1x_2 \\ 1/x_1x_2 & -1/x_2^2 \end{bmatrix} \\
&= -\alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix}^T \\
&\preceq 0.
\end{aligned}$$

f is not convex or quasiconvex.

3.22 *Composition rules.* Show that the following functions are convex.

(a) $f(x) = -\log(-\log(\sum_{i=1}^m e^{a_i^T x+b_i}))$ on $\mathbf{dom}\, f = \{x \mid \sum_{i=1}^m e^{a_i^T x+b_i} < 1\}$. You can use the fact that $\log(\sum_{i=1}^n e^{y_i})$ is convex.

Solution. $g(x) = \log(\sum_{i=1}^m e^{a_i^T x+b_i})$ is convex (composition of the log-sum-exp function and an affine mapping), so $-g$ is concave. The function $h(y) = -\log y$ is convex and decreasing. Therefore $f(x) = h(-g(x))$ is convex.

(b) $f(x,u,v) = -\sqrt{uv - x^T x}$ on $\mathbf{dom}\, f = \{(x,u,v) \mid uv > x^T x,\ u,\ v > 0\}$. Use the fact that $x^T x/u$ is convex in (x,u) for $u > 0$, and that $-\sqrt{x_1 x_2}$ is convex on $\mathbf{R}^2_{++}$.

Solution. We can express f as $f(x,u,v) = -\sqrt{u(v - x^T x/u)}$. The function $h(x_1,x_2) = -\sqrt{x_1 x_2}$ is convex on $\mathbf{R}^2_{++}$, and decreasing in each argument. The functions $g_1(u,v,x) = u$ and $g_2(u,v,x) = v - x^T x/u$ are concave. Therefore $f(u,v,x) = h(g(u,v,x))$ is convex.

(c) $f(x,u,v) = -\log(uv - x^T x)$ on $\mathbf{dom}\, f = \{(x,u,v) \mid uv > x^T x,\ u,\ v > 0\}$.

Solution. We can express f as

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

The first term is convex. The function $v - x^T x/u$ is concave because v is linear and $x^T x/u$ is convex on $\{(x,u) \mid u > 0\}$. Therefore the second term in f is convex: it is the composition of a convex decreasing function $-\log t$ and a concave function.

(d) $f(x,t) = -(t^p - \|x\|_p^p)^{1/p}$ where $p > 1$ and $\mathbf{dom}\, f = \{(x,t) \mid t \geq \|x\|_p\}$. You can use the fact that $\|x\|_p^p/u^{p-1}$ is convex in (x,u) for $u > 0$ (see exercise 3.23), and that $-x^{1/p}y^{1-1/p}$ is convex on $\mathbf{R}^2_+$ (see exercise 3.16).

Solution. We can express f as

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

This is the composition of $h(y_1,y_2) = -y_1^{1-1/p}y_2^{1/p}$ (convex and decreasing in each argument) and two concave functions

$$g_1(x,t) = t, \qquad g_1(x,t) = t - \frac{\|x\|_p^p}{t^{p-1}}.$$

(e) $f(x,t) = -\log(t^p - \|x\|_p^p)$ where $p > 1$ and $\mathbf{dom}\, f = \{(x,t) \mid t > \|x\|_p\}$. You can use the fact that $\|x\|_p^p/u^{p-1}$ is convex in (x,u) for $u > 0$.

Solution. Express f as

$$\begin{aligned} f(x,t) &= -\log t^{p-1} - \log(t - \|x\|_p^p/t^{p-1}) \\ &= -(p-1)\log t - \log(t - \|x\|_p^p/t^{p-1}). \end{aligned}$$

The first term is convex. The second term is the composition of a decreasing convex function and a concave function, and is also convex.

3.49 Show that the following functions are log-concave.

(a) *Logistic function:* $f(x) = e^x/(1+e^x)$ with $\mathbf{dom}\, f = \mathbf{R}$.
Solution. We have

$$\log(e^x/(1+e^x)) = x - \log(1+e^x).$$

The first term is linear, hence concave. Since the function $\log(1+e^x)$ is convex (it is the log-sum-exp function, evaluated at $x_1 = 0$, $x_2 = x$), the second term above is concave. Thus, $e^x/(1+e^x)$ is log-concave.

(b) *Harmonic mean:*

$$f(x) = \frac{1}{1/x_1 + \cdots + 1/x_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}.$$

Solution. The first and second derivatives of

$$h(x) = \log f(x) = -\log(1/x_1 + \cdots + 1/x_n)$$

are

$$\begin{aligned} \frac{\partial h(x)}{\partial x_i} &= \frac{1/x_i^2}{1/x_1 + \cdots + 1/x_n} \\ \frac{\partial^2 h(x)}{\partial x_i^2} &= \frac{-2/x_i^3}{1/x_1 + \cdots + 1/x_n} + \frac{1/x_i^4}{(1/x_1 + \cdots + 1/x_n)^2} \\ \frac{\partial^2 h(x)}{\partial x_i \partial x_j} &= \frac{1/(x_i^2 x_j^2)}{(1/x_1 + \cdots + 1/x_n)^2} \qquad (i \neq j). \end{aligned}$$

We show that $y^T \nabla^2 h(x) y \prec 0$ for all $y \neq 0$, *i.e.*,

$$(\sum_{i=1}^n y_i/x_i^2)^2 < 2(\sum_{i=1}^n 1/x_i)(\sum_{i=1}^n y_i^2/x_i^3)$$

This follows from the Cauchy-Schwarz inequality $(a^T b)^2 \leq \|a\|_2^2 \|b\|_2^2$, applied to

$$a_i = \frac{1}{\sqrt{x_i}}, \qquad b_i = \frac{y_i}{x_i \sqrt{x_i}}.$$

A2.2 *A general vector composition rule.* Suppose

$$f(x) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

where $h : \mathbf{R}^k \to \mathbf{R}$ is convex, and $g_i : \mathbf{R}^n \to \mathbf{R}$. Suppose that for each i, one of the following holds:

- h is nondecreasing in the ith argument, and g_i is convex
- h is nonincreasing in the ith argument, and g_i is concave
- g_i is affine.

Show that f is convex. (This composition rule subsumes all the ones given in the book, and is the one used in software systems such as CVX.) You can assume that $\mathbf{dom}\, h = \mathbf{R}^k$; the result also holds in the general case when the monotonicity conditions listed above are imposed on $\tilde{h}$, the extended-valued extension of h.

Solution. Fix x, y, and $\theta \in [0,1]$, and let $z = \theta x + (1-\theta)y$. Let's re-arrange the indexes so that g_i is affine for $i = 1, \ldots, p$, g_i is convex for $i = p+1, \ldots, q$, and g_i is concave for $i = q+1, \ldots, k$. Therefore we have

$$\begin{array}{ll} g_i(z) = \theta g_i(x) + (1-\theta) g_i(y), & i = 1, \ldots, p, \\ g_i(z) \leq \theta g_i(x) + (1-\theta) g_i(y), & i = p+1, \ldots, q, \\ g_i(z) \geq \theta g_i(x) + (1-\theta) g_i(y), & i = q+1, \ldots, k. \end{array}$$

We then have

$$\begin{array}{rcl} f(z) & = & h(g_1(z), g_2(z), \ldots, g_k(z)) \\ & \leq & h(\theta g_1(x) + (1-\theta) g_1(y), \ldots, \theta g_k(x) + (1-\theta) g_k(y)) \\ & \leq & \theta h(g_1(x), \ldots, g_k(x)) + (1-\theta) h(g_1(y), \ldots, g_k(y)) \\ & = & \theta f(x) + (1-\theta) f(y). \end{array}$$

The second line holds since, for $i = p+1, \ldots, q$, we have increased the ith argument of h, which is (by assumption) nondecreasing in the ith argument, and for $i = q+1, \ldots, k$, we have decreased the ith argument, and h is nonincreasing in these arguments. The third line follows from convexity of h.

A10.2 *Schur complements.* Consider a matrix $X = X^T \in \mathbf{R}^{n \times n}$ partitioned as

$$X = \left[\begin{array}{cc} A & B \\ B^T & C \end{array} \right],$$

where $A \in \mathbf{R}^{k \times k}$. If $\det A \neq 0$, the matrix $S = C - B^T A^{-1} B$ is called the *Schur complement* of A in X. Schur complements arise in many situations and appear in many important formulas and theorems. For example, we have $\det X = \det A \det S$. (You don't have to prove this.)

(a) The Schur complement arises when you minimize a quadratic form over some of the variables. Let $f(u,v) = (u,v)^T X (u,v)$, where $u \in \mathbf{R}^k$. Let $g(v)$ be the minimum value of f over u, *i.e.*, $g(v) = \inf_u f(u,v)$. Of course $g(v)$ can be $-\infty$. Show that if $A \succ 0$, we have $g(v) = v^T S v$.

(b) The Schur complement arises in several characterizations of positive definiteness or semidefiniteness of a block matrix. As examples we have the following three theorems:

- $X \succ 0$ if and only if $A \succ 0$ and $S \succ 0$.
- If $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.
- $X \succeq 0$ if and only if $A \succeq 0$, $B^T(I - AA^\dagger) = 0$ and $C - B^T A^\dagger B \succeq 0$, where $A^\dagger$ is the pseudo-inverse of A. ($C - B^T A^\dagger B$ serves as a generalization of the Schur complement in the case where A is positive semidefinite but singular.)

Prove *one* of these theorems. (You can choose which one.)

Solution.

(a) *If $A \succ 0$, then $g(v) = v^T S v$.*
We have $f(u, v) = u^T A u + 2v^T B u + v^T C v$. If $A \succ 0$, we can minimize f over u by setting the gradient with respect to u equal to zero. We obtain $u^\star(v) = -A^{-1}Bv$, and

$$g(v) = f(u^\star(v), v) = v^T (C - B^T A^{-1} B) v = v^T S v.$$

(b) *Positive definite and semidefinite block matrices.*

- *$X \succ 0$ if and only if $A \succ 0$ and $S \succ 0$.*
Suppose $X \succ 0$. Then $f(u, v) > 0$ for all non-zero (u, v), and in particular, $f(u, 0) = u^T A u > 0$ for all non-zero u (hence, $A \succ 0$), and $f(-A^{-1}Bv, v) = v^T (C - B^T A^{-1} B) v > 0$ (hence, $S = C - B^T A^{-1} B \succ 0$). This proves the 'only if' part.
To prove the "if" part, we have to show that if $A \succ 0$ and $S \succ 0$, then $f(u, v) > 0$ for all nonzero (u, v) (that is, for all u, v that are not both zero). If $v \neq 0$, then it follows from (a) that

$$f(u, v) \geq \inf_u f(u, v) = v^T S v > 0.$$

If $v = 0$ and $u \neq 0$, $f(u, 0) = u^T A u > 0$.
- *If $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.*
From part (a) we know that if $A \succ 0$, then $\inf_u f(u, v) = v^T S v$. If $S \succeq 0$, then

$$f(u, v) \geq \inf_u f(u, v) = v^T S v \geq 0$$

for all u, v, and hence $X \succeq 0$. This proves the 'if'-part.
To prove the 'only if'-part we note that $f(u, v) \geq 0$ for all (u, v) implies that $\inf_u f(u, v) \geq 0$ for all v, *i.e.*, $S \succeq 0$.
- *$X \succeq 0$ if and only if $A \succeq 0$, $B^T(I - AA^\dagger) = 0$, $C - B^T A^\dagger B \succeq 0$.*

Suppose $A \in \mathbf{R}^{k\times k}$ with $\mathbf{rank}(A) = r$. Then there exist matrices $Q_1 \in \mathbf{R}^{k\times r}$, $Q_2 \in \mathbf{R}^{k\times(k-r)}$ and an invertible diagonal matrix $\Lambda \in \mathbf{R}^{r\times r}$ such that

$$A = \left[\begin{array}{cc} Q_1 & Q_2 \end{array}\right] \left[\begin{array}{cc} \Lambda & 0 \\ 0 & 0 \end{array}\right] \left[\begin{array}{cc} Q_1 & Q_2 \end{array}\right]^T,$$

and $[Q_1\ Q_2]^T[Q_1\ Q_2] = I$. The matrix

$$\left[\begin{array}{ccc} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{array}\right] \in \mathbf{R}^{n\times n}$$

is nonsingular, and therefore

$$\begin{aligned}
\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right] \succeq 0 \quad &\Longleftrightarrow \quad \left[\begin{array}{ccc} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{array}\right]^T \left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right] \left[\begin{array}{ccc} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{array}\right] \succeq 0 \\
&\Longleftrightarrow \quad \left[\begin{array}{ccc} \Lambda & 0 & Q_1^T B \\ 0 & 0 & Q_2^T B \\ B^T Q_1 & B^T Q_2 & C \end{array}\right] \succeq 0 \\
&\Longleftrightarrow \quad Q_2^T B = 0, \left[\begin{array}{cc} \Lambda & Q_1^T B \\ B^T Q_1 & C \end{array}\right] \succeq 0.
\end{aligned}$$

We have $\Lambda \succ 0$ if and only if $A \succeq 0$. It can be verified that

$$A^\dagger = Q_1 \Lambda^{-1} Q_1^T, \quad I - AA^\dagger = Q_2 Q_2^T.$$

Therefore

$$Q_2^T B = 0 \Longleftrightarrow Q_2 Q_2^T B = (I - A^\dagger A)B = 0.$$

Moreover, since Λ is invertible,

$$\left[\begin{array}{cc} \Lambda & Q_1^T B \\ B^T Q_1 & C \end{array}\right] \succeq 0 \Longleftrightarrow \Lambda \succ 0, \quad C - B^T Q_1 \Lambda^{-1} Q_1^T B = C - B^T A^\dagger B \succeq 0.$$