EE364a Homework 2 solutions

2.28 *Positive semidefinite cone for* $n = 1,\ 2,\ 3$. Give an explicit description of the positive semidefinite cone $\mathbf{S}^n_+$, in terms of the matrix coefficients and ordinary inequalities, for $n = 1,\ 2,\ 3$. To describe a general element of $\mathbf{S}^n$, for $n = 1,\ 2,\ 3$, use the notation

$$x_1, \qquad \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}, \qquad \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix}.$$

Solution. For $n = 1$ the condition is $x_1 \geq 0$. For $n = 2$ the condition is

$$x_1 \geq 0, \qquad x_3 \geq 0, \qquad x_1x_3 - x_2^2 \geq 0.$$

For $n = 3$ the condition is

$$x_1 \geq 0, \qquad x_4 \geq 0, \qquad x_6 \geq 0, \qquad x_1x_4 - x_2^2 \geq 0, \qquad x_4x_6 - x_5^2 \geq 0, \qquad x_1x_6 - x_3^2 \geq 0$$

and

$$x_1x_4x_6 + 2x_2x_3x_5 - x_1x_5^2 - x_6x_2^2 - x_4x_3^2 \geq 0,$$

i.e., all principal minors must be nonnegative.

We give the proof for $n = 3$, assuming the result is true for $n = 2$. The matrix

$$X = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_4 & x_5 \\ x_3 & x_5 & x_6 \end{bmatrix}$$

is positive semidefinite if and only if

$$z^TXz = x_1z_1^2 + 2x_2z_1z_2 + 2x_3z_1z_3 + x_4z_2^2 + 2x_5z_2z_3 + x_6z_3^2 \geq 0$$

for all z.

If $x_1 = 0$, we must have $x_2 = x_3 = 0$, so $X \succeq 0$ if and only if

$$\begin{bmatrix} x_4 & x_5 \\ x_5 & x_6 \end{bmatrix} \succeq 0.$$

Applying the result for the 2×2-case, we conclude that if $x_1 = 0$, $X \succeq 0$ if and only if

$$x_2 = x_3 = 0, \qquad x_4 \geq 0, \qquad x_6 \geq 0, \qquad x_4x_6 - x_5^2 \geq 0.$$

Now assume $x_1 \neq 0$. We have

$$z^TXz = x_1(z_1 + (x_2/x_1)z_2 + (x_3/x_1)z_3)^2 + (x_4 - x_2^2/x_1)z_2^2 + (x_6 - x_3^2/x_1)z_3^2 + 2(x_5 - x_2x_3/x_1)z_2z_3,$$

so it is clear that we must have $x_1 > 0$ and

$$\begin{bmatrix} x_4 - x_2^2/x_1 & x_5 - x_2x_3/x_1 \\ x_5 - x_2x_3/x_1 & x_6 - x_3^2/x_1 \end{bmatrix} \succeq 0.$$

By the result for 2×2-case studied above, this is equivalent to

$$x_1x_4 - x_2^2 \geq 0, \qquad x_1x_6 - x_3^2 \geq 0, \qquad (x_4 - x_2^2/x_1)(x_6 - x_3^2/x_1) - (x_5 - x_2x_3/x_1)^2 \geq 0.$$

The third inequality simplifies to

$$(x_1x_4x_6 + 2x_2x_3x_5 - x_1x_5^2 - x_6x_2^2 - x_4x_3^2)/x_1 \geq 0.$$

Therefore, if $x_1 > 0$, then $X \succeq 0$ if and only if

$$x_1x_4 - x_2^2 \geq 0, \qquad x_1x_6 - x_3^2 \geq 0, \qquad (x_1x_4x_6 + 2x_2x_3x_5 - x_1x_5^2 - x_6x_2^2 - x_4x_3^2)/x_1 \geq 0.$$

We can combine the conditions for $x_1 = 0$ and $x_1 > 0$ by saying that all 7 principal minors must be nonnegative.

2.33 *The monotone nonnegative cone.* We define the *monotone nonnegative cone* as

$$K_{\mathrm{m+}} = \{x \in \mathbf{R}^n \mid x_1 \geq x_2 \geq \cdots \geq x_n \geq 0\}.$$

i.e., all nonnegative vectors with components sorted in nonincreasing order.

(a) Show that $K_{\mathrm{m+}}$ is a proper cone.

(b) Find the dual cone $K_{\mathrm{m+}}^*$. *Hint.* Use the identity

$$\begin{aligned} \sum_{i=1}^n x_iy_i &= (x_1 - x_2)y_1 + (x_2 - x_3)(y_1 + y_2) + (x_3 - x_4)(y_1 + y_2 + y_3) + \cdots \\ &\quad + (x_{n-1} - x_n)(y_1 + \cdots + y_{n-1}) + x_n(y_1 + \cdots + y_n). \end{aligned}$$

Solution.

(a) The set $K_{\mathrm{m+}}$ is defined by n homogeneous linear inequalities, hence it is a closed (polyhedral) cone.

The interior of $K_{\mathrm{m+}}$ is nonempty, because there are points that satisfy the inequalities with strict inequality, for example, $x = (n, n-1, n-2, \ldots, 1)$.

To show that $K_{\mathrm{m+}}$ is pointed, we note that if $x \in K_{\mathrm{m+}}$, then $-x \in K_{\mathrm{m+}}$ only if $x = 0$. This implies that the cone does not contain an entire line.

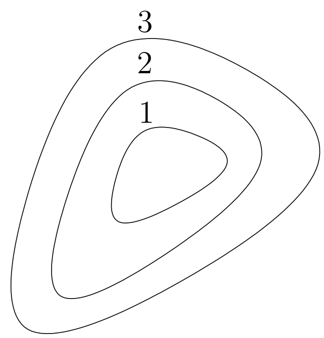
(b) Using the hint, we see that $y^Tx \geq 0$ for all $x \in K_{\mathrm{m+}}$ if and only if

$$y_1 \geq 0, \qquad y_1 + y_2 \geq 0, \quad \ldots, y_1 + y_2 + \cdots + y_n \geq 0.$$

Therefore

$$K_{\mathrm{m+}}^* = \{y \mid \sum_{i=1}^k y_i \geq 0,\ k = 1, \ldots, n\}.$$

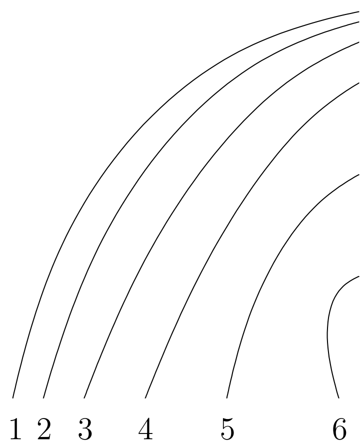
3.2 *Level sets of convex, concave, quasiconvex, and quasiconcave functions.* Some level sets of a function f are shown below. The curve labeled 1 shows $\{x \mid f(x) = 1\}$, etc.



3

1 2 3 4

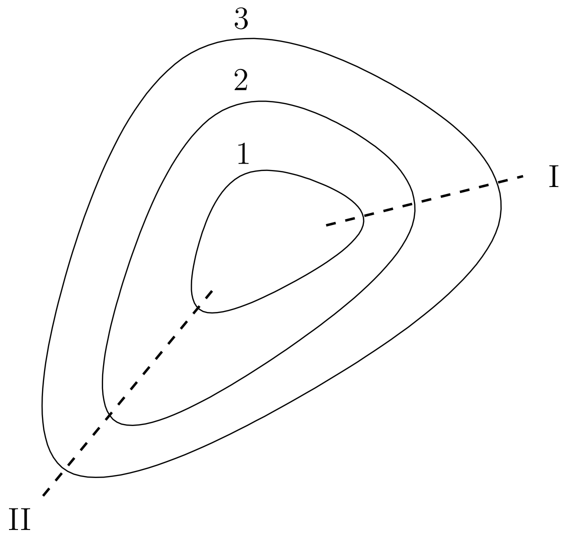
Could f be convex (concave, quasiconvex, quasiconcave)? Explain your answer. Repeat for the level curves shown below.



2 4

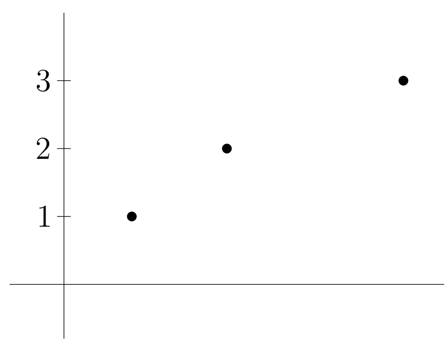
Solution. The first function could be quasiconvex because the sublevel sets appear to be convex. It is definitely not concave or quasiconcave because the superlevel sets are not convex.

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



$f'(x) \uparrow$

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. You can assume f is differentiable.

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 (simpler?) 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. If the function is convex, and it is affine, positively homogeneous ($f(\alpha x) = \alpha f(x)$ for $\alpha \geq 0$), and piecewise-affine, respectively.

3.15 *A family of concave utility functions.* For $0 < \alpha \leq 1$ let

$$
u_\alpha(x) = \frac{x^\alpha - 1}{\alpha},
$$

with $\mathbf{dom}\, u_\alpha = \mathbf{R}_+$. We also define $u_0(x) = \log x$ (with $\mathbf{dom}\, u_0 = \mathbf{R}_{++}$).

(a) Show that for $x > 0$, $u_0(x) = \lim_{\alpha\to 0} u_\alpha(x)$.

(b) Show that u_α are concave, monotone increasing, and all satisfy $u_\alpha(1) = 0$.

These functions are often used in economics to model the benefit or utility of some quantity of goods or money. Concavity of u_α means that the marginal utility (*i.e.*, the increase in utility obtained for a fixed increase in the goods) decreases as the amount of goods increases. In other words, concavity models the effect of *satiation.*

Solution.

(a) In this limit, both the numerator and denominator go to zero, so we use l'Hopital's rule:

$$
\lim_{\alpha\to 0} u_\alpha(x) = \lim_{\alpha\to 0} \frac{(d/d\alpha)(x^\alpha - 1)}{(d/d\alpha)\alpha} = \lim_{\alpha\to 0} \frac{x^\alpha \log x}{1} = \log x.
$$

(b) By inspection we have

$$u_\alpha(1) = \frac{1^\alpha - 1}{\alpha} = 0.$$

The derivative is given by

$$u'_\alpha(x) = x^{\alpha-1},$$

which is positive for all x (since $0 < \alpha < 1$), so these functions are increasing. To show concavity, we examine the second derivative:

$$u''_\alpha(x) = (\alpha - 1)x^{\alpha-2}.$$

Since this is negative for all x, we conclude that u_α is strictly concave.

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

(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).

3.18 Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following.

(b) $f(X) = (\det X)^{1/n}$ is concave on $\mathbf{dom}\, f = \mathbf{S}_{++}^n$.

Solution.

(b) Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in \mathbf{S}^n$.

$$\begin{aligned} g(t) &= (\det(Z + tV))^{1/n} \\ &= \left(\det Z^{1/2} \det(I + tZ^{-1/2} V Z^{-1/2}) \det Z^{1/2}\right)^{1/n} \\ &= (\det Z)^{1/n} \left(\prod_{i=1}^n (1 + t\lambda_i)\right)^{1/n} \end{aligned}$$

where λ_i, $i = 1, \ldots, n$, are the eigenvalues of $Z^{-1/2} V Z^{-1/2}$. From the last equality we see that g is a concave function of t on $\{t \mid Z + tV \succ 0\}$, since $\det Z > 0$ and the geometric mean $(\prod_{i=1}^n x_i)^{1/n}$ is concave on $\mathbf{R}_{++}^n$.

3.24 *Some functions on the probability simplex.* Let x be a real-valued random variable which takes values in $\{a_1, \ldots, a_n\}$ where $a_1 < a_2 < \cdots < a_n$, with $\mathbf{prob}(x = a_i) = p_i$, $i = 1, \ldots, n$. For each of the following functions of p (on the probability simplex $\{p \in \mathbf{R}_+^n \mid \mathbf{1}^T p = 1\}$), determine if the function is convex, concave, quasiconvex, or quasiconcave.

(f) $\mathbf{quartile}(x) = \inf\{\beta \mid \mathbf{prob}(x \leq \beta) \geq 0.25\}$.

Solution. The sublevel and the superlevel sets of $\mathbf{quartile}(x)$ are convex (see problem 2.15), so it is quasiconvex and quasiconcave.

$\mathbf{quartile}(x)$ is not continuous (it takes values in a discrete set $\{a_1, \ldots, a_n\}$, so it is not convex or concave. (A convex or a concave function is always continuous on the relative interior of its domain.)

(g) The cardinality of the smallest set $\mathcal{A} \subseteq \{a_1, \ldots, a_n\}$ with probability $\geq 90\%$. (By cardinality we mean the number of elements in $\mathcal{A}$.)

Solution. f is integer-valued, so it can not be convex or concave. (A convex or a concave function is always continuous on the relative interior of its domain.)

f is quasiconcave because its superlevel sets are convex. We have $f(p) \geq \alpha$ if and only if

$$\sum_{i=1}^{k} p_{[i]} < 0.9,$$

where $k = \max\{i = 1, \ldots, n \mid i < \alpha\}$ is the largest integer less than α, and $p_{[i]}$ is the ith largest component of p. We know that $\sum_{i=1}^{k} p_{[i]}$ is a convex function of p, so the inequality $\sum_{i=1}^{k} p_{[i]} < 0.9$ defines a convex set.

In general, $f(p)$ is not quasiconvex. For example, we can take $n = 2$, $a_1 = 0$ and $a_2 = 1$, and $p^1 = (0.1, 0.9)$ and $p^2 = (0.9, 0.1)$. Then $f(p^1) = f(p^2) = 1$, but $f((p^1 + p^2)/2) = f(0.5, 0.5) = 2$.

(h) The minimum width interval that contains 90% of the probability, *i.e.*,

$$\inf \{\beta - \alpha \mid \mathbf{prob}(\alpha \leq x \leq \beta) \geq 0.9\}.$$

Solution. The minimum width interval that contains 90% of the probability must be of the form $[a_i, a_j]$ with $1 \leq i \leq j \leq n$, because

$$\mathbf{prob}(\alpha \leq x \leq \beta) = \sum_{k=i}^{j} p_k = \mathbf{prob}(a_i \leq x \leq a_j)$$

where $i = \min\{k \mid a_k \geq \alpha\}$, and $j = \max\{k \mid a_k \leq \beta\}$.

We show that the function is quasiconcave. We have $f(p) \geq \gamma$ if and only if all intervals of width less than γ have a probability less than 90%,

$$\sum_{k=i}^{j} p_k < 0.9$$

for all i, j that satisfy $a_j - a_i < \gamma$. This defines a convex set.

Since the function takes values on a finite set, it is not continuous and therefore neither convex nor concave. In addition it is not quasiconvex in general. Consider the example with $n = 2$, $a_1 = 0$, $a_2 = 1$, $p^1 = (0.95, 0.05)$ and $p^2 = (0.05, 0.95)$. Then $f(p^1) = 0$, $f(p^2) = 0$, but $f((p^1 + p^2)/2) = 1$.

3.36 Derive the conjugates of the following functions.

(a) *Max function.* $f(x) = \max_{i=1,\ldots,n} x_i$ on $\mathbf{R}^n$.

Solution. We will show that

$$f^*(y) = \begin{cases} 0 & \text{if } y \succeq 0, \quad \mathbf{1}^T y = 1 \\ \infty & \text{otherwise.} \end{cases}$$

We first verify the domain of f^*. First suppose y has a negative component, say $y_k < 0$. If we choose a vector x with $x_k = -t$, $x_i = 0$ for $i \neq k$, and let t go to infinity, we see that

$$x^T y - \max_i x_i = -t y_k \to \infty,$$

so y is not in $\mathbf{dom}\, f^*$. Next, assume $y \succeq 0$ but $\mathbf{1}^T y > 1$. We choose $x = t\mathbf{1}$ and let t go to infinity, to show that

$$x^T y - \max_i x_i = t\mathbf{1}^T y - t$$

is unbounded above. Similarly, when $y \succeq 0$ and $\mathbf{1}^T y < 1$, we choose $x = -t\mathbf{1}$ and let t go to infinity.

The remaining case for y is $y \succeq 0$ and $\mathbf{1}^T y = 1$. In this case we have

$$x^T y \leq \max_i x_i$$

for all x, and therefore $x^T y - \max_i x_i \leq 0$ for all x, with equality for $x = 0$. Therefore $f^*(y) = 0$.

(d) *Power function.* $f(x) = x^p$ on $\mathbf{R}_{++}$, where $p > 1$. Repeat for $p < 0$.

Solution. We'll use standard notation: we define q by the equation $1/p + 1/q = 1$, *i.e.*, $q = p/(p-1)$.

We start with the case $p > 1$. Then x^p is strictly convex on $\mathbf{R}_+$. For $y < 0$ the function $yx - x^p$ achieves its maximum for $x > 0$ at $x = 0$, so $f^*(y) = 0$. For $y > 0$ the function achieves its maximum at $x = (y/p)^{1/(p-1)}$, where it has value

$$y(y/p)^{1/(p-1)} - (y/p)^{p/(p-1)} = (p-1)(y/p)^q.$$

Therefore we have

$$f^*(y) = \begin{cases} 0 & y \leq 0 \\ (p-1)(y/p)^q & y > 0. \end{cases}$$

For $p < 0$ similar arguments show that $\mathbf{dom}\, f^* = -\mathbf{R}_+$ and $f^*(y) = \frac{p}{q}(y/p)^q$.