Problem 1

Determine whether the following sets are convex.

a) $\left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} 1-x & y & 0 & 0 \\ y & 1+x & 0 & 0 \\ 0 & 0 & 1-z & 0 \\ 0 & 0 & 0 & 1+z \end{bmatrix} \succeq 0 \right\}$.

We can observe that this matrix has entries in which none of the variables depend on each other. Therefore, we can write this as:

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{I} + \begin{bmatrix} -x & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & y & 0 & 0 \\ y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -z & 0 \\ 0 & 0 & 0 & z \end{bmatrix} \succeq 0 .$$

Its clear this is a composition of linear inequalities and therefore it is convex.

b) $\{ x \in \mathbb{R}^n \mid x^T (A^{-1} - aa^T) x - 2ba^T x - b^2 \leq 0,\; a^T x + b > 0 \}$

where $A \succ 0$, $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$.

$$\Rightarrow x^T A^{-1} x - x a a^T x - 2ba^T x - b^2 \leq 0$$

$$= x^T A^{-1} x \leq \underbrace{x a a^T x}_{(a^T x)^2} + 2ba^T x + b^2$$

$$(a^T x)^2 \rightarrow (a^T x + b)^2$$

Since $A > 0$, $A^{1/2} > 0$, this is similar to the hyperbolic cone through an affine transformation.

$\hookrightarrow$ convex

hyperbolic cone:

$\{ x \mid x^T P x \leq (c^T x)^2,\; c^T x \geq 0 \}$

inverse image of 2nd-order cone:

$\{ (z,t) \mid z^T z \leq t^2,\; t \geq 0 \}$,

under affine function:

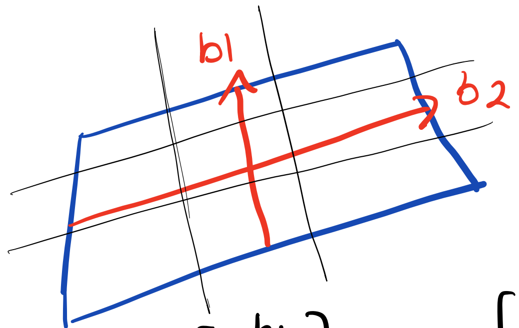
$$\underbrace{f(x) = (P^{1/2} x,\; c^T x)}_{\text{convex}} .$$

Problem 2

Are the sets below polyhedra? If so, give a description in the form of $\{x \mid Ax \preceq b\}$, and if not, explain why.

a) $\{x_1 b_1 + x_2 b_2 \mid 0 \le x_1 \le 1 \text{ and } -1 \le x_2 \le 0\}$, where $b_1, b_2 \in \mathbb{R}^n$, $x_1, x_2 \in \mathbb{R}$.

Polyhedra are the solution set to finitely many linear inequalities and equalites. Intersection of halfspaces. A slab is of the form $\{x \in \mathbb{R}^n \mid \alpha \le a^T x \le \beta\}$, is the intersection of two halfspaces and a polyhedron.



$$\Rightarrow S_1 = \{x_1 b_1 + x_2 b_2 \mid 0 \le x_1 \le 1\}$$

$$S_2 = \{x_1 b_1 + x_2 b_2 \mid -1 \le x_2 \le 0\}$$

S_3: b_1 and b_2 form a plane: $P = x_1 \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} + x_2 \begin{bmatrix} b_2 \\ \vdots \\ b_n \end{bmatrix}$

S_1 & S_2 are slabs that are both orthogonal to S_3

Define a vector c_1 in the plane defined by b_1 & b_2 and orthogonal to b_2 & b_1 respectively:

$$a_1 = b_1 - \frac{b_1^T b_2}{\|b_2\|_2^2} b_2$$

$$a_2 = b_2 - \frac{b_2^T b_1}{\|b_1\|_2^2} b_1$$

We can describe this by the linear inequalities

$$a_1^T x \leq |a_1^T b_1|$$
$$-a_1^T x \leq |a_1^T b_1|$$
$$a_2^T x \leq |a_2^T b_2|$$
$$-a_2^T x \leq |a_2^T b_2|$$

The intersection of finite number of halfspaces is a polyhedron.

linear inequalities

b) $\{x \in \mathbb{R}^n \mid x \succeq 0,\ x^Ty \leq 1$ for all y with $\sum_{i=1}^{n} |y_i| = 1\}$.

$\{x \in \mathbb{R}^n \mid x \succeq 0\}$ is the nonnegative orthant $\mathbb{R}^n_+$.

Then we have $\{x \in \mathbb{R}^n \mid x^Ty \leq 1\ \forall y$ with $\sum_{i=1}^{n} |y_i| = 1\}$.

If $\sum_{i=1}^{n} |y_i| = 1$, then,

$$x^Ty = \sum_{i=1}^{n} x_i y_i \leq \sum_{i=1}^{n} |x_i||y_i| \leq \sum_{i=} |y_i| = 1$$

only holds if $|x_i| \leq 1$.

Therefore, $x^Ty \leq 1\ \forall y$ with $\sum_i |y_i| = 1$ iff $|x_i| \leq 1, i = 1, \ldots, n$.

Hence, our set is the intersection of $\{x \mid |x_i| \leq 1, i = 1, \ldots, n\}$ and the nonnegative orthant which is the intersection of finite many halfspaces and is a polyhedron.

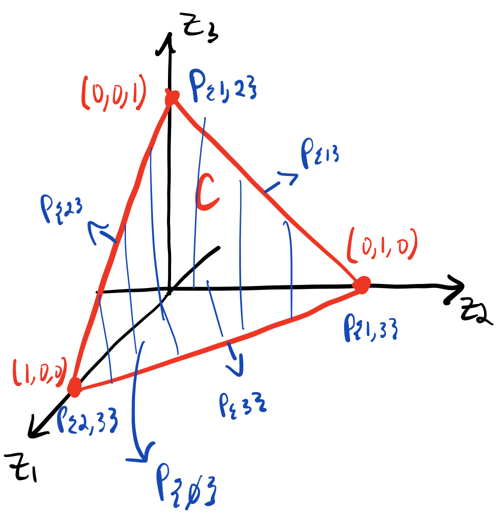
Problem 3

The unit simplex in $\mathbb{R}^n$ is defined as

$$\{x \mid x \succeq 0, \sum_{i=1}^{n} x_i \leq 1\}$$

Describe all the supporting hyperplane of the unit simplex in $\mathbb{R}^3$.

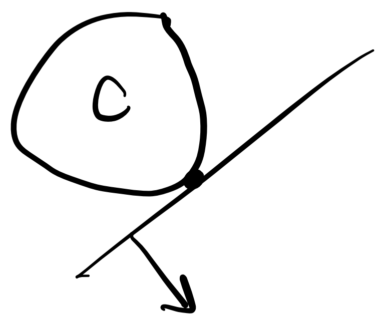
unit simplex in $\mathbb{R}^3$:



Supporting hyperplane for $\mathbb{R}^3$

$$H = \{z \in \mathbb{R}^3 \mid c^T z = cx\}$$

$$\text{s.t. } c^T z \leq c^T x \quad \forall z \in C$$



$$\rightarrow \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad x_1 + x_2 + x_3 = 1$$

$P_\emptyset = P$ itself is a face

- edges intersection of faces

- faces, edges, <u>vertices</u>

faces: $P_{\{1\}} = \{(z_1, z_2, z_3) \in \mathbb{R}^3 \mid z_1 = 0\}$

$P_{\{2\}} = \{(z_1, z_2, z_3) \in \mathbb{R}^3 \mid z_2 = 0\}$ $\longrightarrow$ Normal vector of supporting hyperplanes

$P_{\{3\}} = \{(z_1, z_2, z_3) \in \mathbb{R}^3 \mid z_3 = 0\}$

at these faces: $(-1, 0, 0)$

$(0, -1, 0)$

$(0, 0, -1)$

Vertices: $(1, 0, 0)$

$(0, 1, 0)$

$(0, 0, 1)$

There are an infinite number of hyperplanes at each vertex so we define normal vectors as a rotation vector for each vertex above.

Define a hyperplane $\{x \mid a^T x + b = 0\}$

$$\{x \mid (-\alpha, -\beta, -\gamma) \cdot x \mid \alpha, \beta, \gamma \in [0,1], \sqrt{\alpha^2 + \beta^2 + \gamma^2} = 1\}$$

Then define a hyperplane for each vertex

$$\{x \mid (-\alpha, -\beta, -\gamma) \cdot (x - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix})\}$$

$$\{x \mid (-\alpha, -\beta, -\gamma) \cdot (x - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix})\}$$

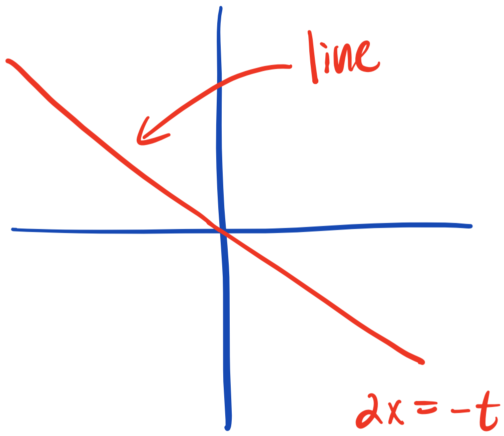
$$\{x \mid (-\alpha, -\beta, -\gamma) \cdot (x - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix})\}$$

Problem 4

Find the dual of the following cones.

The dual cone is described by $K_m^* = \{y \mid x^T y \geq 0 \ \forall x \in K_m\}$.

a) $K = \{(x,t) \mid x \in \mathbb{R}, t \in \mathbb{R}, 2x = -t\}$.



The dual of a subspace $V \subseteq \mathbb{R}^n$ by, ex. 2.22 in the book, is the orthogonal complement $V^T = \{y \mid v^T y = 0 \ \forall v \in V\}$.

ref. Ex. 2.22

b) $K = \{(x,t) \mid x \in \mathbb{R}^n, t \in \mathbb{R}_+, \|x\|_\infty \leq t\}$

Dual norm: The dual of the l_∞-norm is the l_1-norm:

$$\sup\{z^T x \mid \|x\|_\infty \leq 1\} = \sum_{i=1}^n |z_i| = \|z\|_1$$

* def of dual norm

The dual of the norm cone is the cone defined by the dual norm,

$$K^* = \{(u,v) \in \mathbb{R}^{n+1} \mid \|u\|_1 \leq v\}.$$

We want to show that

$$x^T u + tv \geq \text{ when } \|x\| \leq t \iff \|u\|_1 \leq v.$$

Suppose $\|u\|_* \leq v$ and $\|x\| \leq t$ for $t > 0$. Then applying the def. of the dual norm and that $\|-x/t\| \leq 1$, we get,

$$u^T(-x/t) \leq \|u\|_* \leq v,$$

therefore $u^T x + vt > 0$.

Now suppose $\|u\|_* > v$, saying the RHS of iff doesn't hold. Then by def. of dual norm, $\exists\, x$ with $\|x\| \leq$ and $x^T u > v$. Take $t = 1$, we have

$$u^T(-x) + v < 0$$

hence a contradiction. # Reference: Ex. 2.25 in Book

c) $K = \{ X \in S^n \mid X = AA^T,\ A \in \mathbb{R}_+^{n \times m} \}$.

Note that a symmetric matrix $X \in S^n$ is PSD iff it is the sum of outerproducts, i.e., $X = AA^T = \sum_{i=1}^n a_i a_i^T$.

Then, $\{ Y \mid \mathrm{Tr}(XY) \quad \forall\, X \in K_m \}$

$$\Rightarrow \quad Tr(XY)$$

$$= Tr(AA^TY)$$

$$= Tr\left(Y \sum_{i=1}^{n} a_i a_i^T\right)$$

$$= \sum_{i=1}^{n} Tr(Y a_i a_i^T)$$

$$= \sum_{i=1}^{n} Y a_i a_i^T$$

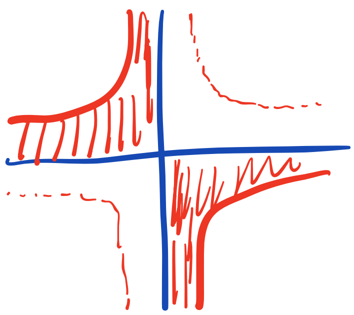
$$= a_i Y a_i^T \geq 0$$

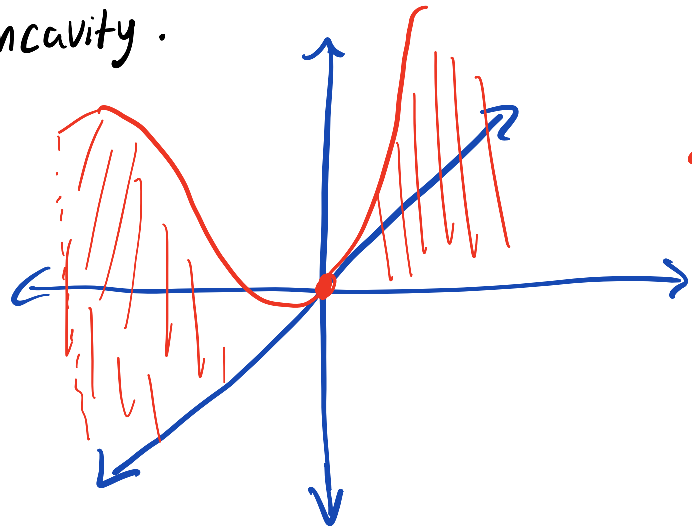
Therefore, $K^* = \{ Y \mid a_i Y a_i^T , \ a \succeq 0 \}$

This means our dual cone is a copositive matrix.

Problem 5

The function $f(x) = x_1 x_2$ where $x = (x_1, x_2) \in \mathbb{R}_+^2$, is neither convex nor concave. Here we will give counter-examples for convexity and concavity.





a) We pick two points $x \in \mathbb{R}^2$ and $y \in \mathbb{R}^2$ as $(1,0)$ and $(0,1)$ and show the following doesn't hold:

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y) \quad \text{for } 0 \leq \theta \leq 1$$

For $x = (1,0)$, $y = (0,1)$ and $\theta = 0.5$

$$\Rightarrow \quad f(\theta x + (1-\theta) y) = f(0.5(1,0) + 0.5(0,1)) = f(0.5, 0.5) = \frac{1}{4}.$$

and

$$\theta f(x) + (1-\theta) f(y) = 0.5 \cdot f(1,0) + 0.5 \cdot f(0,1) = 0.5(0) + 0.5(0)$$
$$= 0$$

Clearly $\frac{1}{4} \nleq 0$.

b)

Now for $x = (0,0)$, $y = (1,1)$, $\theta = 0.5$

$$\Rightarrow f(\theta x + (1-\theta) y) = f(0.5(0,0) + 0.5(1,1)) = \frac{1}{4}$$

and

$$\theta f(x) + (1-\theta) f(y) = 0.5 \cdot f(0,0) + 0.5 f(1,1) = 0.5(0) + 0.5(1) = \frac{1}{2}$$

Clearly $\frac{1}{4} \ngeq \frac{1}{2}$.

Those are counterexamples for both convexity and concavity.

* these do not hold for all $0 \leq \theta \leq 1$.

Problem 6

Consider the function

$$f(X) = -\log\det X - Tr(AX^{-1}),$$

where $dom(f) = \{X \in S^n \mid 0 \prec X \preceq 2A\}$ and $A \succeq 0$ is a given matrix. Show that f is concave in X over this domain.

Restriction to a line: Recall that, $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff $g: \mathbb{R} \to \mathbb{R}$ where $g(t) = f(x+tv)$, $\mathrm{dom}\, g = \{t \mid z + tv \in \mathrm{dom} f\}$ is convex in t for any $z \in \mathrm{dom} f$, $v \in \mathbb{R}^n$.

$$f(x) = -\log\det X - Tr(AX^{-1}), \quad A \succeq 0, \quad 0 \prec X \preceq 2A$$

Let $V \in S^n_{++}$

$$\begin{aligned}
\text{Let } g(t) &= f(x+tV) \\
&= \log(\det(x+tV)) \\
&= \log(\det(X^{1/2}(I + X^{-1/2} V X^{-1/2}) X^{1/2})) \\
&= \log(\det(I + tX^{-1/2} V X^{-1/2})) + 2\log\det(X^{1/2}) \\
&= \log\left(\prod_{i=1}^{n} 1 + t\lambda_i\right) + 2\log\det(X^{1/2}) \\
&= \log\left(\prod_{i=1}^{n} (1 + t\lambda_i)\right) + \log\det(X) \\
&= \sum_{i=1}^{n} \log(1 + t\lambda_i) + \log\det(X)
\end{aligned}$$

$$= \nabla^2\left(\sum_{i=1}^{n} \log(1+t\lambda_i) + \log \cancelto{0}{\det(X)}\right)$$

$$\nabla_t g(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1+t\lambda_i}$$

$$\nabla_t^2 g(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1+t\lambda_i)^2} < 0$$

convex

$$h(t) = \mathrm{Tr}\left(A(X+tV)^{-1}\right)$$

$$= \mathrm{Tr}\left(A\left(X^{1/2}(I+X^{-1/2}VX^{-1/2})X^{1/2}\right)^{-1}\right)$$

$$= \mathrm{Tr}\left(AX^{-1/2}(I+tX^{-1/2}VX^{-1/2})^{-1}X^{-1/2}\right)$$

$$= \mathrm{Tr}\left(X^{-1/2}AX^{-1/2}(I+t\underbrace{X^{-1/2}VX^{-1/2}}_{U\Lambda U^T})^{-1}\right)$$

$$= \mathrm{Tr}\left(X^{-1/2}AX^{-1/2}(I+tU\Lambda U^T)^{-1}\right)$$

$$= \mathrm{Tr}\left(X^{-1/2}AX^{-1/2}\left(U(\cancelto{I}{U^TU}+t\Lambda)U^T\right)^{-1}\right)$$

$$= \mathrm{Tr}\left(X^{-1/2}AX^{-1/2}U(I+t\Lambda)^{-1}U^T\right)$$

$$= \mathrm{Tr}\left(U^TX^{-1/2}AX^{-1/2}U(I+t\Lambda)^{-1}\right)$$

$$= \sum_{i=1}^{n} \underbrace{(U^T X^{-1/2} A X^{-1/2} U)}_{B} \frac{1}{1+\lambda_i}$$

$$= \sum_{i=1}^{n} B_{ii} (1+\lambda_i)^{-1}$$

$$h(t) = \sum_{i=1}^{n} b_{ii} (1+\lambda_i)^{-1}$$

$$\nabla_t h(t) = \sum_{i=1}^{n} b_{ii} \; \nabla_t (1+\lambda_i)^{-1}$$

$$= - \sum_{i=1}^{n} b_{ii} \frac{\lambda_i}{1+t\lambda_i}$$

$$\nabla_t^2 h(t) = \sum_{i=1}^{n} b_{ii} \frac{2\lambda_i^2}{(1+t\lambda_i)^3} > 0$$

convex

thus $-\nabla_t^2 h(t)$ is concave

$$g''(t) - h''(t)$$

$$= -\left(- \sum_{i=1}^{n} \frac{\lambda_i^2}{(1+t\lambda_i)^2}\right) - \left(\sum_{i=1}^{n} b_{ii} \frac{2\lambda_i^2}{(1+t\lambda_i)^3} \right)$$

$$= \left(\sum_{i=1}^{n} \frac{\lambda_i^2}{1+t\lambda_i^2} - \sum_{i=1}^{n} (U^T X^{-1/2} A X^{-1/2} U) \frac{2\lambda_i^2}{(1+\lambda_i)^3} \right.$$

$$= \sum_{i=1}^{n} \frac{\lambda_i^2}{(1+t\lambda_i)^2} - \frac{\lambda_i^2 u_i^T X^{-1/2} 2A X^{-1/2} u_i}{(1+\lambda_i)^3}$$

$$= \sum_{i=1}^{n} \frac{\lambda^2(1+t\lambda_i)}{(1+t\lambda_i)^3} - \frac{\lambda_i^2 u_i^T X^{-1/2} 2A X^{-1/2} u_i}{(1+\lambda_i)^3} .$$

$$= \sum_{i=1}^{n} \frac{\lambda_i^2}{\underbrace{(1+t\lambda_i)^3}_{}} \left(\underbrace{(1+t\lambda_i) - u_i^T X^{-1/2} 2A X^{-1/2} u_i}_{}\right)$$

Lemma

$$\underbrace{0 \prec}_{} X \underbrace{\preceq 2A}_{}$$

$$z^T(X+tV)z > 0$$

$$\underbrace{z^T X^{1/2}}_{y^T}\left(I + t X^{-1/2} V X^{-1/2}\right)\underbrace{X^{1/2} z}_{y} > 0$$

$$y^T\left(I + t X^{-1/2} V X^{-1/2}\right) X^{1/2} z > 0$$

$$y^T(I + tU\Lambda U^T) y > 0$$

$$\underbrace{y^T U}_{q^T}(UU^T + t\Lambda)\underbrace{U^T y}_{q} > 0$$

$$q^T(I + t\Lambda) q > 0 \quad , \text{choose } q = \vec{e_i}$$

$$\Rightarrow (1 + t\lambda_i) > 0$$

Lemma

$$z^T(2A-(X+tV))z \geq 0$$

$$\underbrace{z^T X^{1/2}}_{y^T}(X^{-1/2}2AX^{1/2}-(I+tX^{-1/2}VX^{1/2}))\underbrace{X^{1/2}z}_{y} \geq 0$$

$$y^T(X^{-1/2}2AX^{1/2}-(I+tU\Lambda U^T))y \geq 0$$

$$\underbrace{y^T U}_{q^T}(U^T X^{-1/2}2AX^{-1/2}U-(I+t\Lambda))\underbrace{U^T y}_{q} \geq 0$$

$$q^T(U^T X^{-1/2}2AX^{-1/2}U-(I+t\Lambda))q \geq 0$$

$$q^T((I+t\Lambda)-U^T X^{-1/2}2AX^{-1/2}U)q \leq 0$$

$$q=\vec{e_i}$$

$$\Rightarrow (1+\lambda_i)-u_i^T X^{-1/2}2AX^{-1/2}u_i$$

$$\sum_{i=1}^{n}\underbrace{\frac{\lambda_i^2}{(1+t\lambda_i)^2}}_{>0}\left(\underbrace{(1+t\lambda_i)-u_i^T X^{-1/2}2AX^{-1/2}u_i}_{<1}\right)$$

Thus this is concave and

hence $f(X) = -\log\det(X) - \mathrm{Tr}(AX^{-1})$

where $\mathrm{dom}\, f = \{X \in S^n \mid 0 \prec X \preceq 2A\}$