Convex Optimization - AI2101
Assignment - II

Bhuvan Chandra K
AI23BTECH11013

February 3, 2025

Question 2.11

Show that the hyperbolic set

$$\{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}$$

is convex.

As a generalization, show that

$$\{x \in \mathbb{R}^n_+ \mid \prod_{i=1}^{n} x_i \geq 1\}$$

is convex.

Hint: If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then

$$a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b.$$

Solution:

Given a hyperbolic set $x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1$, We need to prove that this set is convex.
Let this set be S,

$$S = \{x \in \mathbb{R}^2_+ \mid x_1 x_2 \geq 1\}$$

From the definition of convexity of a set,
A set is convex if for any $x, y \in C$ and any θ with $0 \leq \theta \leq 1$, we have:

$$\theta x + (1-\theta) y \in C$$

Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be points in S. From the given condition:

$$x_1 x_2 \geq 1, \quad y_1 y_2 \geq 1$$

We have to show that:

$$[\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2] \geq 1$$

We can use the result: If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$.

Since $x_1, x_2, y_1, y_2 \in \mathbb{R}_+$ and $0 \leq \theta \leq 1$,

$$\theta x_1 + (1-\theta) y_1 \geq (x_1)^\theta (y_1)^{1-\theta}$$

$$\theta x_2 + (1-\theta) y_2 \geq (x_2)^\theta (y_2)^{1-\theta}$$

So we can write:

$$[\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2] \geq [(x_1)^\theta (y_1)^{1-\theta}][(x_2)^\theta (y_2)^{1-\theta}]$$

$$[\theta x_1 + (1-\theta) y_1][\theta x_2 + (1-\theta) y_2] \geq (x_1 x_2)^\theta (y_1 y_2)^{1-\theta}$$

Since $x_1x_2 \geq 1$ and $y_1y_2 \geq 1$,

$$(x_1x_2)^\theta (y_1y_2)^{1-\theta} \geq 1^\theta \cdot 1^{1-\theta}$$
$$(x_1x_2)^\theta (y_1y_2)^{1-\theta} \geq 1$$
$$[\theta x_1 + (1-\theta)y_1][\theta x_2 + (1-\theta)y_2] \geq 1$$

Hence $(\theta x_1 + (1-\theta)y_1, \theta x_2 + (1-\theta)y_2) \in S$.

$$\therefore \text{The set is convex in } \mathbb{R}^2_+.$$

Now let's consider this set $\{x \in \mathbb{R}^n_+ \mid \prod_{i=1}^n x_i \geq 1\}$. We need to prove that this set is convex.

Let this set be M,

$$M = \{x \in \mathbb{R}^n_+ \mid \prod_{i=1}^n x_i \geq 1\}$$

From the definition of convexity of a set,
A set is convex if for any $x, y \in C$ and any θ with $0 \leq \theta \leq 1$, we have:

$$\theta x + (1-\theta)y \in C$$

Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be points in M. From the given condition:

$$\prod_{i=1}^n x_i \geq 1, \quad \prod_{i=1}^n y_i \geq 1$$

We have to show that:

$$\prod_{i=1}^n [\theta x_i + (1-\theta)y_i] \geq 1$$

We can use the result: If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b$.

Since $x_i, y_i \in \mathbb{R}_+ \forall i = 1, \ldots, n$, $0 \leq \theta \leq 1$,

$$\theta x_i + (1-\theta)y_i \geq (x_i)^\theta (y_i)^{1-\theta}.$$

So we can write:

$$\prod_{i=1}^n [\theta x_i + (1-\theta)y_i] \geq \prod_{i=1}^n (x_i)^\theta (y_i)^{1-\theta}.$$
$$\prod_{i=1}^n [\theta x_i + (1-\theta)y_i] \geq \left(\prod_{i=1}^n x_i\right)^\theta \left(\prod_{i=1}^n y_i\right)^{1-\theta}.$$

Since $\prod_{i=1}^n x_i \geq 1$ and $\prod_{i=1}^n y_i \geq 1$,

$$\left(\prod_{i=1}^n x_i\right)^\theta \left(\prod_{i=1}^n y_i\right)^{1-\theta} \geq 1^\theta \cdot 1^{1-\theta}.$$
$$\left(\prod_{i=1}^n x_i\right)^\theta \left(\prod_{i=1}^n y_i\right)^{1-\theta} \geq 1.$$
$$\prod_{i=1}^n [\theta x_i + (1-\theta)y_i] \geq 1.$$

Hence $(\theta x_1 + (1-\theta)y_1, \ldots, \theta x_n + (1-\theta)y_n)$ lies in M.

$$\therefore \text{The set is convex in } \mathbb{R}^n_+.$$

Question 2.16

Show that if S_1 and S_2 are convex sets in $\mathbb{R}^{m+n}$, then so is their partial sum

$$S = \{(x, y_1 + y_2) \mid x \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n, (x, y_1) \in S_1, (x, y_2) \in S_2\}.$$

Solution:

Given: If S_1 and S_2 are convex sets in $\mathbb{R}^{m+n}$, then we have to show that their partial sum

$$S = \{(x, y_1 + y_2) \mid x \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n, (x, y_1) \in S_1, (x, y_2) \in S_2\}$$

is also convex.

From the definition of convexity of a set:
A set S is convex if, for any points P_1 and P_2 from S, where

$$P_1 = (x_1, y_1), \quad P_2 = (x_2, y_2),$$

and for any θ with $0 \leq \theta \leq 1$, we have:

$$\theta P_1 + (1-\theta) P_2 \in S.$$

That is,

$$(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \in S.$$

Since each point in S is the sum of the corresponding points from S_1 and S_2, we defined

$$S = \{(x\ ,\ y_1 + y_2) \mid (x, y_1) \in S_1, (x, y_2) \in S_2\}.$$

Take two points from S, let them be T_1, T_2, and define them as:

$$T_1 = (x_1\ ,\ y_1 + y_2), \quad \text{where } (x_1, y_1) \in S_1, (x_1, y_2) \in S_2.$$

$$T_2 = (x_2\ ,\ y_3 + y_4), \quad \text{where } (x_2, y_3) \in S_1, (x_2, y_4) \in S_2.$$

The convex combination of these two points T_1 and T_2 with parameter θ, where $0 \leq \theta \leq 1$, is:

$$\theta(x_1\ ,\ y_1 + y_2) + (1-\theta)(x_2\ ,\ y_3 + y_4).$$

It can be written as:

$$(\theta x_1 + (1-\theta)x_2\ ,\ \theta(y_1 + y_2) + (1-\theta)(y_3 + y_4))$$

The second component can be written as:

$$\theta(y_1 + y_2) + (1-\theta)(y_3 + y_4) = (\theta y_1 + (1-\theta)y_3) + (\theta y_2 + (1-\theta)y_4)$$

Now, the points (x_1, y_1) and (x_2, y_3) belong to S_1. Since S_1 is convex, so for parameter θ with $0 \leq \theta \leq 1$ we have:

$$(\theta x_1 + (1-\theta)x_2\ ,\ \theta y_1 + (1-\theta)y_3) \in S_1$$

Similarly, the points (x_1, y_2) and (x_2, y_4) belong to S_2. Since S_2 is convex, so for parameter θ with $0 \leq \theta \leq 1$ we have:

$$(\theta x_1 + (1-\theta)x_2\ ,\ \theta y_2 + (1-\theta)y_4) \in S_2$$

We have:

$$(\theta x_1 + (1-\theta)x_2 \ , \ \theta y_1 + (1-\theta)y_3) \in S_1$$

and

$$(\theta x_1 + (1-\theta)x_2 \ , \ \theta y_2 + (1-\theta)y_4) \in S_2$$

Their partial sum will belong to S (since the first components are the same according to the question):

$$(\theta x_1 + (1-\theta)x_2 \ , \ \theta y_1 + (1-\theta)y_3 + \theta y_2 + (1-\theta)y_4) \in S$$

We have shown that the convex combination of any two points in S remains in S.

$$\therefore \text{Given Partial Sum, S is convex.}$$

Question 2.17

In this problem, we study the image of hyperplanes, halfspaces, and polyhedra under the perspective function

$$P(x,t) = \frac{x}{t},$$

with dom $P = \mathbb{R}^n \times \mathbb{R}_{++}$. For each of the following sets C, give a simple description of

$$P(C) = \left\{ \frac{v}{t} \mid (v,t) \in C, \, t > 0 \right\}.$$

(a) The polyhedron $C = \text{conv}\{(v_1, t_1), \ldots, (v_K, t_K)\}$ where $v_i \in \mathbb{R}^n$ and $t_i > 0$.

(b) The hyperplane $C = \{(v,t) \mid f^T v + gt = h\}$ (with f and g not both zero).

(c) The halfspace $C = \{(v,t) \mid f^T v + gt \leq h\}$ (with f and g not both zero).

(d) The polyhedron $C = \{(v,t) \mid Fv + gt \preceq h\}$.

Solution:

(a) The polyhedron $C = \text{conv}\{(V_i, t_i), \ldots, (V_k, t_k)\}$ where $V_i \in \mathbb{R}^n$ and $t_i \geq 0$.
Actually, $P(C)$ could be:

$$P(C) = \text{conv}\left\{ \frac{V_i}{t_i}, \ldots, \frac{V_k}{t_k} \right\}$$

The assumption here is that since C is a convex polyhedron (i.e., a convex set) and we know that applying the convex perspective function $P(V,t) = \frac{V}{t}$ to a convex set results in a convex set, the image $P(C)$ must also be convex.
Let us show

$$P(C) \subseteq \text{conv}\left\{ \frac{V_1}{t_1}, \ldots, \frac{V_k}{t_k} \right\}$$

Let $x = (V,t) \in C$. Since C is the convex hull of points $(V_1, t_1), \ldots, (V_k, t_k)$, any point $(V,t) \in C$ can be written as a convex combination of these points:

$$(V,t) = \sum_{i=1}^{k} \alpha_i (V_i, t_i), \quad \alpha_i \geq 0, \quad \sum_{i=1}^{k} \alpha_i = 1$$

This means that the vector V and the scalar t can be expressed as:

$$V = \sum_{i=1}^{k} \alpha_i V_i, \quad t = \sum_{i=1}^{k} \alpha_i t_i$$

Now, applying the perspective function $P(V,t) = \frac{V}{t}$ to the point (V,t):

$$P(V,t) = \frac{V}{t} = \frac{\sum_{i=1}^{k} \alpha_i V_i}{\sum_{i=1}^{k} \alpha_i t_i}$$

Rewriting $P(V,t)$ as a convex combination of $\frac{V_i}{t_i}$:

$$P(V,t) = \sum_{i=1}^{k} \frac{\alpha_i t_i}{\sum_{j=1}^{k} \alpha_j t_j} \cdot \frac{v_i}{t_i}$$

Let us define β_i as:

$$\beta_i = \frac{\alpha_i t_i}{\sum_{j=1}^{k} \alpha_j t_j}, \quad \text{for } i = 1, \ldots, k$$

Since $\alpha_i \geq 0$ and $t_i > 0$, it follows that $\beta_i \geq 0$ for all i.
Also, the sum of β_i's = 1:

$$\sum_{i=1}^{k} \beta_i = \sum_{i=1}^{k} \frac{\alpha_i t_i}{\sum_{j=1}^{k} \alpha_j t_j} = \frac{\sum_{i=1}^{k} \alpha_i t_i}{\sum_{j=1}^{k} \alpha_j t_j} = 1$$

We can see that $P(V,t)$ is a convex combination of the points $\frac{V_i}{t_i}$ with valid coefficients.

$$\therefore P(V,t) \in \operatorname{conv}\left\{\frac{V_1}{t_1}, \frac{V_2}{t_2}, \ldots, \frac{V_k}{t_k}\right\}$$

Now, Let us show that

$$P(C) \supseteq \operatorname{conv}\left\{\frac{v_1}{t_1}, \frac{v_2}{t_2}, \ldots, \frac{v_k}{t_k}\right\}$$

Let's Consider a point

$$z = \sum_{i=1}^{k} \beta_i \frac{v_i}{t_i}$$

where $\beta_i \geq 0$ and

$$\sum_{i=1}^{k} \beta_i = 1$$

We would like to show that this point lies in the image $P(C)$.
Define γ_i as

$$\gamma_i = \frac{\beta_i t_i}{\sum_{j=1}^{k} \beta_j t_j}, \quad i = 1, \ldots, k$$

Since $\beta_i \geq 0$, the coefficients γ_i are non-negative and sum up to 1 (denominator is the sum of numerators).
Define v' and t' like:

$$v' = \sum_{i=1}^{k} \gamma_i v_i = \sum_{i=1}^{k} \frac{\beta_i t_i}{\sum_{j=1}^{k} \beta_j t_j} v_i$$

$$t' = \sum_{i=1}^{k} \gamma_i t_i = \frac{\sum_{i=1}^{k} \beta_i}{\sum_{j=1}^{k} \beta_j t_j} = \frac{1}{\sum_{j=1}^{k} \beta_j t_j} = 1$$

(v', t') lies in the convex hull of points (v_i, t_i), meaning $(v', t') \in C$.
We can see that

$$\frac{v'}{t'} = \sum_{i=1}^{k} \frac{\beta_i v_i}{t_i} = z$$

We have

$$P(v', t') = \frac{v'}{t'} = z$$

Thus, $z \in P(C)$, so

$$P(C) \supseteq \text{conv}\left\{\frac{v_1}{t_1}, \ldots, \frac{v_k}{t_k}\right\}$$

Since $P(C) \subseteq \text{conv}\left\{\frac{v_1}{t_1}, \ldots, \frac{v_k}{t_k}\right\}$ and

$$P(C) \supseteq \text{conv}\left\{\frac{v_1}{t_1}, \ldots, \frac{v_k}{t_k}\right\}$$

We can conclude that

$$\therefore P(C) = \text{conv}\left\{\frac{v_1}{t_1}, \ldots, \frac{v_k}{t_k}\right\}$$

(b) The hyperplane $C = \{(v, t) \mid f^T v + gt = h\}$ (with f and g not both zero).

Let

$$x = P(v, t) = \frac{v}{t}$$

Then

$$v = xt$$

Substituting:

$$f^T(xt) + gt = h$$
$$t(f^T x + g) = h$$

Since $t > 0$,

$$f^T x + g = \frac{h}{t}$$

The domain of the perspective function is $\mathbb{R}^n \times \mathbb{R}_{++}$, which means the domain consists of points (v, t) where $v \in \mathbb{R}^n$ and $t > 0$. It depends on whether $\frac{h}{t}$ is defined for some $t > 0$.

$$P(C) = \left\{x \mid f^T x + g = \frac{h}{t} \text{ for some } t > 0\right\}$$

- If $h = 0$, then

  $$f^T x + g = 0$$

  i.e., $P(C)$ is a hyperplane in $\mathbb{R}^n$ within the given domain.

- If $h > 0$, then

  $$f^T x + g > 0$$

  i.e., $P(C)$ is the region where $f^T x + g > 0$, meaning $P(C)$ is a half-space within the given domain.

- If $h < 0$, then
$$f^T x + g < 0$$
i.e., $P(C)$ is the region where $f^T x + g < 0$, meaning $P(C)$ is the opposite half-space within the given domain.

(c) The halfspace
$$C = \{(v, t) \mid f^T v + gt \leq h\}$$
(with f and g both not zero).
Let
$$x = P(v, t) = \frac{v}{t}$$
Then
$$v = xt$$
Substituting:
$$f^T(xt) + gt \leq h$$
$$t(f^T x + g) \leq h$$
Since $t > 0$,
$$f^T x + g \leq \frac{h}{t}$$
This defines a halfspace in $\mathbb{R}^n$ and the domain is restricted to the portion where $t > 0$ ($t \in \mathbb{R}_{++}$).
$$P(C) = \left\{ x \mid f^T x + g \leq \frac{h}{t} \text{ for some } t > 0 \right\}$$

- If $h = 0$, the projection results in the halfspace:
$$f^T x + g \leq 0 \quad \text{(halfspace)}$$
- If $h > 0$, h can be arbitrarily large, and so does $\frac{h}{t}$, meaning there is no bound on $f^T x + g$. Hence, the projection covers all of $\mathbb{R}^n$.
- If $h \leq 0$, if h is negative, then so is $\frac{h}{t}$. So the projection will be:
$$f^T x + g < 0 \quad \text{(strict halfspace)}$$

(d) The polyhedron
$$C = \{(v, t) \mid Fv + gt \preceq h\}$$
Let F_i be the i-th row of F, g_i be the i-th element of g, h_i be the i-th element of h.

The symbol $\preceq$ means each row of the inequality holds separately.
Let
$$x = P(v, t) = \frac{v}{t}$$
Then
$$v = xt$$
Substituting into each row i of the polyhedral inequality:
$$F_i v + g_i t \leq h_i$$
$$t(F_i x + g_i) \leq h_i$$
Since $t > 0$ (from the domain of the perspective function),
$$F_i x + g_i \leq \frac{h_i}{t}$$

- If $h_i = 0$, we get:
$$F_i x + g_i \leq 0$$
This is a standard affine inequality in x, defining a halfspace.

- If $h_i < 0$, we get:
$$F_i x + g_i < 0$$
This results in a strict inequality, meaning x must be in the interior of a halfspace.

- If some inequalities have $h_i > 0$ and some inequalities have $h_j < 0$,
$$F_i x + g_i \leq \frac{h_i}{t}, \quad F_j x + g_j \leq \frac{h_j}{t}$$
$$\frac{F_i x + g_i}{h_i} \leq \frac{1}{t}, \frac{F_j x + g_j}{h_j} \leq \frac{1}{t}$$
Since both are bounded by $\frac{1}{t}$,
$$\frac{F_j x + g_j}{h_j} \leq \frac{F_i x + g_i}{h_i}, \quad \text{if } h_i > 0, h_j < 0$$
These are ordered inequalities between polyhedral constraints.
$P(L)$ can be depicted as
$$\bigcap_i \{x \mid F_i x + g_i < 0\} \cup \{x \mid F_i x + g_i = 0 \text{ and } h_i \geq 0\}$$
This is an intersection of modified half-spaces.

Question 2.18

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the linear-fractional function
$$f(x) = \frac{Ax + b}{c^T x + d}, \quad \text{dom } f = \{x \mid c^T x + d > 0\}.$$
Suppose the matrix
$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix}$$
is nonsingular. Show that f is invertible and that f^{-1} is a linear-fractional mapping. Give an explicit expression for f^{-1} and its domain in terms of A, b, c, and d.

Hint: It may be easier to express f^{-1} in terms of Q.

Solution:

Given, $f : \mathbb{R}^n \to \mathbb{R}^n$ be the linear-fractional function

$$f(x) = \frac{Ax + b}{c^T x + d},$$

$$\text{dom } f = \{x \mid c^T x + d > 0\}$$

The matrix

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix}$$

is nonsingular.

$f(x)$ is well defined since its denominator $c^T x + d$ is strictly positive.

Interpreting $f(x)$ geometrically using projective transformations.

We use $P(x) = \begin{bmatrix} x \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1}$ instead of working in $\mathbb{R}^n$. It is a ray in $\mathbb{R}^{n+1}$.

Define the augmented matrix Q:

$$Q = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

Matrix Q acts on this ray $P(x)$ to produce another ray $QP(x)$

$$QP(x) = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} Ax + b \\ c^T x + d \end{bmatrix}$$

The function P^{-1} extracts the first n components and normalizes by the last coordinate (in $\mathbb{R}^{n+1}$):

$$P^{-1} \begin{bmatrix} u \\ v \end{bmatrix} = \frac{u}{v}, \quad \text{for } v > 0$$

Since $x \in \text{dom } f$, i.e., $c^T x + d > 0$, we can normalize by

$$P^{-1}(QP(x)) = P^{-1} \begin{bmatrix} Ax + b \\ c^T x + d \end{bmatrix}$$

$$P^{-1}(QP(x)) = \frac{Ax + b}{c^T x + d}$$

The function $f(x)$ can be obtained by projective transformation.

Since Q is non-singular (invertible), $\det(Q) \neq 0$

The transformation Q is bijective in projective space, so it has a well-defined inverse transformation.

$\therefore$ f(x) must also be invertible.

To find f^{-1}, we use the projective approach again:

$$f^{-1}(y) = P^{-1}(Q^{-1}P(y))$$

As Q is invertible, it can be inverted blockwise.

$$Q^{-1} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}$$

$$A = A, \quad B = b, \quad C = c^T, \quad D = d$$

$$Q^{-1} = \begin{bmatrix} A^{-1} + A^{-1}b(d - c^T A^{-1}b)^{-1}c^T A^{-1} & -A^{-1}b(d - c^T A^{-1}b)^{-1} \\ -(d - c^T A^{-1}b)^{-1}c^T A^{-1} & (d - c^T A^{-1}b)^{-1} \end{bmatrix}$$

Let

$$Q^{-1} = \begin{bmatrix} A' & b' \\ c' & d' \end{bmatrix}$$

for simplicity.

For any y, projective representation:

$$P(y) = \begin{bmatrix} y \\ 1 \end{bmatrix}$$

Multiplying by Q^{-1},

$$Q^{-1}P(y) = \begin{bmatrix} A' & b' \\ c' & d' \end{bmatrix} \begin{bmatrix} y \\ 1 \end{bmatrix} = \begin{bmatrix} A'y + b' \\ c'y + d' \end{bmatrix}$$

We normalize by doing $P^{-1}(Q^{-1}P(y))$.

$$P^{-1}(Q^{-1}P(y)) = \frac{A'y + b'}{c'y + d'}$$

$$f^{-1}(y) = \frac{A'y + b'}{c'y + d'}$$

Thus, f^{-1} is also a linear-fractional function.

The inverse function is only valid where the denominator is positive:

$$\text{i.e., } \operatorname{dom} f^{-1} = \{y \mid c'y + d' > 0\}$$

This ensures $f^{-1}(y)$ is well defined.

$$\therefore f \text{ is invertible and } f^{-1} \text{ is a linear-fractional mapping, and}$$

$$f^{-1}(y) = \frac{A'y + b'}{c'y + d'}$$

where

$$A' = A^{-1} + A^{-1}b(d - c^T A^{-1}b)^{-1}c^T A^{-1}$$

$$b' = -A^{-1}b(d - c^T A^{-1}b)^{-1}$$

$$c' = -(d - c^T A^{-1}b)^{-1}c^T A^{-1}$$

$$d' = (d - c^T A^{-1}b)^{-1}$$

$$f^{-1}(y) = \frac{(A^{-1} + A^{-1}b(d - c^T A^{-1}b)^{-1}c^T A^{-1})y + (-A^{-1}b(d - c^T A^{-1}b)^{-1})}{(-(d - c^T A^{-1}b)^{-1}c^T A^{-1})y + ((d - c^T A^{-1}b)^{-1})}$$

Question 2.19

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be the linear-fractional function

$$f(x) = \frac{Ax+b}{c^T x + d}, \quad \text{dom } f = \{x \mid c^T x + d > 0\}.$$

In this problem, we study the inverse image of a convex set C under f, i.e.,

$$f^{-1}(C) = \{x \in \text{dom } f \mid f(x) \in C\}.$$

For each of the following sets $C \subseteq \mathbb{R}^n$, give a simple description of $f^{-1}(C)$:

(a) The halfspace $C = \{y \mid g^T y \leq h\}$ (with $g \neq 0$).

(b) The polyhedron $C = \{y \mid Gy \preceq h\}$.

(c) The ellipsoid $C = \{y \mid y^T P^{-1} y \leq 1\}$ (where $P \in \mathbb{S}^n_{++}$).

(d) The solution set of a linear matrix inequality, $C = \{y \mid y_1 A_1 + \cdots + y_n A_n \preceq B\}$, where $A_1, \ldots, A_n, B \in \mathbb{S}^p$.

Solution:

Given the linear fractional function

$$f(x) = \frac{Ax+b}{c^T x + d}$$

with domain

$$\text{dom } f = \{x \mid c^T x + d > 0\}.$$

For each convex set C, we want to identify the pre-image:

$$f^{-1}(C) = \{x \in \text{dom } f \mid f(x) \in C\}.$$

(a) The halfspace $C = \{y \mid g^T y \leq h\}$ (with $g \neq 0$). Putting $y = f(x)$,

$$f^{-1}(C) = \{x \in \text{dom } f \mid g^T f(x) \leq h\}.$$

$$g^T \left(\frac{Ax+b}{c^T x + d} \right) \leq h, \quad c^T x + d > 0.$$

Since $c^T x + d > 0$, multiplying it on both sides of the inequality:

$$g^T(Ax+b) \leq h(c^T x + d), \quad c^T x + d > 0.$$

$$g^T Ax + g^T b \leq hc^T x + hd, \quad c^T x + d > 0.$$

$$(g^T A - hc^T)x \leq hd - g^T b, \quad c^T x + d > 0.$$

Since this is a linear inequality, meaning $f^{-1}(C)$ is a halfspace:

$$f^{-1}(C) = \{x \mid (g^T A - hc^T)x \leq hd - g^T b, \, c^T x + d > 0\}.$$

This is a halfspace intersected with the domain $c^T x + d > 0$ (dom f).

(b) The polyhedron $C = \{y \mid Gy \preceq h\}$. Putting $y = f(x)$,

$$f^{-1}(C) = \{x \in \text{dom } f \mid Gf(x) \preceq h\}.$$

$$G\left(\frac{Ax+b}{c^Tx+d}\right) \preceq h, \quad c^Tx+d>0.$$

Multiplying both sides by c^Tx+d of the inequality:

$$G(Ax+b) \preceq h(c^Tx+d), \quad c^Tx+d>0.$$

$$GAx+Gb \preceq hc^Tx+hd, \quad c^Tx+d>0.$$

$$(GA-hc^T)x \preceq hd-Gb, \quad c^Tx+d>0.$$

This is a polyhedron on x:

$$f^{-1}(C) = \{x \mid (GA-hc^T)x \preceq hd-Gb,\, c^Tx+d>0\}.$$

This is a polyhedral set intersected with the domain.

(c) The ellipsoid $C = \{y \mid y^TP^{-1}y \leq 1\}$ (where $P \in S^n_{++}$).

Ellipsoid is a quadratic constraint on y. We transform it into a quadratic constraint on x.
Putting $y = f(x)$,

$$f^{-1}(C) = \{x \in \mathrm{dom}\, f \mid f(x)^TP^{-1}f(x) \leq 1\}.$$

$$\left(\frac{Ax+b}{c^Tx+d}\right)^T P \left(\frac{Ax+b}{c^Tx+d}\right) \leq 1.$$

Multiplying by $(c^Tx+d)^2$ on both sidess :

$$(Ax+b)^TP^{-1}(Ax+b) \leq (c^Tx+d)^2$$

$$(x^TA^T+b^T)P(Ax+b) \leq (c^Tx+d)^2$$

$$x^TA^TP^{-1}Ax + x^TA^TP^{-1}b + b^TP^{-1}Ax + b^TP^{-1}b \leq (c^Tx)^2 + 2d(c^Tx) + d^2$$

Notice that the second and third terms in the above expression are the same, except for the transpose. Since these are both scalars and the transpose of a scalar is just itself.
So, It can be written as :

$$x^TA^TP^{-1}Ax + 2b^TP^{-1}Ax + b^TP^{-1}b \leq (c^Tx)^2 + 2d(c^Tx) + d^2$$

Since $(c^Tx)^T = x^Tc$,

$$x^TA^TP^{-1}Ax + 2b^TP^{-1}Ax + b^TP^{-1}b \leq x^Tcc^Tx + 2dc^Tx + d^2$$

$$x^T(A^TP^{-1}A - cc^T)x + 2(b^TP^{-1}A - dc^T)x + (b^TP^{-1}b - d^2) \leq 0$$

Let,

$$J = A^TP^{-1}A - cc^T$$

$$k = b^TP^{-1}A - dc$$

$$h = d^2 - b^TP^{-1}b$$

Inequality can be written as :

$$x^TJx + 2q^Tk \leq h, \quad c^Tx+d \geq 0$$

In this expression, $J \succ 0$ (it is positive definite), thus it is an ellipsoid.

This is an ellipsoid intersected with dom f ($c^Tx+d>0$).

(d) The solution set of a linear matrix inequality,

$$C = \{y \mid y_1 A_1 + \cdots + y_n A_n \preceq B\}$$

where $A_1, A_2, \ldots, A_n, B \in \mathbb{S}^p$.
Putting $y = f(x)$,

$$\left(\frac{(Ax+b)_1}{c^T x + d}\right) A_1 + \left(\frac{(Ax+b)_2}{c^T x + d}\right) A_2 + \cdots + \left(\frac{(Ax+b)_n}{c^T x + d}\right) A_n \preceq B.$$

Multiplying both sides with C^T and C,

$$(Ax+b)_1 A_1 + \cdots + (Ax+b)_n A_n \preceq (c^T x + d)B$$

$$\left[(Ax+b)_i = (a_i^T x + b_i)\right]$$

$$\sum_{i=1}^{n} (a_i^T x + b_i) A_i \preceq (c^T x + d)B$$

$$\sum_{i=1}^{n} a_i^T x A_i + \sum_{i=1}^{n} b_i A_i \preceq c^T x B + dB$$

$$\left(\sum_{i=1}^{n} a_i^T A_i - c^T B\right) x \preceq dB - \sum_{i=1}^{n} b_i A_i$$

$$(a_1^T A_1 + a_1^T A_2 + \cdots + a_n^T A_n - c^T B)x \preceq dB - \sum_{i=1}^{n} b_i A_i$$

$(a_1^T A_1 + \cdots + a_n^T A_n - c^T B)x$ can be written as
Product of first columns of A or first rows of A^T multiplied with x_i .

$$\sum (a_{1i} A_1 + \cdots + a_{ni} A_n)\, x_i - \left(\sum c_i B\right) x_i$$

Let

$$P_i = a_{1i} A_1 + \cdots + a_{ni} A_n - c_i B$$

$$Q = dB - (b_1 A_1 + \cdots + b_n A_n)$$

The expression becomes

$$P_1 x_1 + \cdots + P_n x_n \preceq Q, \quad c^T x d > 0$$

This depicts the inverse image of the LMI constraint under the function $f(x)$. $f^{-1}(c)$ is the intersection of dom f $(c^T x + d > 0)$ with the solution set of an LMI.