Sample Questions

All problems are from [BV04].

Chapter 2
Problem 2.8a
Here I first offer an alternate proof from the solution from the book
S is a **polyhedron**. It is the parallelogram with corners $\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_1 - \mathbf{a}_2$, $-\mathbf{a}_1 + \mathbf{a}_2$, $-\mathbf{a}_1 - \mathbf{a}_2$, as shown in Fig. 1 for an example in $\mathbb{R}^2$. Here it shall be assumed that $\mathbf{a}_1$ and $\mathbf{a}_2$ are independent (else S_1 below will be a line, i.e. a degenerate plane, so it is less general than a plane). This parallelogram can be construction by intersection

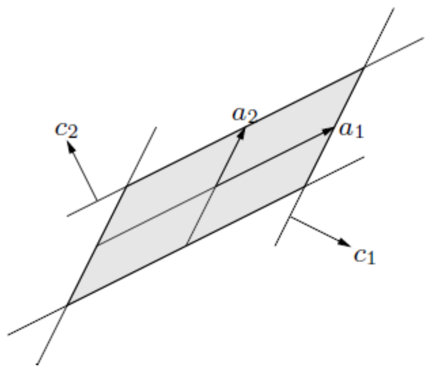

Fig. 1: Problem 2.8a.

of plane and slab, thus convex. The sets are

- $S_1 = \{\mathbf{x}\,|\mathbf{x} = y_1\mathbf{a}_1 + y_2\mathbf{a}_2,\ y_1, y_2 \in \mathbb{R}\}$: the plane defined by $\mathbf{a}_1$ and $\mathbf{a}_2$
- $S_2 = \left\{\mathbf{x} = \mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 \,\middle|\mathbf{a}_1^T\mathbf{z} = \mathbf{a}_2^T\mathbf{z} = 0,\ -1 \le y_1 \le 1\right\}$. This is a slab parallel to $\mathbf{a}_2$ and orthogonal to S_1.
- $S_3 = \left\{\mathbf{x} = \mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 \,\middle|\mathbf{a}_1^T\mathbf{z} = \mathbf{a}_2^T\mathbf{z} = 0,\ -1 \le y_2 \le 1\right\}$. This is a slab parallel to $\mathbf{a}_1$ and orthogonal to S_1.

Since S_1 is a plane formed by the linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$, it can also be expressed as the row space of $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}$, i.e.

$$S_1 = \left\{\mathbf{x} \,\middle|\, \mathbf{v}_k^T\mathbf{x} = 0, \mathbf{A}\mathbf{v}_k = \mathbf{0}, \mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}, k = 1, \dots, n-2\right\} \tag{1}$$

That is, the $dim(\mathcal{N}(\mathbf{A})) = n - 2$ since $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^n$ and assumed to be indepedent. From (1), it is concluded that S_1 is a polyhedron.

To understand why S_2 and S_3 are slabs (and thus polyhedra), let's start with S_2 (the derivation is similar to S_3. Left multiply $\mathbf{x} = \mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2$ by $\mathbf{a}_1^T$ and using the other set conditions,

$$\begin{aligned} \mathbf{a}_1^T\mathbf{x} &= \mathbf{a}_1^T\left(\mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2\right) \\ &= y_1\left\|\mathbf{a}_1\right\|_2^2 + y_2\mathbf{a}_1^T\mathbf{a}_2 \end{aligned}$$

Then

$$
\begin{aligned}
y_1 &= \frac{\mathbf{a}_1^T\mathbf{x} - y_2\mathbf{a}_1^T\mathbf{a}_2}{\|\mathbf{a}_1\|_2^2} \\
&= \mathbf{a}_1^T\left(\frac{\mathbf{x} - y_2\mathbf{a}_2}{\|\mathbf{a}_1\|_2^2}\right) \\
&= \mathbf{a}_1^T\mathbf{v} \in [-1, 1] \\
&\iff \left|\mathbf{a}_1^T\mathbf{v}\right| \leq 1.
\end{aligned}
$$

Similarly for S_3, left-multiply $\mathbf{x}$ by $\mathbf{a}_2^T$,

$$
\begin{aligned}
y_2 &= \mathbf{a}_2^T\left(\frac{\mathbf{x} - y_1\mathbf{a}_1}{\|\mathbf{a}_2\|_2^2}\right) \\
&= \mathbf{a}_2^T\mathbf{w} \in [-1, 1] \\
&\iff \left|\mathbf{a}_2^T\mathbf{w}\right| \leq 1.
\end{aligned}
$$

Hence

$$
\begin{aligned}
S_2 &= \left\{\mathbf{v} \,\middle|\, \left|\mathbf{a}_1^T\mathbf{v}\right| \leq -1\right\} \\
S_3 &= \left\{\mathbf{w} \,\middle|\, \left|\mathbf{a}_1^T\mathbf{w}\right| \leq -1\right\}
\end{aligned}
$$

which are both slabs and polyhedra. Since the intersection of polyhedra is also a polyhedra, thus S is a polyhedra.

Below is the original solution from the book

For simplicity, we assume that $\mathbf{a}_1$ and $\mathbf{a}_2$ are independent. We can express S as the intersection of three sets:

- S_1: the plane defined by $\mathbf{a}_1$ and $\mathbf{a}_2$
- $S_2 = \left\{\mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 \,\middle|\, \mathbf{a}_1^T\mathbf{z} = \mathbf{a}_2^T\mathbf{z} = 0,\ -1 \leq y_1 \leq 1\right\}$. This is a slab parallel to $\mathbf{a}_2$ and orthogonal to S_1.
- $S_3 = \left\{\mathbf{z} + y_1\mathbf{a}_1 + y_2\mathbf{a}_2 \,\middle|\, \mathbf{a}_1^T\mathbf{z} = \mathbf{a}_2^T\mathbf{z} = 0,\ -1 \leq y_2 \leq 1\right\}$. This is a slab parallel to $\mathbf{a}_1$ and orthogonal to S_1.

Each of these sets can be described with linear inequalities.

- S_1 can be described as

$$
\left\{\mathbf{x} \,\middle|\, \mathbf{v}_k^T\mathbf{x} = 0, \quad k = 1, \ldots, n-2, \mathbf{v}_k \in \mathcal{N}\left(\begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}\right)\right\}
$$

- Let $\mathbf{c}_1$ be a vector in the plane defined by $\mathbf{a}_1$ and $\mathbf{a}_2$, and orthogonal to $\mathbf{a}_2$. For example, we can take

$$
\mathbf{c}_1 = \mathbf{a}_1 - \frac{\mathbf{a}_1^T\mathbf{a}_2}{\|\mathbf{a}_2\|_2^2}\mathbf{a}_2.
$$

  Then $\mathbf{x} \in S_2$ iff

$$
-\left|\mathbf{c}_1^T\mathbf{a}_1\right| \leq \mathbf{c}_1^T\mathbf{x} \leq \left|\mathbf{c}_1^T\mathbf{a}_1\right|.
$$

- Similarly, let $\mathbf{c}_2$ be a vector in the plane defined by $\mathbf{a}_1$ and $\mathbf{a}_2$, and orthogonal to $\mathbf{a}_1$, e.g.

$$
\mathbf{c}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2^T\mathbf{a}_1}{\|\mathbf{a}_1\|_2^2}\mathbf{a}_1.
$$

  Then $\mathbf{x} \in S_3$ iff

$$
-\left|\mathbf{c}_2^T\mathbf{a}_2\right| \leq \mathbf{c}_2^T\mathbf{x} \leq \left|\mathbf{c}_2^T\mathbf{a}_2\right|.
$$

Putting it all together, we can describe S as the solution set of $2n$ linear inequalities

$$\begin{aligned} \mathbf{v}_k^T\mathbf{x} &\le 0,\ k=1,\dots,n-2 \\ -\mathbf{v}_k^T\mathbf{x} &\le 0,\ k=1,\dots,n-2 \\ \mathbf{c}_1^T\mathbf{x} &\le \left|\mathbf{c}_1^T\mathbf{a}_1\right| \\ -\mathbf{c}_1^T\mathbf{x} &\le \left|\mathbf{c}_1^T\mathbf{a}_1\right| \\ \mathbf{c}_2^T\mathbf{x} &\le \left|\mathbf{c}_2^T\mathbf{a}_2\right| \\ -\mathbf{c}_2^T\mathbf{x} &\le \left|\mathbf{c}_2^T\mathbf{a}_2\right|. \end{aligned}$$

Problem 2.8b
S is a **polyhedron**, defined by linear inequalities $x_k \ge 0$ and three equality constraints.

Problem 2.8c
S is **not a polyhedron**. It is the intersection of the unit ball $\{\mathbf{x}\,|\,\|\mathbf{x}\|_2 \le 1\}$ and the nonnegative orthant $\mathbb{R}_+^n$. This follows from the following fact, which follows from the Cauchy-Schwarz inequality:

$$\mathbf{x}^T\mathbf{y} \le 1,\ \forall \mathbf{y} \text{ with } \|\mathbf{y}\|_2 = 1 \ \Leftrightarrow \|\mathbf{x}_2\|_2 \le 1.$$

Although in this example we define S as an intersection of halfspaces, it is not a polyhedron because the definition requires infinitely many halfspaces.

Problem 2.8d
S is a **polyhedron**. S is the intersection of the set $\{\,||x_k| \le 1,\ k=1,\dots,n\}$ and the nonnegative orthant $\mathbb{R}_+^n$. This follows from the following fact:

$$\mathbf{x}^T\mathbf{y} \le 1,\ \forall \mathbf{y} \text{ with } \sum_{i=1}^n |y_i| = 1 \ \Leftrightarrow\ |x_i| \le 1,\ i=1,\dots,n.$$

We can prove this as follows. First suppose that $|x_i| \le 1$ for all i. Then

$$\mathbf{x}^T\mathbf{y} = \sum_i x_i y_i \le \sum_i |x_i||y_i| \le \sum_i |y_i| = 1$$

if $\sum_i |y_i| = 1$.
Conversely, suppose that $\mathbf{x}$ is a nonzero vector that satisfies $\mathbf{x}^T\mathbf{y} \le 1$ for all $\mathbf{y}$ with $\sum_i |y_i| = 1$. In particular we can make the following choice for $\mathbf{y}$: let k be an index for which $|x_k| = \max_i |x_i|$, and take $y_k = 1$ if $x_k > 0$, $y_k = -1$ if $x_k < 0$, and $y_i = 0$ for $i \ne k$. With this choice of $\mathbf{y}$ we have

$$\mathbf{x}^T\mathbf{y} = \sum_i x_i y_i = y_k x_k = |x_k| = \max_i |x_i|.$$

Therefore we must have $\max_i |x_i| \le 1$.
All this implies that we can describe S by a finite number of linear inequalities: it is the intersection of the nonnegative orthant with the set $\{\mathbf{x}\,|\,-\mathbf{1} \preceq \mathbf{x} \preceq \mathbf{1}\}$, i.e. the solution of $2n$ linear inequalities

$$\begin{aligned} -x_i &\le 0,\ i=1,\dots,n \\ x_1 &\le 1,\ i=1,\dots,n. \end{aligned}$$

Note that as in part (c) the set S was given as an intersection of an infinite number of halfspaces. The difference is that here most of the linear inequalities are redundant, and only a finite number are needed to characterize S.

None of these sets are affine sets or subspaces, except in some trivial cases. For example, the set defined in part (a) is a subspace (hence an affine set), if $\mathbf{a}_1 = \mathbf{a}_2 = \mathbf{0}$; the set defined in part (b) is an affine set if $n = 1$ and $S = \{1\}$; etc.

Problem 2.10

A set is convex iff its intersection with an arbitrary line $\{\widehat{\mathbf{x}} + t\mathbf{v}\,|t \in \mathbb{R}\}$ is convex.

Problem 2.10a
We have

$$(\widehat{\mathbf{x}} + t\mathbf{v})^T \mathbf{A} (\widehat{\mathbf{x}} + t\mathbf{v}) + \mathbf{b}^T (\widehat{\mathbf{x}} + t\mathbf{v}) + c = \alpha t^2 + \beta t + \gamma,$$

where

$$\alpha = \mathbf{v}^T\mathbf{A}\mathbf{v}, \quad \beta = \mathbf{b}^T\mathbf{v} + 2\widehat{\mathbf{x}}^T\mathbf{A}\mathbf{v}, \quad \gamma = c + \mathbf{b}^T\widehat{\mathbf{x}} + \widehat{\mathbf{x}}^T\mathbf{A}\widehat{\mathbf{x}}.$$

The intersection of C with the line defined by $\widehat{\mathbf{x}}$ and $\mathbf{v}$ is the set

$$\left\{\widehat{\mathbf{x}} + t\mathbf{v} |\, \alpha t^2 + \beta t + \gamma \le 0\right\},$$

which is convex if $\alpha \ge 0$. This is true for any $\mathbf{v}$, if $\mathbf{v}^T\mathbf{A}\mathbf{v} \ge 0$ for all $\mathbf{v}$, i.e. $\mathbf{A} \succeq \mathbf{0}$. The converse does not hold; for example, take $\mathbf{A} = -\mathbf{1}$, $\mathbf{b} = \mathbf{0}$, $c = -1$. Then $\mathbf{A} \not\succeq \mathbf{0}$, but $C = \mathbb{R}$ is convex.

Problem 2.10b
Let $H = \left\{\mathbf{x}\,|\mathbf{g}^T\mathbf{x} + h = 0\right\}$. We define α, β, and γ as in the solution of part (a), and in addition,

$$\delta = \mathbf{g}^T\mathbf{v}, \quad \epsilon = \mathbf{g}^T\widehat{\mathbf{x}} + h.$$

Without loss of generality, we can assume that $\widehat{\mathbf{x}} \in H$, i.e. $\epsilon = 0$. The intersection $C \bigcap H$ with the line defined by $\widehat{\mathbf{x}}$ and $\mathbf{v}$ is

$$\left\{\widehat{\mathbf{x}} + t\mathbf{v} |\, \alpha t^2 + \beta t + \gamma \le 0,\ \delta t = 0\right\}.$$

If $\delta = \mathbf{g}^T\mathbf{v} \ne 0$, the intersection is the singleton $\{\widehat{\mathbf{x}}\}$, if $\gamma \le 0$, or it is empty. In either case it is a convex set. If $\delta = \mathbf{g}^T\mathbf{v} = 0$, the set reduces to

$$\left\{\widehat{\mathbf{x}} + t\mathbf{v} |\, \alpha t^2 + \beta t + \gamma \le 0\right\},$$

which is convex if $\alpha \ge 0$. Therefore $C \bigcap H$ is convex if

$$\mathbf{g}^T\mathbf{v} = 0 \Longrightarrow \mathbf{v}^T\mathbf{A}\mathbf{v} \ge 0. \tag{2}$$

This is true if there exists λ such that $\mathbf{A} = \lambda \mathbf{g}\mathbf{g}^T \succeq \mathbf{0}$; then (2) holds, because then

$$\mathbf{v}^T\mathbf{A}\mathbf{v} = \mathbf{v}^T \left(\mathbf{A} + \lambda\mathbf{g}\mathbf{g}^T\right)\mathbf{v} \ge 0$$

for all $\mathbf{v}$ satisfying $\mathbf{g}^T\mathbf{v} = 0$.

Again, the converse is not true.

Problem 2.19a

$$\begin{aligned}
\boldsymbol{f}^{-1}(C) &= \left\{\mathbf{x} \in \operatorname{dom} \boldsymbol{f}\,\middle|\mathbf{g}^T\boldsymbol{f}(\mathbf{x}) \le h\right\} \\
&= \left\{\mathbf{x}\,\middle|\,\frac{\mathbf{g}^T(\mathbf{A}\mathbf{x} + \mathbf{b})}{\mathbf{c}^T\mathbf{x} + d} \le h,\ \mathbf{c}^T\mathbf{x} + d > 0\right\} \\
&= \left\{\mathbf{x}\,\middle|\left(\mathbf{A}^T\mathbf{g} - h\mathbf{c}\right)^T\mathbf{x} \le hd - \mathbf{g}^T\mathbf{b},\ \mathbf{c}^T\mathbf{x} + d > 0\right\},
\end{aligned}$$

which is another halfspce, intersected with $\operatorname{dom} \boldsymbol{f}$.

Problem 2.19b
The polyhedron

$$\begin{aligned}
\boldsymbol{f}^{-1}(C) &= \{\mathbf{x} \in \operatorname{dom} \boldsymbol{f}\,|\mathbf{G}\boldsymbol{f}(\mathbf{x}) \preceq \mathbf{h}\} \\
&= \left\{\mathbf{x}\,\middle|\,\frac{\mathbf{G}(\mathbf{A}\mathbf{x} + \mathbf{b})}{\mathbf{c}^T\mathbf{x} + d} \preceq \mathbf{h},\ \mathbf{c}^T\mathbf{x} + d > 0\right\} \\
&= \left\{\mathbf{x}\,\middle|\left(\mathbf{G}\mathbf{A} - \mathbf{h}\mathbf{c}^T\right)\mathbf{x} \le \mathbf{h}d - \mathbf{G}\mathbf{b},\ \mathbf{c}^T\mathbf{x} + d > 0\right\},
\end{aligned}$$

is a polyhedron intersected with dom $\boldsymbol{f}$.

Problem 2.19c

$$\begin{aligned}\boldsymbol{f}^{-1}(C) &= \left\{\mathbf{x}\in \text{dom } \boldsymbol{f} \,\middle|\, \boldsymbol{f}(\mathbf{x})^T\mathbf{P}^{-1}\boldsymbol{f}(\mathbf{x}) \le 1\right\} \\ &= \left\{\mathbf{x}\in \text{dom } \boldsymbol{f} \,\middle|\, (\mathbf{A}\mathbf{x}+\mathbf{b})^T\mathbf{P}^{-1}(\mathbf{A}\mathbf{x}+\mathbf{b}) \le \left(\mathbf{c}^T\mathbf{x}+d\right)^2\right\} \\ &= \left\{\mathbf{x} \,\middle|\, \mathbf{x}^T\mathbf{Q}\mathbf{x}+2\mathbf{q}^T\mathbf{x}\le r,\ \mathbf{c}^T\mathbf{x}+d>0\right\},\end{aligned}$$

where $\mathbf{Q}\triangleq \mathbf{A}^T\mathbf{P}^{-1}\mathbf{A}-\mathbf{c}\mathbf{c}^T$, $\mathbf{q}\triangleq \mathbf{A}^T\mathbf{P}^{-1}\mathbf{b}+d\mathbf{c}$, $r\triangleq d^2-\mathbf{b}^T\mathbf{P}^{-1}\mathbf{b}$. If $\mathbf{A}^T\mathbf{P}^{-1}\mathbf{A}\succ \mathbf{c}\mathbf{c}^T$, this is an ellipsoid intersected with dom $\boldsymbol{f}$.

Problem 2.19d
We denote by $\mathbf{a}_i^T$ as the ith row of $\mathbf{A}$.

$$\begin{aligned}\boldsymbol{f}^{-1}(C) &= \left\{\mathbf{x}\in \text{dom } \boldsymbol{f} \,\middle|\, f_1(\mathbf{x})\mathbf{A}_1+f_2(\mathbf{x})\mathbf{A}_2+\cdots+f_n(\mathbf{x})\mathbf{A}_n\preceq \mathbf{B}\right\} \\ &= \left\{\mathbf{x}\in \text{dom } \boldsymbol{f} \,\middle|\, \left(\mathbf{a}_1^T\mathbf{x}+b_1\right)\mathbf{A}_1+\cdots+\left(\mathbf{a}_n^T\mathbf{x}+b_n\right)\mathbf{A}_n\preceq \left(\mathbf{c}^T\mathbf{x}+d\right)\mathbf{B}\right\} \\ &= \left\{\mathbf{x}\in \text{dom } \boldsymbol{f} \,\middle|\, \mathbf{G}_1x_1+\cdots+\mathbf{G}_mx_m\preceq \mathbf{H},\ \mathbf{c}^T\mathbf{x}+d>0\right\}\end{aligned}$$

where

$$\begin{aligned}\mathbf{G}_i &\triangleq a_{1i}\mathbf{A}_1+a_{2i}\mathbf{A}_2+\cdots+a_{ni}\mathbf{A}_n-c_i\mathbf{B} \\ \mathbf{H} &\triangleq d\mathbf{B}-b_1\mathbf{A}_1-b_2\mathbf{A}_2-\cdots-b_n\mathbf{A}_n.\end{aligned}$$

$\boldsymbol{f}^{-1}(C)$ is the intersection of dom $\boldsymbol{f}$ with the solution set of an LMI.

Problem 2.29a
In $\mathbb{R}^2$, a cone K is a "pie slice" (see Fig. 2) In terms of polar coordinates, a pointed closed convex cone K can

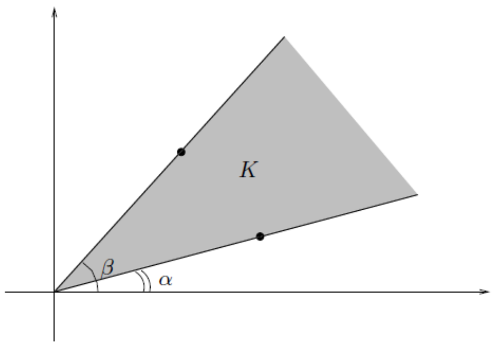

Fig. 2: Problem 2.29a.

be expressed as

$$K=\left\{(r\cos\phi, r\sin\phi)\,|\,r\ge 0, \alpha\le\phi\le\beta\right\},$$

where $0\le\beta-\alpha<180°$. When $\beta-\alpha=180°$, this gives a non-pointed cone (a halfspace). Other possible non-pointed cones are the entire plane

$$K=\left\{(r\cos\phi, r\sin\phi)\,|\,r\ge 0, 0\le\phi\le 2\pi\right\}=\mathbb{R}^2,$$

and lines through the origin

$$K=\left\{(r\cos\alpha, r\sin\alpha)\,|\,r\in\mathbb{R}\right\}.$$

Problem 2.29b
By definition, K^* is the intersection of all halfspaces $\mathbf{x}^T\mathbf{y} \geq 0$ where $\mathbf{x} \in K$. However, as can be seen from Fig. 3), the two halfspaces defined by the extreme ways are sufficient to define K^*, i.e.

$$K^* = \{\mathbf{y}\,| y_1 \cos\alpha + y_2 \sin\alpha \geq 0,\; y_1 \cos\beta + y_2 \sin\beta \geq 0\}\,.$$

If K is a halfspace, $K = \left\{\mathbf{x}\,\middle|\mathbf{v}^T\mathbf{x} \geq 0\right\}$, the dual cone is the ray $K^* = \{t\mathbf{v}|\, t \geq 0\}$. If $K = \mathbb{R}^2$, the dual

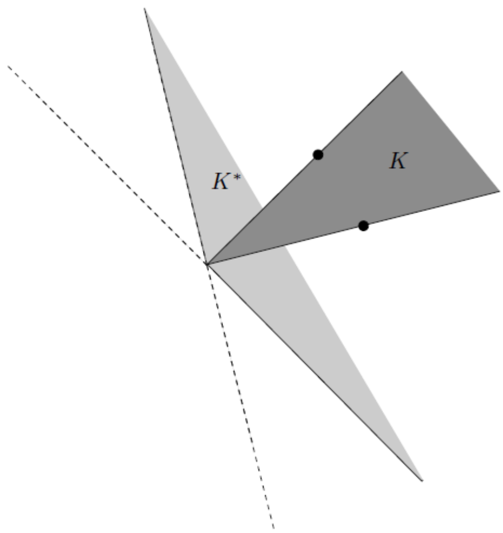

Fig. 3: Problem 2.29b.

cone is $K^* = \{0\}$. If K is a line $\{t\mathbf{v}\,|t \in \mathbb{R}\}$ through the origin, the dual cone is the line perpendicular to $\mathbf{v}$, $K^* = \left\{\mathbf{y}\,\middle|\mathbf{v}^T\mathbf{y} = 0\right\}$.

Chapter 3
Problem 3.9a
The Hessian of $\widetilde{f}$ must be positive semidefinite everywhere:

$$\nabla^2 \widetilde{f}(\mathbf{z}) = \mathbf{F}^T \nabla^2 f\left(\mathbf{F}\mathbf{z} + \widehat{\mathbf{x}}\right)\mathbf{F} \succeq \mathbf{0}.$$

Problem 3.9b
The condition in (a) means that $\mathbf{v}^T \nabla^2 f\left(\mathbf{F}\mathbf{z} + \widehat{\mathbf{x}}\right)\mathbf{v} \geq 0$ for all $\mathbf{v}$ with $\mathbf{A}\mathbf{v} = \mathbf{0}$, i.e.

$$\mathbf{v}^T\mathbf{A}^T\mathbf{A}\mathbf{v} = 0 \Longrightarrow \mathbf{v}^T f\left(\mathbf{F}\mathbf{z} + \widehat{\mathbf{x}}\right)\mathbf{v} \geq 0.$$

The result immediately follows from the hint.
Problem 3.10a
Define $f : \mathbb{R} \longrightarrow \mathbb{R}$ as

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0, \end{cases}$$

$x_0 = 0$, and scalar random variables

$$w = \begin{cases} 1, & \text{with probability } 1/2 \\ -1, & \text{with probability } 1/2 \end{cases} \qquad v = \begin{cases} 4, & \text{with probability } 1/10 \\ -4/9, & \text{with probability } 9/10. \end{cases}$$

w and v are zero-mean and

$$\text{var}(v) = \frac{16}{9} > 1 = \text{var}(w).$$

However,

$$\mathbf{E}\left[f(v)\right] = \frac{2}{5} < \frac{1}{2} = E\left[f(w)\right].$$

Problem 3.10b

$f\left(x_0 + tv\right)$ is convex in t for fixed v, hence if v is a random variable, $g(t) = E\left[f\left(x_0 + tv\right)\right]$ is a convex function of t. From Jensen's inequality,

$$g(t) = E\left[f\left(x_0 + tv\right)\right] \geq f(x_0) = g(0).$$

Now consider two points a, b, with $0 < a < b$. If $g(b) < g(a)$, then

$$\frac{b-a}{b}g(0) + \frac{a}{b}g(b) < \frac{b-a}{b}g(a) + \frac{a}{b}g(a) = g(a)$$

which contradics Jensen's inequality. Therefore we must have $g(b) \geq g(a)$.

Problem 3.14a

The condition follows directly from the 2nd-order conditions for convexity and concavity: it is

$$\nabla^2_{xx} f(\mathbf{x}, \mathbf{z}) \succeq \mathbf{0}, \quad \nabla^2_{zz} f(\mathbf{x}, \mathbf{z}) \preceq \mathbf{0},$$

for all $\mathbf{x}, \mathbf{z}$. In terms of $\nabla^2 f$, this means its $(1,1)$ block is positive semidefinite, and its $(2,2)$ block is negative semidefinite.

Problem 3.14b

Let us fix $\widetilde{\mathbf{z}}$. Since $\nabla_x f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) = \mathbf{0}$ and $f(\mathbf{x}, \widetilde{\mathbf{z}})$ is convex in $\mathbf{x}$, we conclude that $\widetilde{\mathbf{x}}$ minimizes $f(\mathbf{x}, \widetilde{\mathbf{z}})$ over $\mathbf{x}$, i.e. for all $\mathbf{z}$, we have

$$f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) \leq f(\mathbf{x}, \widetilde{\mathbf{z}}).$$

This is one of the inequalities in the saddle-point condition. We can argue in the same way about $\widetilde{\mathbf{z}}$. Fix $\widetilde{\mathbf{x}}$, and note that $\nabla_z f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) = \mathbf{0}$, together with concavity of this function in $\mathbf{z}$, means that $\widetilde{\mathbf{z}}$ maximizes the function, i.e. for any $\mathbf{x}$ we have

$$f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) \geq f(\widetilde{\mathbf{x}}, \mathbf{z}).$$

Problem 3.14c

To establish this we argue the same way. If the saddle point condition holds, then $\widetilde{\mathbf{x}}$ minimizes $f(\mathbf{x}, \widetilde{\mathbf{z}})$ over all $\mathbf{x}$. Therefore, we have $\nabla_x f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) = \mathbf{0}$. Similarly, since $\widetilde{\mathbf{z}}$ maximizes $f(\widetilde{\mathbf{x}}, \mathbf{z})$ over all $\mathbf{z}$, we have $\nabla_z f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{z}}) = 0$.

Problem 3.22a

$g(\mathbf{x}) = \log\left(\sum_{i=1}^m e^{\mathbf{a}_i^T \mathbf{x} + b_i}\right)$ 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 $f(\mathbf{x}) = h(-g(\mathbf{x}))$ is convex.

Problem 3.22b

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

Problem 3.22c

We can express f as

$$f(\mathbf{x}, u, v) = -\log u - \log\left(v - \frac{\mathbf{x}^T\mathbf{x}}{u}\right).$$

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

Problem 3.22d
We can express f as

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

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

$$g_1(\mathbf{x},t) = t^{1-\frac{1}{p}} \quad g_2(\mathbf{x},t) = t - \frac{\|\mathbf{x}\|_p^p}{t^{p-1}}.$$

Problem 3.22e
Express f as

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

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