Convex cones and generalized inequalities

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

\[
x_1, \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \\ x_3 & x_4 & x_5 \\ x_4 & x_5 & x_6 \end{bmatrix}
\]

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

\[
x_1 \geq 0, \quad x_3 \geq 0, \quad x_1 x_3 - x_2 \geq 0.
\]

For \( n = 3 \) the condition is

\[
x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_1 x_4 - x_2^2 \geq 0, \quad x_4 x_6 - x_5^2 \geq 0, \quad x_1 x_6 - x_3^2 \geq 0
\]

and

\[
x_1 x_4 x_6 + 2x_2 x_3 x_5 - x_1 x_4^2 - x_6 x_3^2 - x_4 x_5^2 \geq 0,
\]

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

\[
z^T X z = x_1 z_1^2 + 2x_2 z_1 z_2 + 2x_3 z_1 z_3 + x_4 z_2^2 + 2x_5 z_2 z_3 + x_6 z_3^2 \geq 0
\]

for all \( z \).

If \( x_1 = 0 \), we must have \( x_2 = x_3 = 0 \), so \( X \geq 0 \) if and only if

\[
\begin{bmatrix} x_4 & x_5 \\ x_5 & x_6 \end{bmatrix} \geq 0.
\]

Applying the result for the \( 2 \times 2 \)-case, we conclude that if \( x_1 = 0 \), \( X \geq 0 \) if and only if

\[
x_2 = x_3 = 0, \quad x_4 \geq 0, \quad x_6 \geq 0, \quad x_4 x_6 - x_3^2 \geq 0.
\]

Now assume \( x_1 \neq 0 \). We have

\[
z^T X z = x_1 (z_1 + (x_2/x_1) z_2 + (x_3/x_1) z_3)^2 + (x_4 - x_2^2/x_1) z_2^2 + (x_6 - x_3^2/x_1) z_3^2 + 2(x_5 - x_2 x_3/x_1) z_2 z_3,
\]

so it is clear that we must have \( x_1 > 0 \) and

\[
\begin{bmatrix} x_4 - x_2^2/x_1 & x_5 - x_2 x_3/x_1 \\ x_5 - x_2 x_3/x_1 & x_6 - x_3^2/x_1 \end{bmatrix} \geq 0.
\]

By the result for \( 2 \times 2 \)-case studied above, this is equivalent to

\[
x_1 x_4 - x_2^2 \geq 0, \quad x_1 x_6 - x_3^2 \geq 0, \quad (x_4 - x_2^2/x_1)(x_6 - x_3^2/x_1) - (x_5 - x_2 x_3/x_1)^2 \geq 0.
\]

The third inequality simplifies to

\[
(x_1 x_4 x_6 - 2x_2 x_3 x_5 - x_1 x_4^2 - x_6 x_3^2 - x_4 x_5^2)/x_1 \geq 0.
\]

Therefore, if \( x_1 > 0 \), then \( X \geq 0 \) if and only if

\[
x_1 x_4 - x_2^2 \geq 0, \quad x_1 x_6 - x_3^2 \geq 0, \quad (x_1 x_4 x_6 - 2x_2 x_3 x_5 - x_1 x_4^2 - x_6 x_3^2 - x_4 x_5^2)/x_1 \geq 0.
\]

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

Examples

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

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

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

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

(b) Show that $u_\alpha$ 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 $u_\alpha$ 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:

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

(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 $u_\alpha$ is strictly concave.

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

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

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

(b) $f(x_1, x_2) = x_1x_2$ on $\mathbb{R}_+^2$.

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

$$\{(x_1, x_2) \in \mathbb{R}_+^2 \mid x_1x_2 \geq \alpha\}$$

are convex. It is not quasiconvex.

(c) $f(x_1, x_2) = 1/(x_1x_2)$ on $\mathbb{R}_+^2$.

Solution. The Hessian of $f$ is

$$\nabla^2 f(x) = \frac{1}{x_1x_2} \begin{bmatrix} 2/(x_1^2) & 1/(x_1x_2) \\ 1/(x_1x_2) & 2/x_2^2 \end{bmatrix} \succeq 0$$

Therefore, $f$ is convex and quasiconvex. It is not quasiconcave or concave.
(d) \( f(x_1, x_2) = x_1/x_2 \) on \( \mathbb{R}^2_{++} \).

**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 \( \mathbb{R} \times \mathbb{R}^+ \).

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

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

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

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

**Solution.** Concave and quasiconcave. The Hessian is

\[
\nabla^2 f(x) = \begin{bmatrix}
\alpha(1-\alpha)x_1^{-\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{-\alpha-1}x_2^{-\alpha} \\
\alpha(1-\alpha)x_1^{-\alpha-1}x_2^{1-\alpha} & (1-\alpha)(-\alpha)x_1^{\alpha}x_2^{-\alpha-1} \\
\end{bmatrix}
\]

\[
= \alpha(1-\alpha)x_1^{-\alpha}x_2^{1-\alpha} \begin{bmatrix}
-1/x_1^2 & 1/x_1 x_2 \\
1/x_1 x_2 & -1/x_2^2 \\
\end{bmatrix}
\]

\[
= -\alpha(1-\alpha)x_1^{\alpha}x_2^{1-\alpha} \begin{bmatrix}
1/x_1 & 1/x_1 \\
-1/x_2 & -1/x_2 \\
\end{bmatrix}^T 
\]

\[ \succeq 0. \]

\( f \) is not convex or quasiconvex.

### 3.17

Suppose \( p < 1 \), \( p \neq 0 \). Show that the function

\[
f(x) = \left( \sum_{i=1}^n x_i^p \right)^{1/p}
\]

with \( \text{dom } f = \mathbb{R}^n_{++} \) is concave. This includes as special cases \( f(x) = (\sum_{i=1}^n x_i^{1/2})^2 \) and the harmonic mean \( f(x) = (\sum_{i=1}^n 1/x_i)^{-1} \). Hint. Adapt the proofs for the log-sum-exp function and the geometric mean in §3.1.5.

**Solution.** The first derivatives of \( f \) are given by

\[
\frac{\partial f(x)}{\partial x_i} = \left( \sum_{i=1}^n x_i^p \right)^{(1-p)/p} x_i^{-p-1} = \left( \frac{f(x)}{x_i} \right)^{1-p}.
\]

The second derivatives are

\[
\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{x_i} \left( \frac{f(x)}{x_i} \right)^{1-p} \left( \frac{f(x)}{x_j} \right)^{1-p} = \frac{1-p}{f(x)} \left( \frac{f(x)^2}{x_i x_j} \right)^{1-p}
\]

for \( i \neq j \), and

\[
\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{1-p}{f(x)} \left( \frac{f(x)^2}{x_i^2} \right)^{1-p} - \frac{1-p}{x_i} \left( \frac{f(x)}{x_i} \right)^{1-p}
\]

for \( i = j \).
We need to show that
\[ gy^T \nabla^2 f(x)y = \frac{1-p}{f(x)} \left( \sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}} \right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} \leq 0 \]

This follows by applying the Cauchy-Schwarz inequality \( a^T b \leq ||a||_2 ||b||_2 \) with
\[ a_i = \left( \frac{f(x)}{x_i} \right)^{-p/2}, \quad b_i = y_i \left( \frac{f(x)}{x_i} \right)^{1-p/2}, \]
and noting that \( \sum_{i} a_i^2 = 1. \)

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

(a) \( f(X) = \text{tr} \left( X^{-1} \right) \) is convex on \( \text{dom} \ f = \mathbb{S}^n_{++} \).

(b) \( f(X) = (\det X)^{1/n} \) is concave on \( \text{dom} \ f = \mathbb{S}^n_{++} \).

**Solution.**

(a) Define \( g(t) = f(Z + tV) \), where \( Z > 0 \) and \( V \in \mathbb{S}^n \).

\[
g(t) = \text{tr}((Z + tV)^{-1}) = \text{tr}(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}) = \text{tr}(Z^{-1}Q(I + t\lambda)^{-1}Q^T) = \text{tr}(Q^TZ^{-1}Q(I + t\lambda)^{-1}) = \sum_{i=1}^{n} (Q^TZ^{-1}Q)_{ii}(1 + t\lambda_i)^{-1},
\]
where we used the eigenvalue decomposition \( Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T \). In the last equality we express \( g \) as a positive weighted sum of convex functions \( 1/(1 + t\lambda_i) \), hence it is convex.

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

\[
g(t) = (\det(Z + tV))^{1/n} = (\det Z^{1/2} \det(I + tZ^{-1/2}VZ^{-1/2}) \det Z^{1/2})^{1/n} = (\det Z)^{1/n} \left( \prod_{i=1}^{n} (1 + t\lambda_i) \right)^{1/n}
\]
where \( \lambda_i, i = 1, \ldots, n \), are the eigenvalues of \( Z^{-1/2}VZ^{-1/2} \). From the last equality we see that \( g \) is a concave function of \( t \) on \( \{ t \mid Z + tV > 0 \} \), since \( \det Z > 0 \) and the geometric mean \( (\prod_{i=1}^{n} x_i)^{1/n} \) is concave on \( \mathbb{R}^n_{++} \).

**3.19** Nonnegative weighted sums and integrals.

(a) Show that \( f(x) = \sum_{i=1}^{r} \alpha_i x_i \) is a convex function of \( x \), where \( \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_r \geq 0 \), and \( x_i \) denotes the \( i \)th largest component of \( x \). (You can use the fact that \( f(x) = \sum_{i=1}^{r} x_i \) is convex on \( \mathbb{R}^n \).

**Solution.** We can express \( f \) as
\[
f(x) = \alpha_r(x_{[1]} + x_{[2]} + \cdots + x_{[r]}) + (\alpha_{r-1} - \alpha_r)(x_{[1]} + x_{[2]} + \cdots + x_{[r-1]})
+ (\alpha_{r-2} - \alpha_{r-1})(x_{[1]} + x_{[2]} + \cdots + x_{[r-2]}) + \cdots + (\alpha_1 - \alpha_2)x_{[1]},
\]
which is a nonnegative sum of the convex functions
\[ x[1], \quad x[1] + x[2], \quad x[1] + x[2] + x[3], \quad \ldots, \quad x[1] + x[2] + \cdots + x[v]. \]

(b) Let \( T(x, \omega) \) denote the trigonometric polynomial
\[ T(x, \omega) = x_1 + x_2 \cos \omega + x_3 \cos 2\omega + \cdots + x_n \cos (n-1)\omega. \]
Show that the function
\[ f(x) = - \int_0^{2\pi} \log T(x, \omega) \, d\omega \]
is convex on \( \{ x \in \mathbb{R}^n \mid T(x, \omega) > 0, \ 0 \leq \omega \leq 2\pi \} \).

**Solution.** The function
\[ g(x, \omega) = - \log(x_1 + x_2 \cos \omega + x_3 \cos 2\omega + \cdots + x_n \cos (n-1)\omega) \]
is convex in \( x \) for fixed \( \omega \). Therefore
\[ f(x) = \int_0^{2\pi} g(x, \omega) \, d\omega \]
is convex in \( x \).

### 3.20 Composition with an affine function
Show that the following functions \( f : \mathbb{R}^n \to \mathbb{R} \) are convex.

(a) \( f(x) = \| Ax - b \| \), where \( A \in \mathbb{R}^{m \times n} \), \( b \in \mathbb{R}^m \), and \( \| \cdot \| \) is a norm on \( \mathbb{R}^n \).

**Solution.** \( f \) is the composition of a norm, which is convex, and an affine function.

(b) \( f(x) = -(\det(A_0 + x_1 A_1 + \cdots + x_n A_n))^{1/m} \), on \( \{ x \mid A_0 + x_1 A_1 + \cdots + x_n A_n > 0 \} \), where \( A_i \in \mathbb{S}^m \).

**Solution.** \( f \) is the composition of the convex function \( h(X) = -(\det X)^{1/m} \) and an affine transformation. To see that \( h \) is convex on \( \mathbb{S}^m_+ \), we restrict \( h \) to a line and prove that \( g(t) = -(\det(Z + tV))^{1/m} \) is convex:
\[
g(t) = -(\det(Z + tV))^{1/m} = -(\det Z)^{1/m} (\det(I + tZ^{-1/2}VZ^{-1/2}))^{1/m} = -(\det Z)^{1/m} (\prod_{i=1}^{m}(1 + t\lambda_i))^{1/m}
\]
where \( \lambda_1, \ldots, \lambda_m \) denote the eigenvalues of \( Z^{-1/2}VZ^{-1/2} \). We have expressed \( g \) as the product of a negative constant and the geometric mean of \( 1 + t\lambda_i, \ i = 1, \ldots, m \). Therefore \( g \) is convex. (See also exercise 3.18.)

(c) \( f(X) = \text{tr}(A_0 + x_1 A_1 + \cdots + x_n A_n)^{-1} \), on \( \{ x \mid A_0 + x_1 A_1 + \cdots + x_n A_n > 0 \} \), where \( A_i \in \mathbb{S}^m \). (Use the fact that \( \text{tr}(X^{-1}) \) is convex on \( \mathbb{S}_+^m \); see exercise 3.18.)

**Solution.** \( f \) is the composition of \( \text{tr}(X^{-1}) \) and an affine transformation
\[ x \mapsto A_0 + x_1 A_1 + \cdots + x_n A_n. \]

### 3.21 Pointwise maximum and supremum
Show that the following functions \( f : \mathbb{R}^n \to \mathbb{R} \) are convex.
2.26 Support function. The support function of a set $C \subseteq \mathbb{R}^n$ is defined as

$$S_C(y) = \sup\{y^T x \mid x \in C\}.$$  

(We allow $S_C(y)$ to take on the value $+\infty$.) Suppose that $C$ and $D$ are closed convex sets in $\mathbb{R}^n$. Show that $C = D$ if and only if their support functions are equal.

**Solution.** Obviously if $C = D$ the support functions are equal. We show that if the support functions are equal, then $C = D$, by showing that $D \subseteq C$ and $C \subseteq D$.

We first show that $D \subseteq C$. Suppose there exists a point $x_0 \in D$, $x \notin C$. Since $C$ is closed, $x_0$ can be strictly separated from $C$, i.e., there exists an $a \neq 0$ with $a^T x_0 > b$ and $a^T x < b$ for all $x \in C$. This means that

$$\sup_{x \in C} a^T x \leq a^T x_0 \leq \sup_{x \in D} a^T x,$$

which implies that $S_C(a) \neq S_D(a)$. By repeating the argument with the roles of $C$ and $D$ reversed, we can show that $C \subseteq D$.

2.27 Converse supporting hyperplane theorem. Suppose the set $C$ is closed, has nonempty interior, and has a supporting hyperplane at every point in its boundary. Show that $C$ is convex.

**Solution.** Let $H$ be the set of all halfspaces that contain $C$. $H$ is a closed convex set, and contains $C$ by definition.

The support function $S_C$ of a set $C$ is defined as $S_C(y) = \sup_{x \in C} y^T x$. The set $H$ and its interior can be defined in terms of the support function as

$$H = \bigcap_{y \neq 0} \{x \mid y^T x \leq S_C(y)\}, \quad \text{int } H = \bigcap_{y \neq 0} \{x \mid y^T x < S_C(y)\},$$

and the boundary of $H$ is the set of all points in $H$ with $y^T x = S_C(y)$ for at least one $y \neq 0$.

By definition $\text{int } C \subseteq \text{int } H$. We also have $\text{bd } C \subseteq \text{bd } H$: if $\bar{x} \in \text{bd } C$, then there exists a supporting hyperplane at $\bar{x}$, i.e., a vector $a \neq 0$ such that $a^T \bar{x} = S_C(a)$, i.e., $\bar{x} \in \text{bd } H$.

We now show that these properties imply that $C$ is convex. Consider an arbitrary line intersecting $\text{int } C$. The intersection is a union of disjoint open intervals $I_k$, with endpoints in $\text{bd } C$ (hence also in $\text{bd } H$), and interior points in $\text{int } C$ (hence also in $\text{int } H$). Now $\text{int } H$ is a convex set, so the interior points of two different intervals $I_1$ and $I_2$ can not be separated by boundary points (since boundary points are in $\text{bd } H$, not in $\text{int } H$). Therefore there can be at most one interval, i.e., $\text{int } C$ is convex.