Notes on Convex Analysis and Nonlinear Optimization by Borwein and Lewis

Forest Yang

May 30 2019

1 Chapter 1: Background

1.1 Euclidean spaces, basic definitions

Consider a finite dimensional vector space equipped with inner product $\langle \cdot, \cdot \rangle$ and norm defined by $\|x\| = \sqrt{\langle x, x \rangle}$.

- $B := \{x \in \mathbb{R}^n : \|x\| \leq 1\}$.
- Sums, differences, and products of their sets are defined by doing the operation on all pairs of elements s.t. one is from each set.
- Cone: a nonempty set C such that $\mathbb{R}_+ C = C$. Does this mean cones can be nonconvex? For example, the set $\mathrm{Span}\{u\} \cup \mathrm{Span}\{v\}$ for linearly independent vectors $u, v \in \mathbb{R}^n$ seems to be a cone under this definition, but it is not convex.
- $f : D \to \mathbb{R}$ is continuous if $f(x^i) \to x$ whenever $x^i \to x$ in D. For any $\alpha \in \mathbb{R}$, the level set $\{x \in D : f(x) \leq \alpha\}$ is closed as long as D is closed. To see this, take a sequence x^i in the level set converging to x. Since D is closed, $x \in D$. Since f is continuous, $f(x^i) \to f(x)$. Since $f(x^i) \leq \alpha$ for every i, $f(x) \leq \alpha$, so x is also in the level set. Therefore, the level set contains all its limit points, i.e. it is closed.
- Use the convention $+\infty - \infty = +\infty$, so that for any two sets $C, D \in \mathbb{R}$, we have $\inf C + \inf D = \inf(C + D)$. Also, $0(\pm\infty) = 0$.
- Given $\delta > 0$ and $g : (0, \delta) \to \mathbb{R}$,

  $$\liminf_{t \downarrow 0} g(t) := \lim_{t \downarrow 0} \inf_{(0,t)} g$$

  and the same thing for sup.
- Given a *convex* set C, $f : C \to \mathbb{R}$ is convex if ...

- Requiring f to have bounded level sets is a "growth condition," i.e. it means f must grow. It is implied by the condition

$$\liminf_{\|x\|\to\infty} \frac{f(x)}{\|x\|} > 0,$$

and surprisingly for convex functions these two growth conditions are equivalent.

1.2 Exercises for 1.1

1. Prove the intersection of an arbitrary collection of convex sets is convex. Deduce that the convex hull of a set $D \subset \mathbb{E}$ is well-defined as the intersection of all convex sets containing D.

Proof. Let $C = \bigcap_{\alpha\in I} C_\alpha$ be the intersection of convex sets C_α indexed by $\alpha \in I$. Suppose $x, y \in C$. Then, $x, y \in C_\alpha$ for all $\alpha \in I$. Thus, for any $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in C_\alpha$ for all $\alpha \in I$. i.e., $\lambda x + (1-\lambda)y \in C$. Therefore, C is convex. Thus, defining the convex hull of a set $S \subset \mathbb{E}$ as

$$\operatorname{conv} S := \bigcap_{C\subset\mathbb{E}\,:\,C \text{ cvx},\ S\subset C} C$$

is well-defined since $\operatorname{conv} S$ is guaranteed to be convex. $\square$

2.

(a) Prove that if the set $C \subset \mathbb{E}$ is convex and if $x^1, x^2, \ldots, x^m \in C$, $0 \leq \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$, and $\sum \lambda_i = 1$, then $\sum \lambda_i x^i \in C$. Prove furthermore that if $f : C \to \mathbb{R}$ is a convex function then $f\left(\sum \lambda_i x^i\right) \leq \sum \lambda_i f(x^i)$.

Proof. Proof by induction on m, assuming $\lambda_i \neq 0$ for all $i \in [m]$. If some $\lambda_i = 0$, then we may reduce m. The case $m = 1$ is trivial.
Now suppose $m \geq 2$. By the induction hypothesis applied to the case $m-1$, denoting $\lambda_i' = \frac{\lambda_i}{1-\lambda_1}$,

$$\begin{aligned}
&\sum_{i=2}^m \lambda_i' = \frac{1}{1-\lambda_1}\sum_{i=2}^m \lambda_i = \frac{1}{1-\lambda_1}(1-\lambda_1) = 1,\\
&\sum_{i=2}^m \lambda_i' x^i \in C,\\
&f\left(\sum_{i=2}^m \lambda_i' x^i\right) \leq \sum_{i=2}^m \lambda_i' f(x_i).
\end{aligned}$$

Applying convexity to x^1 and $y = \sum_{i=2}^m \lambda_i' x^i$, we get

$$\begin{aligned}
&\sum_{i=1}^m \lambda_i x^i = \lambda_1 x^1 + (1-\lambda_1)y \in C,\\
&f\left(\sum_{i=1}^m \lambda_i x^i\right) \leq \lambda_1 f(x_1) + (1-\lambda_1)f(y) \leq \sum_{i=1}^m \lambda_i f(x^i).
\end{aligned}$$

$\square$

(b) We see later (Theorem 3.1.11) that the function $-\log$ is convex on the strictly positive reals. Deduce, for any strictly positive reals $x^1, x^2, \ldots, x^m$, and any nonnegative reals $\lambda_1, \lambda_2, \ldots, \lambda_m$ with sum 1, the *arithmetic-geometric mean inequality*

$$\sum_i \lambda_i x^i \geq \prod_i (x^i)^{\lambda_i}.$$

Proof. Applying convexity of $-\log$,

$$\begin{aligned} -\log\left(\sum_{i=1}^m \lambda_i x^i\right) &\leq -\sum_{i=1}^m \lambda_i \log(x^i) = -\log\left(\prod_{i=1}^m (x^i)^{\lambda_i}\right) \\ \implies \log\left(\sum_{i=1}^m \lambda_i x^i\right) &\geq \log\left(\prod_{i=1}^m (x^i)^{\lambda_i}\right) \\ \sum_{i=1}^m \lambda_i x^i &\geq \prod_{i=1}^m (x^i)^{\lambda_i}. \end{aligned}$$

□

(c) Prove that for any set $D \subset \mathbb{E}$, $\operatorname{conv} D$ is the set of all convex combinations of elements of D.

Proof. Let C be the set of convex combinations of elements of D. Suppose $x \in C$. Then x is a convex combination of $x^1, \ldots, x^m \in D$. Note $x^1, \ldots, x^m \in \operatorname{conv} D$, and $\operatorname{conv} D$ is convex. Therefore, $x \in \operatorname{conv} D$. We have shown that $C \subseteq \operatorname{conv} D$.
Now C is a convex set containing D. Since $\operatorname{conv} D$ is the smallest convex set containing D, $\operatorname{conv} D \subseteq C$. Thus, $C = \operatorname{conv} D$. □

3. Prove that a convex set $D \subset \mathbb{E}$ has convex closure, and deduce that $\operatorname{cl} \operatorname{conv} D$ is the smallest closed convex set containing D.

Proof. Let x^i be a sequence in D converging to x and y^i a sequence in D converging to y. Ergo, x and y are in $\operatorname{cl} D$. By properties of limits of sequences under sums and products,

$$\lim_{i\to\infty} \left[\lambda x^i + (1-\lambda) y^i\right] = \lambda \lim_i x^i + (1-\lambda) \lim_i y_i = \lambda x + (1-\lambda) y.$$

Thus, convex combinations of x and y are in $\operatorname{cl} D$, i.e. $\operatorname{cl} D$ is convex. Let C be a closed convex set containing D. Thus, it contains $\operatorname{conv} D$. Because it is closed, it contains the limit points of $\operatorname{conv} D$, so it contains $\operatorname{cl} \operatorname{conv} D$. By the first part, $\operatorname{cl} \operatorname{conv} D$ is a closed convex set. Thus, it is the smallest closed convex set containing D. □

4 (Radstrom cancellation). Suppose sets $A, B, C \subset \mathbb{E}$ satisfy $A + C \subset B + C$.

(a) If A and B are convex, B is closed, and C is bounded, prove $A \subset B$.

Proof. Intuitively, the "frame" of B must contain the "frame" of A, if you drag around B and the outline after dragging it around according to adding C contains the outline of dragging A around according to adding C. For a simple example, if $\mathbb{E} = \mathbb{R}$ and $\max A > \max B$ or $\min A < \min B$, then $A + C \subset B + C$ cannot hold, because then $\max(A + C) > \max(B + C)$ or $\min(A + C) < \min(B + C)$.
We generalize/formalize this argument with support functions, which uniquely define closed convex sets. If there exists a direction $w \in \mathbb{E}^*$ such that

$$h_A(w) := \sup_{x\in A}\langle w, x\rangle > \sup_{x\in B}\langle w, x\rangle =: h_B(w),$$

then

$$\sup_{x\in A+C}\langle w, x\rangle = \sup_{x\in A}\langle w, x\rangle + \sup_{x\in C}\langle w, x\rangle > \sup_{x\in B}\langle w, x\rangle + \sup_{x\in C}\langle w, x\rangle = \sup_{x\in B+C}\langle w, x\rangle.$$

This contradicts $A + C \subset B + C$. Therefore, we must have $h_A \le h_B$. Since taking closure doesn't change the support function, $h_{\mathrm{cl}\,A} \le h_B$. By the theory of support functions applied to closed convex sets, since h_B dominates $h_{\mathrm{cl}\,A}$, $\mathrm{cl}\,A \subset B$. This implies that $A \subset B$. $\square$

(b) Show this result can fail if B is not convex.

Proof. Let A be a filled in square and let B be the capital I formed by taking the perimeter of the square, deleting the two vertical sides, and adding a vertical line in the middle. Let C be a vertical line with base at the origin that has height greater than the height of the square. Then, both $A+C$ and $B+C$ evaluate to the same filled in rectangle with height greater than twice the height of the square and width equal to the width of the square. Yet, $A \not\subset B$. $\square$

5* (Strong separation). Suppose that the set $C \subset \mathbb{E}$ is closed and convex, and that the set $D \subset \mathbb{E}$ is compact and convex.

(a) Prove the set $D - C$ is closed and convex.

Proof. If A and B are convex, it is easy to see that $A + B$ is convex. Thus, $D - C$ is convex, as it is the sum of D and $-C$ with are both convex.
Furthermore, if A is closed and B is compact, then $A + B$ is closed. Let $a_n + b_n$ be a sequence converging to $z \in \mathrm{cl}(A + B)$. Then, because B is compact, there is a convergent subsequence of b_n, i.e. $b_{n_k} \to b$. Subsequences of convergent sequences converge to the same limit, so $a_{n_k} + b_{n_k} \to z$. We have $b_{n_k} \to b \in B$ and $a_{n_k} \to z - b \in A$ because A and B are closed. Thus, $z = z - b + b \in A + B$.
This suffices to show that $D-C$ is closed and convex by taking $B = D$ and $A = -C$. $\square$

(b) Deduce that if in addition D and C are disjoint then there exists a nonzero element a in $\mathbb{E}$ with $\inf_{x\in D}\langle a, x\rangle > \sup_{y\in C}\langle a, y\rangle$. Interpret geometrically.

Proof. We know that $D - C$ is closed, and because $D \cap C = \emptyset$, does not contain 0. We claim there exists an element of $D - C$ of minimum norm. Because we are minimizing the norm, we may intersect $D - C$ with the ball of radius r, where r is the norm of an arbitrary element in $D - C$. This yields a compact set, and since the norm is continuous there is an element of minimum norm $a \neq 0$.
Now we claim that $\langle a, x\rangle \geq \|a\|^2$ for every $x \in D - C$. This would imply that $\langle a, d\rangle - \langle a, c\rangle \geq \|a\|^2 > 0$ for all $d \in D$, $c \in C$. We can take the inf over d and sup over c to get the desired statement. To prove the claim, suppose $\langle a, b\rangle < \|a\|^2$ for some $b \in D - C$. Consider

$$f(t) = \|(1-t)a + tb\|^2.$$

We have

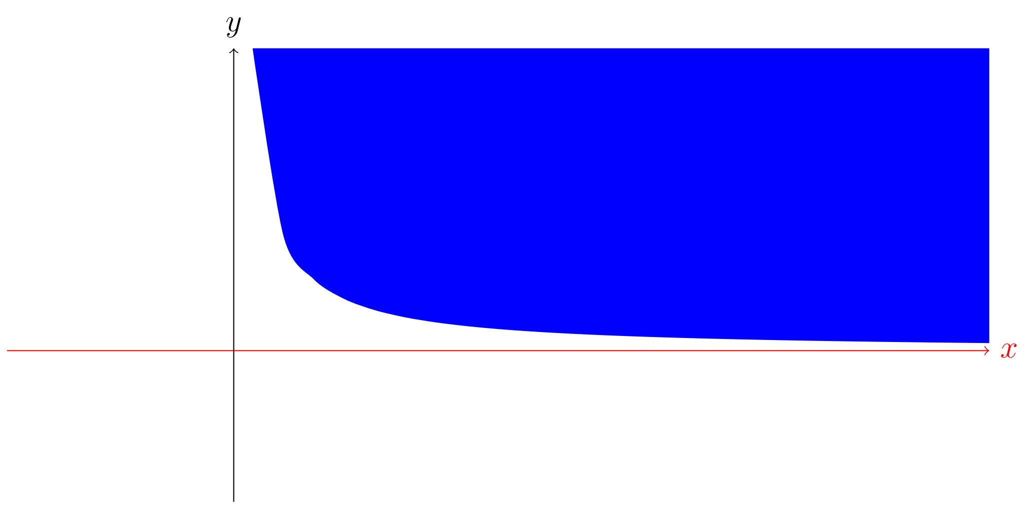
$$f'(0) = \langle a, b - a\rangle = \langle a, b\rangle - \|a\|^2 < 0.$$

This implies that for some $t > 0$, $f(t) < f(0)$. This contradicts $\|a\| = \min_{x \in D-C} \|x\|$. The interpretation of the statement is there is a hyperplane separating D and C (with normal parallel to a) without touching either of them. ∎

(c) Show part (b) fails for the closed convex sets in $\mathbb{R}^2$,

$$D = \{x \mid x_1 > 0, x_1 x_2 \geq 1\}, \quad C = \{x \mid x_2 = 0\}.$$

Proof. In the below, the blue shape is D and the red line is C. Visually it is pretty clear that there is no line strictly separating D and C, extending the x axis to infinity. Any such line must be above the x axis, and eventually there is an x_1 small enough for $x_1 y \geq 1$ where y is the height of the line.



∎

6 ** (Recession cones) Consider a nonempty closed convex set $C \subset \mathbb{E}$. We define the *recession cone* of C by

$$0^+(C) = \{d \in \mathbb{E} \mid C + \mathbb{R}_+ d \subset C\}.$$

(a) Prove $0^+(C)$ is a closed convex cone.

Proof. If d^i is a sequence in $0^+(C)$ converging to d, then $C + \mathbb{R}_+ d^i \subset C$ for every i. Because C is closed, $C + \mathbb{R}_+ d \subset C$. More explicitly, for any $r \in \mathbb{R}_+$ and $c \in C$, $c + rd^i \to c + rd$, and so $c + rd \in C$.
To show that $0^+(C)$ is convex, let $x, y \in 0^+(C)$. Then, for any $\lambda \in [0,1], r \in \mathbb{R}_+, c \in C$,

$$\lambda(c + rx) + (1 - \lambda)(c + ry) = c + r(\lambda x + (1 - \lambda)y).$$

Since c, r were arbitrary, we have $\lambda x + (1 - \lambda)y \in 0^+(C)$. Since λ, x, y were arbitrary, $0^+(C)$ is convex. $\square$

(b) Prove $d \in 0^+(C)$ iff $x + \mathbb{R}_+ d \subset C$ for some point x in C. Show this equivalence can fail if C is not closed.

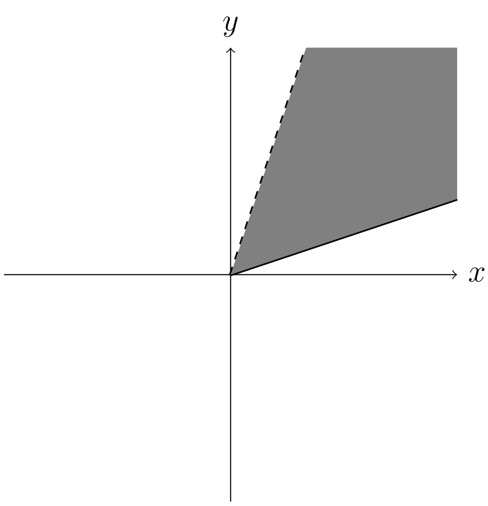
Proof. The forward implication is obvious. For the backwards implication, suppose $x + \mathbb{R}_+ d \subset C$ for some $x \in C$. Now let $y \in C, r \in \mathbb{R}_+$ be arbitrary. We know that for all $r' \in \mathbb{R}_+$ and $\lambda \in [0,1]$,

$$(1 - \lambda)y + \lambda(x + r'd) \in C.$$

If we let $\lambda = \frac{r}{r'}$ for $r' \geq r$ and then take $r' \to \infty$,

$$(1 - \lambda)y + \lambda(x + r'd) \to y + rd.$$

Because C is closed, $y + rd \in C$. This proves that $C + \mathbb{R}_+ d \subset C$.
For an example of where this is not true, take



$\square$

(c) Consider a family of closed convex sets $C_\gamma (\gamma \in \Gamma)$ with nonempty intersection. Prove $0^+(\bigcap C_\gamma) = \bigcap 0^+(C_\gamma)$.

Proof. First, the reverse inclusion. Suppose $d \in \bigcap 0^+(C_\gamma)$. Then, for every C_γ, $C_\gamma + \mathbb{R}_+ d \subset C_\gamma$. Then, if $c \in \bigcap C_\gamma$, we have $c + \mathbb{R}_+ d \subset C_\gamma$ for every $\gamma \in \Gamma$, i.e. $d \in 0^+(\bigcap C_\gamma)$.
For the forward inclusion, suppose $d \in 0^+ (\bigcap C_\gamma)$. Then for some $x \in \bigcap C_\gamma$, $x + \mathbb{R}_+ d \subset \bigcap C_\gamma$. In other words, for every C_γ, there is $x \in C_\gamma$ s.t. $x + \mathbb{R}_+ d \subset C_\gamma$. By the previous part, this implies $d \in 0^+(C_\gamma)$. Thus, $d \in \bigcap 0^+(C_\gamma)$. $\square$

(d) For a unit vector $u \in \mathbb{E}$, prove $u \in 0^+(C)$ iff there is a sequence (x^r) in C satisfying $\|x^r\| \to \infty$ and $\|x^r\|^{-1}x^r \to u$. Deduce C is unbounded if and only if $0^+(C)$ is nontrivial.

Proof. The proof is similar to that of part (b). Let $y \in C$ and $r \in \mathbb{R}_+$.

$$\forall \lambda \in [0,1],\ (1-\lambda)y + \lambda x^i \in C.$$

Let $\lambda = \frac{r}{\|x^i\|}$ for $\|x^i\| \geq r$. Then, the above converges to $y + ru$, which is in C because C is closed. Since y, r were arbitrary, $u \in 0^+(C)$.
Nontriviality of $0^+(C)$ implies there is such a unit vector u, which implies the existence of $\|x^i\| \to \infty$, i.e. unboundedness. If C is unbounded, then there is a sequence x^i s.t. $\|x^i\| \to \infty$ where no element is 0. The sequence $x^i/\|x^i\|$ is contained in the compact $\{x \in \mathbb{E} : \|x\| = 1\}$. Thus, there is a convergent subsequence. The existence of such a sequence implies its limit is in $0^+(C)$, i.e. $0^+(C)$ is nontrivial. □

(e) If $\mathbb{Y}$ is a Euclidean space, the map $A : \mathbb{E} \to \mathbb{Y}$ is linear, and $N(A) \cap 0^+(C)$ is a linear subspace, prove AC is closed. Show this result can fail without the last assumption.

Proof. Let y^i be a sequence in AC converging to y. Then, there exists a sequence $x^i \in C$ such that $Ax^i \to y$. Each x^i can be decomposed into $x^i = u^i + v^i$ where $u^i \in \text{range}(A^\top)$ and $v^i \in N(A)$. Note $Ax^i = Au^i \to y$. The pseudoinverse A^+ satisfies $A^+Ax^i = u^i$ and is linear. A^+ is continuous because it is linear. Thus,

$$u^i = A^+y^i \to A^+y.$$

In other words, the sequence u^i has a limit $u = A^+y$. Next, consider the sequence v^i. If it is bounded, we can take a convergent subsequence to obtain $u^i + v^i \to u + v \in C$, which would imply the statement. Otherwise, v^i is unbounded (while u^i is bounded). We may assume $v^i/\|v^i\|$ converges to $\hat{v}$; otherwise we can take a convergent subsequence. So, $\|u^i + v^i\| \to \infty$ and $\|u^i + v^i\|^{-1}(u^i + v^i) \to \hat{v}$. By part (d), we have $\hat{v} \in 0^+(C) \cap N(A)$. By assumption, $-\hat{v} \in 0^+(C) \cap N(A)$. This allows us to subtract off the component of $\hat{v}$ in each v^i, making v^i perpendicular to $\hat{v}$. We repeat this process until the sequence v^i is bounded – the number of repeats is finite, because $\mathbb{E}$ is finite dimensional and we are subtracting off an orthogonal component each time. Then, we can take a convergent subsequence of v^i to obtain $u^i + v^i \to u + v \in C$ satisfying $A(u+v) = y$.
This result may fail without the last assumption. Let $A : \mathbb{R}^2 \to \mathbb{R}$ be the projection onto the y coordinate, and let C be the closed convex set $\{(x,y) : x > 0, y \geq 1/x\}$. 0 is a limit point of AC, but it is not contained in AC. □

(f) Consider another nonempty closed convex set $D \subset \mathbb{E}$ such that $0^+(C) \cap 0^+(D)$ is a linear subspace. Prove $C - D$ is closed.

Proof. Let $A : C \times D \to \mathbb{E}$ be the map $A(c,d) = c - d$. Then, $C - D = A(C \times D)$. By the above part, if we can show that $N(A) \cap 0^+(C \times D)$ is a linear subspace, $C - D$ is closed. If $u \in 0^+(C)$ and $v \in 0^+(D)$, then for any $(x, y) \in C \times D, (x, y) + \mathbb{R}_+(u, v) \subset C \times D$. Also, if $(u, v) \in 0^+(C \times D)$, then $x + \mathbb{R}_+ u \subset C$ and $y + \mathbb{R}_+ v \subset D$. Thus, $0^+(C \times D) = 0^+(C) \times 0^+(D)$. Now we note

$$N(A) = \left\{ \begin{bmatrix} x \\ x \end{bmatrix} \,\middle|\, x \in C \cap D \right\}.$$

Now let $(x, x) \in N(A) \cap 0^+(C \times D)$. By the above, $x \in 0^+(C)$ and $x \in 0^+(D)$. Thus, $x \in 0^+(C) \cap 0^+(D)$. Now if $x \in (C \cap D) \cap 0^+(C) \cap 0^+(D)$, then $(x, x) \in N(A) \cap 0^+(C \times D)$. Also, for any $\alpha \in \mathbb{R}$, $\alpha x \in (C \cap D) \cap 0^+(C) \cap 0^+(D)$, because $x + (\alpha - 1)x \in C \cap D$ due to $0^+(C) \cap 0^+(D)$ being linear, which also gives $\alpha x \in 0^+(C) \cap 0^+(D)$. Thus, $(C \cap D) \cap 0^+(C) \cap 0^+(D)$ is linear. Since "doubling" each element gives $N(A) \cap 0^+(C \times D)$, $N(A) \cap 0^+(C \times D)$ is linear. Thus, $D - C$ is closed. □

7. For any set of vectors $a^1, a^2, \ldots, a^m$ in $\mathbb{E}$, prove the function $f(x) = \max_i \langle a^i, x \rangle$ is convex on $\mathbb{E}$.

Proof. Let $x, y \in \mathbb{E}, \lambda \in [0, 1]$. Let i^* be such that $f(\lambda x + (1 - \lambda)y) = \langle a^{i^*}, \lambda x + (1 - \lambda)y \rangle$. We have

$$f(\lambda x + (1 - \lambda)y) = \lambda \langle a^{i^*}, x \rangle + (1 - \lambda)\langle a^{i^*}, y \rangle \leq \lambda f(x) + (1 - \lambda)f(y).$$

□

8. Prove Proposition 1.1.3 (Weierstrass).

Proposition 1.1 (Weierstrass). *Let $D \subset \mathbb{E}$ be a nonempty closed set and $f : D \to \mathbb{R}$ continuous and with bounded level sets. Then, f has a global minimizer.*

Proof. Since D is nonempty, there exists some $\alpha \in \mathbb{R}$ such that the level set $\{x \in D : f(x) \leq \alpha\}$ is nonempty. By assumption, this level set is bounded. Furthermore, it is closed, because f is continuous. Thus, it is compact, and f attains a minimum over this set. This minimum is also a global minimum of f because elements of D outside this set evaluate to greater than α. □

9 (Composing convex functions). Suppose that the set $C \subset \mathbb{E}$ is convex and that the functions $f_1, f_2, \ldots, f_n : C \to \mathbb{R}$ are convex, and define a function $f : C \to \mathbb{R}^n$ with components f_i. Suppose further that $f(C)$ is convex and that the function $g : f(C) \to \mathbb{R}$ is convex and *isotone*: any points $y \leq z$ in $f(C)$ satisfy $g(y) \leq g(z)$. Prove the composition $g \circ f$ is convex.

Proof. Let $x, y \in C$ and $\lambda \in [0, 1]$. Since

$$\forall\, i \in [n], \quad f_i(\lambda x + (1 - \lambda)y) \leq \lambda f_i(x) + (1 - \lambda)f_i(y),$$

it follows that $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ elementwise. Therefore,

$$g(f(\lambda x + (1 - \lambda)y)) \leq g(\lambda f(x) + (1 - \lambda)f(y)) \leq \lambda g(f(x)) + (1 - \lambda)g(f(y)).$$

This proves that $g \circ f$ is convex. □

10 * (Convex growth conditions).

(a) Find a function with bounded level sets which does not satisfy the growth condition

$$\liminf_{r\to\infty} \left\{ \frac{f(x)}{\|x\|} \,\middle|\, 0 \neq x \in C,\ \|x\| = r \right\} > 0.$$

Take $f(x) = \sqrt{x}$ over the nonnegative reals.

(b) Prove any function satisfying the growth condition has bounded level sets.

Proof. Let $r = \liminf f(x)/\|x\| > 0$. There exists a constant τ such that if $\|x\| \geq \tau$, then $f(x)/\|x\| \geq r/2$. For such x, $f(x) \geq r\|x\|/2$. Therefore, letting $x \in C$ be arbitrary, if $f(x) \leq \alpha$, then either $x \leq \tau$ or $\|x\|r/2 \leq f(x) \leq \alpha$ and $\|x\| \leq 2\alpha/r$. Hence, $\|x\| \leq \max\{\tau, 2\alpha/r\}$. □

(c) Suppose the convex function $f : C \to \mathbb{R}$ has bounded level sets but that the growth condition fails. Deduce the existence of a sequence (x^m) in C with $f(x^m) \leq \|x^m\|/m \to \infty$. For a fixed point $\bar{x} \in C$, derive a contradiction by considering the sequence

$$\bar{x} + (\|x^m\|/m)^{-1}(x^m - \bar{x}).$$

Hence complete the proof that a convex function has bounded level sets iff it satisfies the growth condition.

Proof. If the growth condition fails, then

$$\liminf_{r\to\infty} \left\{ \frac{f(x)}{\|x\|} \,\middle|\, 0 \neq x \in C,\ \|x\| = r \right\} \to 0.$$

Then there exists an r such that if $\|x\| \geq r$, then $f(x)/\|x\| \leq 1/m$. We can pick such an x that also satisfies $\|x\| \geq m^2$. Thus, we deduce the existence of the sequence x^m such that $f(x^m) \leq \|x^m\|/m$, and $\|x^m\|/m \geq m \to \infty$.
The sequence $\bar{x} + (\|x^m\|/m)^{-1}(x^m - \bar{x})$ is unbounded because

$$\bar{x} + \left(\frac{\|x^m\|}{m}\right)^{-1}(x^m - \bar{x}) = \left(1 - \frac{m}{\|x^m\|}\right)\bar{x} + m\frac{x^m}{\|x^m\|}.$$

Since $m \to \infty$ and the left term goes to $\bar{x}$, the sequence goes to infinite norm. However, on this input, f evaluates to something less than

$$\left(1 - \frac{m}{\|x^m\|}\right) f(\bar{x}) + (\|x^m\|/m)^{-1} f(x^m) \leq 2|f(\bar{x})| + 1.$$

Thus, the level set corresponding to $2|f(\bar{x})| + 1$ is unbounded, giving a contradiction. □

11 ** (Accessbility lemma). Suppose C is a convex set in $\mathbb{E}$.

(a) Prove $\operatorname{cl} C \subset C + \epsilon B$ for any real $\epsilon > 0$.

Proof. For any sequence $x^i \to x$ in C, eventually x^i is ϵ close to x. Ergo, $x \in C+\epsilon B$. $\square$

(b) For sets D and F in $\mathbb{E}$ with D open, prove $D + F$ is open.

Proof.

$$D + F = \bigcup_{f \in F} D + f.$$

Each $D + f$ is an open set, and the union of open sets is open. $\square$

(c) For x in $\operatorname{int} C$ and $0 < \lambda \leq 1$, prove $\lambda x+(1-\lambda) \operatorname{cl} C \subset C$. Deduce $\lambda \operatorname{int} C+(1-\lambda) \operatorname{cl} C \subset \operatorname{int} C$.

Proof. Let $x \in \operatorname{int} C, y \in \operatorname{cl} C$. Since $x \in \operatorname{int} C$, there is a ball of radius some $\epsilon > 0$ centered at x contained in C. Let $0 < \lambda \leq 1$, and denote

$$z = \lambda x + (1 - \lambda)y.$$

Thus,

$$x = \frac{z - (1 - \lambda)y}{\lambda}.$$

Take a sequence y^i in C converging to y. Let's compute the x' for which the convex combination of x' and y^i according to λ equals z.

$$z = \lambda x' + (1 - \lambda)y^i \implies x' = \frac{z - (1 - \lambda)y^i}{\lambda}. \tag{1.2.1}$$

We have

$$\|x - x'\| = \frac{1 - \lambda}{\lambda}\|y - y^i\|.$$

Because $y^i \to y$, eventually $\|y - y^i\| \leq \epsilon\lambda/(1 - \lambda)$. This implies $\|x - x'\| \leq \epsilon$. Thus, $x' \in C$, since $x+\epsilon B \subset C$. Then (1.2.1) expresses z as a convex combination of elements in C. Thus, $z = \lambda x + (1 - \lambda)y \in C$. Thus, $\lambda x + (1 - \lambda) \operatorname{cl} C \subset C$.
$\lambda \operatorname{int} C + (1 - \lambda) \operatorname{cl} C \subset \operatorname{int} C$ then because $\lambda \operatorname{int} C + (1 - \lambda) \operatorname{cl} C$ is open, and a subset of C. $\square$

(d) Deduce $\operatorname{int} C$ is convex.

Proof. Since $(1 - \lambda) \operatorname{int} C \subset (1 - \lambda) \operatorname{cl} C$,

$$\lambda \operatorname{int} C + (1 - \lambda) \operatorname{int} C \subset \operatorname{int} C.$$

Thus, $\operatorname{int} C$ is convex. $\square$

(e) Deduce further that if $\operatorname{int} C$ is nonempty then $\operatorname{cl}\operatorname{int} C = \operatorname{cl} C$. Is convexity necessary?

Proof. Recall by part (c),

$$\lambda \operatorname{int} C + (1-\lambda)\operatorname{cl} C \subset \operatorname{int} C.$$

Given $x \in \operatorname{cl} C$, take some $y \in \operatorname{int} C$ (because $\operatorname{int} C$ assumed nonempty). Then

$$\operatorname{int} C \ni \lambda y + (1-\lambda)x \xrightarrow{\lambda \to \infty} x \in \operatorname{cl} C.$$

We have given a sequence in $\operatorname{int} C$ converging to x. Thus, $x \in \operatorname{cl}\operatorname{int} C$. This proves $\operatorname{cl} C \subset \operatorname{cl}\operatorname{int} C$. w The other inclusion is trivial. □

12 ** (Affine sets) A set L in $\mathbb{E}$ is *affine* if the entire line through any distinct points x and y in L lies in L: algebraically, $\lambda x + (1-\lambda)y \in L$ for any real λ. The *affine hull* of a set D in $\mathbb{E}$, denoted aff D, is the smallest affine set containing D. An *affine combination* of points $x^1, x^2, \ldots, x^m$ is a point of the form $\sum_{i=1}^m \lambda_i x^i$, for reals λ_i summing to 1.

(a) Prove the intersection of an arbitrary collection of affine sets is affine.

Proof. The proof is similar to the one for convex sets. Let $x, y \in \bigcap C$, the intersection of a collection of affine sets $\mathcal{C}$. For each $C \in \mathcal{C}$, $\lambda x + (1-\lambda)y \in C$ for any $\lambda \in \mathbb{R}$. Thus, $x + (1-\lambda)y \in \bigcap C$. Thus, $\bigcap C$ is affine. □

(b) Prove that a set is affine if and only if it is a translate of a linear subspace.

Proof. Suppose D is a translate of a linear subspace, i.e. $D = V + b$ for some linear subspace V and $b \in \mathbb{E}$. For any $x+b, y+b \in D$, we have $\alpha x + \beta y \in V$ for all $\alpha, \beta \in \mathbb{R}$. Thus, for any $\lambda \in \mathbb{R}$, $\lambda x + (1-\lambda)y \in V$. Thus,

$$\lambda(x+b) + (1-\lambda)(y+b) = \lambda x + (1-\lambda)y + b \in D.$$

This proves that D is affine.
Now suppose that D is affine. If it is empty, we are done. Otherwise, take some element $b \in D$, and consider the set $V = D - b$. This set is still affine, because if $x, y \in V$, then $x+b, y+b \in D$, and so $\lambda x + (1-\lambda)y + b \in D$ implying $\lambda x + (1-\lambda)y \in V$ for all $\lambda \in \mathbb{R}$. Furthermore, $0 \in V$. Thus,

$$\forall x \in V,\ \alpha \in \mathbb{R},\ \alpha x + (1-\alpha)0 = \alpha x \in V.$$

So, for any $x, y \in V$,

$$x + y = \frac{1}{2}(2x) + \left(1 - \frac{1}{2}\right) 2y \in V,$$

because we have written $x+y$ as an affine combination of elements of V.
We have shown V is linear. Then $D = V + b$ is a translate of a linear subspace. □

(c) Prove aff D is the set of all affine combinations of elements of D.

Proof. Denote $C = \{\sum_{i=1}^m \lambda_i x^i : m \in \mathbb{N},\, \sum_{i=1}^m \lambda_i = 1,\, x^1, \dots, x^m \in D\}$ as the set of affine combinations.
First we prove that C is affine. Take $a, b \in C$. Then

$$a = \sum_{i=1}^m \lambda_i x^i, \quad b = \sum_{i=1}^{m'} \lambda_i' x'^i, \quad \sum_{i=1}^m \lambda_i = \sum_{i=1}^{m'} \lambda_i' = 1,\, x^i, x'^i \in D.$$

For any $\lambda \in \mathbb{R}$,

$$\lambda a + (1-\lambda) b = \sum_{i=1}^m \lambda \lambda_i x^i + \sum_{i=1}^{m'} (1-\lambda)\lambda_i' x'^i \in C,$$

because $\sum_{i=1}^m \lambda\lambda_i + \sum_{i=1}^{m'} (1-\lambda)\lambda_i' = 1$ and $x^i, x'^i \in D$. Thus, C is affine.
Now we prove that an arbitrary affine set B contains all affine combinations of its elements. Let $x^1, \dots, x^m \in B$ and $\sum_{i=1}^m \lambda_i = 1$. We wish to show that $\sum_{i=1}^m \lambda_i x^i \in B$. If $m = 2$, this is immediate. Now we induct. If $\lambda_1 = 0$, we may reduce m. Otherwise, $z = (1-\lambda_1)^{-1} \sum_{i=2}^m \lambda_i x^i \in B$ because it is an affine combination of $m-1$ elements. Then $\lambda_1 x^1 + (1-\lambda) z \in B$ because B is affine. But this is the element we wanted to show was in B to begin with. Thus, any affine combination of elements in B is in B.
Since C is affine and contains D, we have $\operatorname{aff} D \subset C$ by minimality of $\operatorname{aff} D$. Since $\operatorname{aff} D$ contains D and is affine, it contains all affine combinations of elements of D. Thus, $C \subset \operatorname{aff} D$. Therefore, $\operatorname{aff} D = C$. ☐

(d) Prove $\operatorname{cl} D \subset \operatorname{aff} D$ and deduce $\operatorname{aff} D = \operatorname{aff}(\operatorname{cl} D)$.

Proof. To show $\operatorname{cl} D \subset \operatorname{aff} D$, it suffices to show that linear subspaces are closed, because this implies affine sets are closed and then the result follows because $\operatorname{aff} D \supset D$.
Let V be a linear subspace and suppose $u \notin V$. Then, V has an orthonormal basis $v^1, \dots, v^m$. We have

$$u = w + \sum_{i=1}^m c_i v^i, \quad w \neq 0.$$

For any $v \in V$, we then have $\|u - v\| \geq \|w\| > 0$. This eliminates the possibility of u being a limit point of V. Therefore, V is closed.
$\operatorname{aff} D$ is an affine set containing $\operatorname{cl} D$, so $\operatorname{aff} \operatorname{cl} D \subset \operatorname{aff} D$. But clearly $\operatorname{aff} D \subset \operatorname{aff} \operatorname{cl} D$, since $D \subset \operatorname{cl} D$. This proves $\operatorname{aff} D = \operatorname{aff} \operatorname{cl} D$. ☐

(e) For any point $x \in D$, prove $\operatorname{aff} D = x + \operatorname{Span}(D - x)$, and deduce the linear subspace $\operatorname{Span}(D - x)$ is independent of x.

Proof. $\operatorname{aff} D \subset x + \operatorname{Span}(D - x)$: Let $\sum_{i=1}^m \lambda_i x^i \in \operatorname{aff} D$.

$$\sum_{i=1}^m \lambda_i = 1, \forall\, i \in [m], x^i \in D \implies \sum_{i=1}^m \lambda_i x^i = x + \sum_{i=1}^m \lambda_i (x^i - x) \in x + \operatorname{Span}(D - x).$$

$x + \operatorname{Span}(D - x) \subset \operatorname{aff} D$: Let $u - x \in D - x$. We have for all $\lambda \in \mathbb{R}$,

$$x + \lambda(u - x) = (1 - \lambda)x + \lambda u \in \operatorname{aff} D.$$

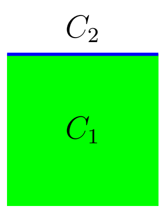
Then, for any $c_1, \ldots, c_m \in \mathbb{R}$ and $u_1, \ldots, u_m \in D$, because $\operatorname{aff} D$ is convex,

$$x + \sum_{i=1}^m c_i(u_i - x) = \frac{1}{m}\left(\sum_{i=1}^m x + mc_i(u_i - x)\right) \in \operatorname{aff} D.$$

This suffices to show that $x + \operatorname{Span}(D - x) \subset \operatorname{aff} D$. $\square$

13. ** (The relative interior) The *relative interior* of a convex set C in $\mathbb{E}$ is its interior relative to its affine hull, $\operatorname{aff} C$, denoted $\operatorname{ri} C$. In other words, a point x lies in $\operatorname{ri} C$ if there is a real $\delta > 0$ with $(x + \delta B) \cap \operatorname{aff} C \subset C$.

(a) Find convex sets $C_1 \subset C_2$ with $\operatorname{ri} C_1 \not\subset \operatorname{ri} C_2$.



(b) Suppose $\dim \mathbb{E} > 0, 0 \in C$ and $\operatorname{aff} C = \mathbb{E}$. Prove C contains a basis $\{x^1, x^2, \ldots, x^n\}$ of $\mathbb{E}$. Deduce $\frac{1}{n+1}\sum_{i=1}^n x^i \in \operatorname{int} C$. Hence deduce that any nonempty convex set in $\mathbb{E}$ has nonempty relative interior.

Proof. By exercise 12 part (e), $\operatorname{aff} C = \operatorname{Span}(C) = \mathbb{E}$. This implies that C contains a basis of $\mathbb{E}$, because we can keep choosing linearly independent vectors from C (if we stop early, this implies that $\mathbb{E}$ is the span of fewer vectors than the dimension of $\mathbb{E}$, a contradiction).

$$\frac{1}{n+1}\sum_{i=1}^n x^i \in \operatorname{int} C$$

because it is in C, and there is a ball around it contained in C. The latter follows because there exists $\epsilon > 0$ such that $\left\|\frac{1}{n+1}\sum_{i=1}^n x^i - x\right\| \leq \epsilon$ implies that the unique representation

$$x = \sum_{i=1}^n c_i x^i$$

satisfies $\sum_{i=1}^n c_i \leq 1$. This is because there is linear mapping $A : \mathbb{E} \to \mathbb{R}^n$ which maps each x to the coefficients of x in the basis, and thus $\|A(x_1 - x_2)\| \leq \|A\|\|x_1 - x_2\|$. Then any nonempty convex set in $\mathbb{E}$ has nonempty relative interior, because we can shift the set to contain 0, and then consider it as a subset of its affine hull, then apply the above argument. $\square$

(c) Prove that for $0 < \lambda \leq 1$ we have $\lambda \operatorname{ri} C + (1-\lambda) \operatorname{cl} C \subset \operatorname{ri} C$, and hence $\operatorname{ri} C$ is convex with $\operatorname{cl}(\operatorname{ri} C) = \operatorname{cl} C$.

Proof. Apply parts (c),(d),(e) of 11, setting $\mathbb{E} = \operatorname{aff} C$, so that $\operatorname{int} c = \operatorname{ri} C$. $\square$

(d) Prove that for a point $x \in C$, TFAE:

(i) $x \in \operatorname{ri} C$.

(ii) For any point $y \in C$ there exists a real $\epsilon > 0$ with $x + \epsilon(x - y)$ in C.

(iii) $\mathbb{R}_+(C - x)$ is a linear subspace.

Proof. (i) $\implies$ (ii): Because $x \in \operatorname{ri} C$, there exists an $\delta > 0$ such that $(x + \delta B) \cap \operatorname{aff} C \subset C$. If $x = y$, the result is immediate. Otherwise, $(x-y)/\|x-y\| \subset B$. Therefore,

$$x + \frac{\delta}{\|x-y\|}(x-y) \subset (x + \delta B) \cap \operatorname{aff} C \subset C.$$

Thus, we can take $\epsilon = \delta/\|x - y\|$ to prove the statement. The inclusion in $\operatorname{aff} C$ is because $x + \epsilon(x-y) = (1 - (-\epsilon))x + (-\epsilon)y$.
(ii) $\implies$ (iii): For any $y \in C$, we automatically have $y - x \in C - x$. By (ii), we also have $\epsilon(x - y) \in C - x$ for some $\epsilon > 0$. This implies that $\mathbb{R}_+(C-x)$ contains $\mathbb{R}(y-x)$. Furthermore, $\mathbb{R}_+(C - x)$ is convex, because $\mathbb{R}_+ D$ is convex for any convex set D: if $\alpha x, \beta y \in \mathbb{R}_+ D$, then for any $\lambda \in [0,1]$,

$$\lambda\alpha x + (1-\lambda)\beta y = (\lambda\alpha + (1-\lambda)\beta)\left[\frac{\lambda\alpha x}{\lambda\alpha + (1-\lambda)\beta} + \frac{(1-\lambda)\beta y}{\lambda\alpha + (1-\lambda)\beta}\right] \in \mathbb{R}_+ D.$$

This suffices to show that $\mathbb{R}_+(C-x)$ is a linear subspace.
(iii) $\implies$ (i): $\mathbb{R}_+(C - x)$ is a linear subspace implies that for any $y \in C$, in addition to $y - x \in C - x$ we have $\epsilon(x - y) \in C - x$. Now for any $z \in \operatorname{aff} C = x + \operatorname{Span} C - x$ by 12. (e),

$$z = x + \sum_{i=1}^m c_i(x^i - x), \quad \forall i \in [m],\, x^i \in C.$$

For every x^i, there exists $\epsilon_i > 0$ such that $\epsilon_i(x - x^i) \in C - x$. If $c_i \geq 0$, define $z_i = x^i - x$ and $d_i = c_i$, otherwise define $z^i = \epsilon_i(x - x^i)$ and $d_i = c_i/\epsilon_i$. Therefore,

$$z - x = \left(\sum_{i=1}^m d_i\right) \sum_{i=1}^m \frac{d_i}{\sum_{j=1}^m d_j} z^i.$$

In other words, $z - x$ is equal to a positive constant times a convex combination of elements of $C - x$. But $C - x$ is convex. Therefore, it is equal to a positive constant times an element of $C - x$, i.e. $z - x \in \mathbb{R}_+(C - x)$. Now let $e^1, \ldots, e^n$ be an orthonormal basis for aff $C - x = \text{Span}\,(C - x)$. By the previous, there exist $c_1, \ldots, c_n > 0$ such that $c_i e^i, -c_i e^i \in C - x$ for every $i \in [n]$. This implies that $0 \in \text{ri}(C - x)$, i.e. $x \in \text{ri}(C)$. □

(e) If $\mathbb{F}$ is another Euclidean space and the map $A : \mathbb{E} \to \mathbb{F}$ is linear, prove $\text{ri}\, AC \supset A\, \text{ri}\, C$.

Proof. Let $y \in A\, \text{ri}\, C$. Thus there exists a $x \in \text{ri}\, C$ such that $Ax = y$. We wish to show $y \in \text{ri}\, AC$. Take $z \in AC$. Then, $z = Av$ for some $v \in C$. Then, for some $\epsilon > 0$,

$$x + \epsilon(x - v) \in C.$$

Then,

$$y + \epsilon(y - z) = A(x + \epsilon(x - v)) \in AC.$$

This proves that for any $z \in AC$, there exists $\epsilon > 0$ such that $y + \epsilon(y - z) \in AC$. By part (d), this implies that $y \in \text{ri}\, AC$. □

1.3 1.2 Symmetric matrices

1. $\mathbb{S}^n$ is the set of $n \times n$ symmetric matrices.

2. $\mathbb{S}^n_+$ is the set of positive semidefinite matrices, matrices $X \in \mathbb{S}^n$ such that $x^\top X x \geq 0$ for all $x \in \mathbb{R}^n$. $\mathbb{S}^n_{++}$ is the set of positive definite matrices, for which the inequality is strict for all $x \neq 0$.

3. $\mathbb{S}^n$ is made into a Euclidean space by the inner product

$$\langle X, Y \rangle = \text{Tr}(XY).$$

4. Any matrix $X \in \mathbb{S}^n$ has n real eigenvalues (counted by multiplicity), which we write in nonincreasing order $\lambda_1(X) \geq \ldots \geq \lambda_n(X)$. This defines a function $\lambda : \mathbb{S}^n \to \mathbb{R}^n$.

5. Let $\text{diag} : \mathbb{R}^n \to \mathbb{S}^n$ be the linear map taking x to $\text{diag}\, x$, a diagonal matrix with the entries of x on its diagonal.

6. Let $\mathbb{O}^n$ denote the group of $n \times n$ orthogonal matrices.

7. A doubly stochastic matrix is a square matrix with all nonnegative entries, and each row and column summing to one (the set of such matrices is denoted Γ^n).

8. The *adjoint* of the linear map A is the linear map $A^* : \mathbb{Y} \to \mathbb{E}$ defined by the property

$$\langle A^* y, x \rangle = \langle y, Ax \rangle \text{ for all points } x \in \mathbb{E} \text{ and } y \text{ in } \mathbb{Y}.$$

Theorem 1.2 (1.2.1, (Fan))**.** *Any matrices X and Y in $\mathbb{S}^n$ satisfy the inequality*

$$Tr(XY) \leq \lambda(X)^\top \lambda(Y).$$

Equality holds if and only if X and Y have a simultaneous ordered spectral decomposition: there is an orthogonal matrix $U \in \mathbb{O}^n$ with

$$X = U^\top \operatorname{diag}(\lambda(X))U, \quad Y = U^\top \operatorname{diag}(\lambda(Y))U.$$

For a vector $x \in \mathbb{R}^n$, denote $[x]$ as the vector with components permuted into nonincreasing order.

Proposition 1.3 (1.2.4 (Hardy-Littlewood-Polya))**.** *Any vectors x and y in $\mathbb{R}^n$ satisfy the inequality*

$$x^\top y \leq [x]^\top [y].$$

Note: this is called the rearrangement inequality.

Theorem 1.4 (1.2.5 (Birkhoff))**.** *Doubly stochastic matrices are convex combinations of permutation matrices.*

1.4 Exercises for 1.2

1. Prove $\mathbb{S}^n_+$ is a closed convex cone with interior $\mathbb{S}^n_{++}$.

Proof. Suppose X and Y in $\mathbb{S}^n_+$, i.e. $x^\top Xx \geq 0$ and $x^\top Yx \geq 0$ for all $x \in \mathbb{R}^n$. Then for and $\lambda \in [0,1]$, $x \in \mathbb{R}^n$,

$$x^\top(\lambda X + (1-\lambda)Y)x = \lambda x^\top Xx + (1-\lambda)x^\top Yx \geq 0.$$

Therefore, $\mathbb{S}^n_+$ is convex. Also, $\mathbb{R}_+\mathbb{S}^+_n = \mathbb{S}^+_n$, as $r \geq 0$ implies $rx^\top Xx \geq 0$. To see that $\mathbb{S}^n_+$ is closed, take a sequence $X^n \to X$. Let $x \in \mathbb{R}^n$ be arbitrary. Because $x^\top(\cdot)x : \mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous,

$$0 \leq x^\top X^n x \to x^\top Xx.$$

This implies that $x^\top Xx \geq 0$. Since x was arbitrary, $X \in \mathbb{S}^n_+$. Thus, $\mathbb{S}^n_+$ is closed.
Now we show that $\mathbb{S}^n_{++} = \operatorname{int} \mathbb{S}^n_+$. If $X \in \mathbb{S}^n_{++}$, then let

$$\tau = \min_{x\in\mathbb{R}^n : \|x\|=1} x^\top Xx > 0.$$

The minimum exists because $\cdot X \cdot$ is continuous and it is taken over a compact set. Now for any $\epsilon > 0$, consider the Y in the ball of radius ϵ/n around X. For any $x \in \mathbb{R}^n$ with $\|x\| = 1$,

$$x^\top Yx = x^\top Xx + x^\top(Y-X)x \geq x^\top Xx - |x^\top(Y-X)x|. \geq \tau - |x^\top(Y-X)x|.$$

Now since $\|Y - X\| \leq \epsilon/n$, each column of $Y - X$ has norm $\leq \epsilon/n$. Let z_i denote the ith column of $Y - X$. Since $\|x\| \leq 1$, each entry $|x_i| \leq 1$. Putting these facts together,

$$|x^\top(Y-X)x| = \left| x^\top(\sum_{i=1}^n x_i z_i) \right| \leq \sum_{i=1}^n |x_i||x^\top z_i| \leq \sum_{i=1}^n |x^\top z_i| \leq n(\epsilon/n) = \epsilon.$$

If we take $\epsilon = \tau$, i.e. consider the τ/n ball around X, we get

$$x^\top Yx \geq \tau - |x^\top (Y-X)x| \geq \tau - \tau = 0.$$

Therefore, the τ/n ball around X is contained in $\mathbb{S}^n_+$. Since $X \in \mathbb{S}^n_{++}$ was arbitrary, $\mathbb{S}^n_{++}$ is open. (Actually using $\|Y-X\| = \max_{\|x\|=1} x^\top (Y-X)x$ we could have taken $\epsilon = \tau$.
Now we prove that any $X \in \mathbb{S}^n_+ \setminus \mathbb{S}^n_{++}$ does not have a ball around it is contained in $\mathbb{S}^n_+$. Let $\epsilon > 0$ be arbitrary. Let $x \neq 0$ be such that $x^\top Xx = 0$. Notice that $X - \epsilon I/\sqrt{n}$ is in the ϵ ball around X. Then

$$x^\top (X - \frac{1}{\sqrt{n}}\epsilon I)x = -\frac{\epsilon}{\sqrt{n}} x^\top Ix = -\frac{\epsilon\|x\|^2}{\sqrt{n}} < 0.$$

This proves the claim. So, any open subset of $\mathbb{S}^n_+$ does not contain any element of $\mathbb{S}^n_+ \setminus \mathbb{S}^n_{++}$. Therefore, $\mathbb{S}^n_{++} = \operatorname{int} \mathbb{S}^n_+$. □

2. Explain why $\mathbb{S}^2_+$ is not a polyhedron.

Proof. Suppose $\langle W, \cdot\rangle = C$ is the equation of a bounding hyperplane H of $\mathbb{S}^2_+$, and $X \in \mathbb{S}^2_+ \cap H$. By the previous problem, this implies that $X \in \mathbb{S}^2_+ \setminus \mathbb{S}^2_{++}$, i.e. there exists $x \neq 0$ such that $x^\top Xx = 0$. Note that this implies $Xx = 0$. To see why, if $Xx \neq 0$, then $X^{1/2}x \neq 0$. and then $x^\top Xx = \|X^{1/2}x\| > 0$. Note

$$x^\top(\cdot)x : \mathbb{S}^2 \to \mathbb{R}, \quad \dim\operatorname{range}(x^\top(\cdot)x) = 1 \implies \dim\operatorname{null}(x^\top(\cdot)x) = 2,$$

$x^\top(\cdot)x$ is a linear map on the vector space $\mathbb{S}^2$ of dimension 3; $\dim\operatorname{range}(x^\top(\cdot)x) = 1$ because $x \neq 0$, and by the rank-nullity theorem $\dim\operatorname{null}(x^\top(\cdot)x) = \dim \mathbb{S}^2 - \dim\operatorname{range}(x^\top(\cdot)x) = 2$. Furthermore, $\dim(H - X) = \dim\operatorname{null} W = 2$ ($W \neq 0$). Therefore, there are two cases.
1. $\operatorname{null} W \neq \operatorname{null}(x^\top(\cdot)x)$. Since both are dimension 2, there is a nonzero matrix in one that is not in the other. This implies $\exists Y \in \operatorname{null} W$ s.t. $x^\top Yx < 0$, so that $\forall \epsilon > 0, X + \epsilon Y \in H$ and $x^\top(X + \epsilon Y)x < 0$, meaning that the affine dimension of $\mathbb{S}^2_+ \cap H$ is at most 1.
2. $\operatorname{null} W = \operatorname{null}(x^\top(\cdot)x)$. Subspaces of dimension 2 define a unique orthogonal direction, in a dimension 3 vector space. Thus, this implies $W = cxx^\top$ for some $c \neq 0 \in \mathbb{R}$, as $\langle xx^\top, Y\rangle = x^\top Yx$ for any $Y \in \mathbb{S}^n$. We can assume $c = 1$ by replacing C with C/c. If $C < 0$, we have a contradiction as $\langle W, X\rangle = x^\top Xx = C < 0$, but $x^\top Xx \geq 0$. If $C > 0$, then $2Cxx^\top/\|x\|^4$ and $Cxx^\top/2\|x\|^4$ are on opposite sides of the hyperplane, yet are both in $\mathbb{S}^n_+$, contradicting H being a bounding hyperplane. If $C = 0$, then $X \in \mathbb{S}^2_+ \cap H$ implies $X = \sigma x_\perp x_\perp^\top$ where $x_\perp$ points in the unique direction perpendicular to x (unique because $x \in \mathbb{R}^2$) and σ is some nonnegative real number. In other words, $\mathbb{S}^2_+ \cap H$ has affine dimension 1 (it consists of nonnegative multiples of $x_\perp x_\perp^\top$).
In either case, if H is a valid bounding hyperplane then $\mathbb{S}^2_+ \cap H$ has affine dimension 1. If $\mathbb{S}^2_+$ is a polyhedron, then it is the intersection of a finite number of hyperplanes, meaning its boundary is the union of a finite number of sets of affine dimension 1 (implying the boundary is contained in the union of a finite number of lines). This seems preposterous. □

3 ($\mathbb{S}^3_+$ is not strictly convex). Find nonzero matrices X and Y in $\mathbb{S}^3_+$ such that $\mathbb{R}_+X \neq \mathbb{R}_+Y$ and $(X+Y)/2 \notin \mathbb{S}^3_{++}$.

Proof. Let $\{x, y, z\}$ be an orthogonal basis. Take $X = xx^\top$ and $Y = yy^\top$. $z^\top W z = 0$ for all $W \in \{X, Y, 1/2(X+Y)\}$, so all of them are not in the interior. But all of them are PSD. $\square$

4 (A nonlattice ordering). Suppose the matrix Z in $\mathbb{S}^2$ satisfies

$$W \succeq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } W \succeq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \iff W \succeq Z.$$

(a) By considering diagonal W, prove

$$Z = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$

for some real a.

Proof. If a diagonal entry of Z is less than 1, then $W = Z$ satisfies $W \succeq Z$, but W is not $\succeq$ both the matrices with 0s and 1s, because after subtracting the one with a one in that diagonal entry the resulting matrix has a negative entry on its diagonal.
If a diagonal entry of Z is strictly greater than one, then the identity I is $\succeq$ the matrices with 0s and 1s but is not $\succeq Z$. $\square$

(b) By considering $W = I$, prove $Z = I$.

Proof. We must have $I \succeq Z$. If $a > 0$, then $x^\top(I - Z)x < 0$ with $x = [1\ -1]$. If $a < 0$, then $x^\top(I - Z)x < 0$ with $x = [1\,1]$. Therefore, $a = 0$. $\square$

(c) Derive a contradiction by considering

$$W = \frac{2}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Proof.

$$W - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 2 \end{bmatrix},$$

which has positive trace and determinant 0, implying PSD.

$$W - Z = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3}. \end{bmatrix}$$

This has negative determinant, implying not PSD. $\square$

5 (Order preservation).

(a) Prove any matrix $X \in \mathbb{S}^n$ satisfies $(X^2)^{1/2} \succeq X$.

Proof. Since X is symmetric, it has an orthogonal eigendecomposition $X = UDU^\top$:

$$(X^2)^{1/2} = (UD^2U^\top) = U|D|U^\top \succeq UDU^\top = X,$$

where $|D|$ is the diagonal matrix with the absolute values of the entries of D. $\square$

(b) Find matrices $X \succeq Y$ in $\mathbb{S}^2_+$ such that $X^2 \not\succeq Y^2$.

Proof.

$$X = \begin{bmatrix} 1+\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad Y = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \implies X - Y = \begin{bmatrix} \frac{1}{2}+\frac{1}{\sqrt{2}} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{\sqrt{2}}-\frac{1}{2} \end{bmatrix}.$$

X and Y are both PSD, and $\mathrm{Tr}(X-Y) = \sqrt{2}$ and $\det(X-Y) = 0$, which implies that $X - Y$ is PSD. On the other hand, noting $Y^2 = Y$,

$$X^2 = \begin{bmatrix} \frac{3}{2}+\sqrt{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \implies X^2 - Y^2 = \begin{bmatrix} 1+\sqrt{2} & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}.$$

Now $X^2 - Y^2$ has a negative determinant. This implies it is not PSD. $\square$

(c) For matrices $X \succeq Y$ in $\mathbb{S}^n_+$ prove $X^{1/2} \succeq Y^{1/2}$. (Hint: consider the relationship

$$\langle (X^{1/2}+Y^{1/2})x, (X^{1/2}-Y^{1/2})x\rangle = \langle (X-Y)x, x\rangle \geq 0,$$

for eigenvectors x of $X^{1/2} - Y^{1/2}$.)

Proof. The relationship follows because

$$\begin{aligned}\langle (X^{1/2}+Y^{1/2})x, (X^{1/2}-Y^{1/2})x\rangle &= \langle Xx, x\rangle - \langle X^{1/2}Y^{1/2}x, x\rangle + \langle Y^{1/2}X^{1/2}x, x\rangle - \langle Yx, x\rangle \\ &= \langle Xx, x\rangle - \langle Y^{1/2}x, X^{1/2}x\rangle + \langle X^{1/2}x, Y^{1/2}x\rangle - \langle Yx, x\rangle \\ &= \langle (X-Y)x, x\rangle \geq 0.\end{aligned}$$

Now let x be an eigenvector of $X^{1/2} - Y^{1/2}$ with eigenvalue λ. By the relationship,

$$\lambda\langle (X^{1/2}+Y^{1/2})x, x\rangle \geq 0.$$

Suppose $\langle (X^{1/2}+Y^{1/2})x, x\rangle = 0$. This implies that x is in $\mathrm{null}(X^{1/2}) \cap \mathrm{null}(Y^{1/2})$, because we showed in an earlier question that the existence of a square root for PSD A implies $x^\top Ax = 0$ implies $x \in \mathrm{null}(A)$, and since $X^{1/2}$ and $Y^{1/2}$ are both PSD if either $x^\top X^{1/2}x$ or $x^\top Y^{1/2}x$ are positive the sum is positive. This further implies that $(X^{1/2} - Y^{1/2})x = 0$, i.e. $\lambda = 0$.
Otherwise, $\langle (X^{1/2}+Y^{1/2})x, x\rangle > 0$. This implies that $\lambda \geq 0$.
We have shown that any eigenvalue of $X^{1/2}-Y^{1/2}$ is nonnegative. Since it is symmetric, this implies that $X^{1/2} - Y^{1/2}$ is PSD. $\square$

6 * (Square-root iteration). Suppose a matrix A in $\mathbb{S}^n_+$ satisfies $I \succeq A$. Prove that the iteration

$$Y_0 = 0, \quad Y_{n+1} = \frac{1}{2}(A + Y_n^2) \quad (n = 0, 1, 2, \dots)$$

is nondecreasing (that is, $Y_{n+1} \succeq Y_n$ for all n) and converges to the matrix $I - (I - A)^{1/2}$. (Hint: consider diagonal matrices A.)

Proof. First consider diagonal matrices A. It is clear that Y_n is diagonal for every n. Then we just need to prove that for each $i = 1, \dots, n$, the ith diagonal entry of Y_n, denoted $Y_{n,ii}$, converges to $1 - \sqrt{1 - A_{ii}}$, and furthermore that each $(Y_{n,ii})_{n=0}^{\infty}$ is a nondecreasing sequence. For simplicity, denote $a = A_{ii}$ and $y_n = Y_{n,ii}$. We have

$$y_{n+1} = \frac{1}{2}(a + y_n^2) =: f(y_n).$$

Notice that if $\ell = 1 - \sqrt{1 - a}$, then

$$f(\ell) = \frac{1}{2}(a + \ell^2) = \frac{1}{2}(a + 1 - 2\sqrt{1 - a} + 1 - a) = 1 - \sqrt{1 - a} = \ell.$$

Therefore,

$$\ell - y_{n+1} = f(\ell) - f(y_n) = \frac{1}{2}(\ell^2 - y_n^2) = \frac{1}{2}(\ell + y_n)(\ell - y_n).$$

Assuming $\ell - y_n \geq 0$ (as is true for $n = 0$), we have $\ell - y_{n+1} \geq 0$, and since $(\ell + y_n)/2 \leq 1$ since $\ell \leq 1$ and $y_n \leq 1$, we have $\ell - y_{n+1} \leq \ell - y_n$ i.e. $y_{n+1} \geq y_n$. By Bernoulli's inequality

$$\ell = 1 - \sqrt{1 - a} \leq 1 - (1 - a/2) = \frac{a}{2} \leq \frac{1}{2} \implies \ell - y_{n+1} \leq \frac{3}{4}(\ell - y_n).$$

Coupled with nonnegativity of $\ell - y_n$ for all $n \geq 0$, this implies that $y_n \to \ell$.
Backtracking, we have really shown that $Y_n \to I - (I - A)^{1/2}$ for diagonal A. But notice that we can consider diagonal A WLOG: If $A = UDU^\top$ is an orthogonal eigendecomposition,

$$U^\top Y_{n+1} U = \frac{1}{2}(U^\top A U + U^\top Y_n^2 U) = \frac{1}{2}(D + (U^\top Y_n U)^2).$$

Therefore, $U^\top Y_n U \to I - (I - D)^{1/2}$. Therefore,

$$Y_n \to U(I - (I - D)^{1/2})U^\top = I - U(I - D)^{1/2}U^\top = I - (I - A)^{1/2}.$$

Also, $U^\top Y_{n+1} U \succeq U^\top Y_n U \implies Y_{n+1} \succeq Y_n$. □

7 (The Fan and Cauchy-Schwarz inequalities).

(a) For any matrices $X \in \mathbb{S}^n$ and $U \in \mathbb{O}^n$, prove $\|U^\top X U\| = \|X\|$.

Proof.

$$\mathrm{Tr}(U^\top X U U^\top X U) = \mathrm{Tr}(U^\top X X U) = \mathrm{Tr}(X X U U^\top) = \mathrm{Tr}(X X).$$

□

(b) Prove the function λ is norm-preserving.

$$\|\lambda(X)\| = \|\operatorname{diag}\lambda(X)\| = \|U\operatorname{diag}\lambda(X)U^\top\| = \|X\|,$$

because there is an eigendecomposition $X = U\operatorname{diag}\lambda(X)U^\top$.

(c) Explain why Fan's inequality is a refinement of the Cauchy-Schwarz inequality.

Proof. By Cauchy Schwarz, we know $\langle X, Y\rangle \leq \|X\|\|Y\| = \|\lambda(X)\|\|\lambda(Y)\|$. Fan's inequality is a refinement because $\langle X, Y\rangle \leq \lambda(X)^\top\lambda(Y)$ is stronger – $\lambda(X)^\top\lambda(Y) \leq \|\lambda(X)\|\|\lambda(Y)\|$. □

8. Prove the inequality $\operatorname{Tr} Z + \operatorname{Tr} Z^{-1} \geq 2n$ for all matrices $Z \in \mathbb{S}^n_{++}$, with equality if and only if $Z = I$.

Proof. Consider the matrices

$$A = \begin{bmatrix} Z & 0 \\ 0 & I \end{bmatrix}, \quad B = -\begin{bmatrix} I & 0 \\ 0 & Z^{-1} \end{bmatrix}.$$

If the eigenvalues of Z are $\lambda_1 \geq \ldots \geq \lambda_n > 0$, we see that the eigenvalues of A are $\lambda_1, \ldots, \lambda_n, \overbrace{1, \ldots, 1}^{n \text{ times}}$ and those of B are $-\lambda_1^{-1}, \ldots, -\lambda_n^{-1}, \overbrace{-1, \ldots, -1}^{n \text{ times}}$. Notice for any $r_1, r_2 > 0 \in \mathbb{R}$, $r_1 \geq r_2 \iff -r_1^{-1} \geq -r_2^{-1}$, and the eigenvalues of B are precisely those of A, negated and inverted. Thus, the ranked eigenvalues of B are the ranked eigenvalues of A, negated and inverted. Now applying Fan's inequality,

$$-\operatorname{Tr} Z - \operatorname{Tr} Z^{-1} = \operatorname{Tr}(AB) \leq \lambda(A)^\top\lambda(B) = -2n \implies 2n \leq \operatorname{Tr} Z + \operatorname{Tr} Z^{-1}.$$

Equality occurs iff A and B have the same ordered eigendecomposition, which happens iff $Z = I$. □

9. Prove the Hardy-Littlewood-Polya inequality (Proposition 1.2.4) directly.

Proof. Let $a, b \in \mathbb{R}^n$ and WLOG, suppose $a_1 \geq a_2 \geq \ldots \geq a_n$. Let i be the first index such that $b_i \neq [b]_i$. Then, $b_i < [b]_i$ and $\exists j > i$, $b_j = [b]_i$. We have $a^\top b \leq a^\top b'$, where b' has b_j and b_i swapped, since

$$a^\top b' - a^\top b = b_j a_i + b_i a_j - b_j a_j - b_i a_i = (b_j - b_i)(a_i - a_j) \geq 0.$$

We have increased the number of entries that b agrees with $[b]$ in a row. By doing this repeatedly, we obtain $a^\top b \leq [a]^\top[b]$. □

10. Given a vector $x \in \mathbb{R}^n_+$ satisfying $x_1 x_2 \ldots x_n = 1$, define numbers $y_k = 1/x_1 x_2 \ldots x_k$ for each index $k = 1, 2, \ldots, n$. Prove

$$x_1 + x_2 + \ldots + x_n = \frac{y_n}{y_1} + \frac{y_1}{y_2} + \ldots + \frac{y_{n-1}}{y_n}.$$

By applying the Hardy-Littlewood-Polya inequality to suitable vectors, prove $x_1+x_2+\ldots+x_n \geq n$. Deduce the inequality

$$\frac{1}{n}\sum_{i=1}^{n} z_i \geq \left(\prod_{i=1}^{n} z_i\right)^{1/n}$$

for any vector $z \in \mathbb{R}^n_+$.

Proof. For $i \geq 1$, we have

$$\frac{y_i}{y_{i+1}} = \frac{1/x_1\ldots x_i}{1/x_1\ldots x_{i+1}} = x_{i+1},$$

and $y_n/y_1 = 1/y_1 = x_1$. This proves the first equation. We apply HLP to vectors $a = [y_n, y_1, \ldots, y_{n-1}]$ and $b = [-y_1^{-1}, \ldots, -y_n^{-1}]$. Since $r_1 \geq r_2 \iff r_1^{-1} \geq r_2^{-1}$ for $r_1, r_2 > 0$, we have

$$a^\top b = -\frac{y_n}{y_1} - \ldots - \frac{y_{n-1}}{y_n} = -x_1 - \ldots - x_n \leq -n = [a]^\top [b].$$

In other words, $x_1+\ldots+x_n \geq n$. Now let $x \in \mathbb{R}^n_+$. If some entry is 0, the AM-GM inequality is immediate. Otherwise, divide each entry by the GM of the entries. Now we may apply:

$$\sum_{i=1}^{n} \frac{x_i}{(\prod_{j=1}^{n} x_j)^{1/n}} \geq n \implies \frac{1}{n}\sum_{i=1}^{n} x_i \geq \left(\prod_{i=1}^{n} x_i\right)^{1/n}.$$

□

11. For a fixed column vector $s \in \mathbb{R}^n$, define a linear map $A : \mathbb{S}^n \to \mathbb{R}^n$ by setting $AX = Xs$ for any matrix $X \in \mathbb{S}^n$. Calculate the adjoint map A^*.

Proof. Notice that for $y \in \mathbb{R}^n$,

$$\langle y, AX\rangle = y^\top X s = \langle y s^\top X\rangle = \langle s y^\top, X\rangle = \langle \frac{1}{2}(s y^\top + y s^\top), X\rangle.$$

Therefore, $\langle y, AX\rangle = \langle A^* y, X\rangle$ with $A^* : \mathbb{R}^n \to \mathbb{S}^n$ defined as $A^* y = \frac{1}{2}(s y^\top + y s^\top)$. □

12 * (Fan's inequality). For vectors x and y in $\mathbb{R}^n$ and a matrix U in $\mathbb{O}^n$, define

$$\alpha = \langle \operatorname{diag} x, U^\top \operatorname{diag} y U\rangle.$$

(a) Prove $\alpha = x^\top Z y$ for some doubly stochastic matrix Z.

Proof.

$$\begin{aligned}\alpha &= \operatorname{Tr}(\operatorname{diag} x U^\top \operatorname{diag} y U) = \operatorname{Tr}\left(\sum_{i=1}^{n} x_i e_i e_i^\top U^\top \sum_{j=1}^{n} y_j e_j e_j^\top U\right)\\
&= \sum_{i,j=1}^{n} x_i y_j \operatorname{Tr} e_i e_i^\top U^\top e_j e_j^\top U = \sum_{i,j=1}^{n} x_i y_j \operatorname{Tr} e_i^\top U^\top e_j e_j^\top U e_i\\
&= \sum_{i,j=1}^{n} x_i y_j (e_i^\top U^\top e_j)(e_j^\top U e_i) = \sum_{i,j=1}^{n} x_i y_j U_{ji}^2 = x^\top (U^\top)^2 y.\end{aligned}$$

$U^\top$ is orthonormal, so U is orthonormal, so after squaring each entry, each column and row sums to 1. In other words, $(U^\top)^2$ is doubly stochastic. □

(b) Use Birkhoff's theorem and Proposition 1.2.4 to deduce the inequality $\alpha \leq [x]^\top [y]$.

Proof. By Birkhoff's theorem, $(U^\top)^2 = \sum_\sigma \lambda_\sigma P_\sigma$ where P_σ is the permutation matrix corresponding to permutation σ and $\sum_\sigma \lambda = 1, \lambda_\sigma \geq 0$. Now we can apply HLP:

$$\alpha = \sum_\sigma \lambda_\sigma x^\top P_\sigma y \leq \sum_\sigma \lambda_\sigma [x]^\top [y] = [x]^\top [y].$$

□

(c) Deduce Fan's inequality (1.2.2).

Proof. Given symmetric matrices X and Y, we have

$$\begin{aligned}\langle X, Y\rangle &= \langle U \operatorname{diag}\lambda(X) U^\top, V \operatorname{diag}\lambda(Y) V^\top\rangle = \langle \operatorname{diag}\lambda(X), U^\top V \operatorname{diag}\lambda(Y) V^\top U\rangle \\ &\leq [\lambda(X)]^\top [\lambda(Y)] = \lambda(X)^\top \lambda(Y)\end{aligned}$$

by orthogonality of $U^\top V$ and the previous parts. □

13 (A lower bound). Use Fan's inequality for two matrices X and Y in $\mathbb{S}^n$ to prove a *lower* bound for $\operatorname{Tr}(XY)$ in terms of $\lambda(X)$ and $\lambda(Y)$.

Proof.

$$-\operatorname{Tr}(XY) = \operatorname{Tr}(-XY) \leq \lambda(-X)^\top \lambda(Y) = -]\lambda(X)[^\top \lambda(Y) \implies \operatorname{Tr}(XY) \geq\,]\lambda(X)[^\top \lambda(Y),$$

where $]x[$ denotes the vector x with entries sorted in nondecreasing order. □

14 * (Level sets of perturbed log matrices).

(a) For $\delta \in \mathbb{R}_{++}$, prove the function

$$t \in \mathbb{R}_{++} \mapsto \delta t - \log t$$

has compact level sets.

Proof. because $f(t) = \delta t - \log t$ is continuous, its level sets $\{t \in \mathbb{R}_{++} : f(t) \leq \alpha\}$ are closed. Now we just need to show they are bounded. Notice that $f(t)/t = \delta - \log t/t \to \delta > 0$ as $t \to \infty$, because $\log t/t \to 0$. Therefore, f satisfies the growth condition

$$\lim_{r\to\infty} \inf \left\{ \frac{f(t)}{t} \,\middle|\, t \in \mathbb{R}_{++}, \|t\| > r \right\} > 0,$$

and thus has bounded level sets. □

(b) For $c \in \mathbb{R}^n_{++}$ prove the function

$$x \in \mathbb{R}^n_{++} \mapsto c^\top x - \sum_{i=1}^n \log x_i$$

has compact level sets.

Proof. Observe that, denoting $\underline{c} = \min_{i \in [n]} c_i > 0$ and x^* as the maximum entry of x,

$$g(x) := c^\top x - \sum_{i=1}^n \log(x_i) \geq \underline{c}x^* - n\log(x^*). \tag{1.4.1}$$

By the previous part, $\underline{c}x^* - n\log(x^*)$ has bounded level sets as a function of $x^* \in \mathbb{R}_{++}$. This is equivalent to saying, given $\alpha \in \mathbb{R}$, $\exists \tau > 0$ such that

$$x^* > \tau \implies \underline{c}x^* - n\log(x^*) > \alpha.$$

By (1.4.1), we have

$$x^* > \tau \implies g(x) > \alpha,$$

or $g(x) \leq \alpha \implies x^* \leq \tau$. In other words, $g(x) \leq \alpha$ implies the maximum entry of x is bounded. In other words, x is bounded. This implies the sublevel sets of g are bounded, thus compact by continuity. $\square$

(c) For C in $\mathbb{S}^n_{++}$, prove the function

$$X \in \mathbb{S}^n_{++} \mapsto \langle C, X\rangle - \log\det X$$

has compact level sets. (Hint: Use Exercise 13.)

Proof. Note that $\det X = \prod_{i=1}^n \lambda(X)_i$. Using Exercise 13,

$$\langle C, X\rangle - \log\det X = \langle C, X\rangle - \sum_{i=1}^n \log\lambda(X)_i \geq]\lambda(C)[^\top \lambda(X) - \sum_{i=1}^n \log\lambda(X)_i.$$

Note that $\lambda(C), \lambda(X) \in \mathbb{R}^n_{++}$ since $C, X \in \mathbb{S}^n_{++}$. By the previous part, if $\langle C, X\rangle - \log\det X \leq \alpha$, then we must have $\|\lambda(X)\| \leq \tau$ for some $\tau > 0$. In fact, $\|X\| = \|\lambda(X)\|$. This implies the compactness of sublevel sets of $\langle C, X\rangle - \log\det X$. $\square$

15 * (Theobald's condition). Assuming Fan's inequality, complete the proof of Fan's theorem as follows. Suppose equality holds in Fan's inequality, and choose a spectral decomposition

$$X + Y = U^\top(\operatorname{diag}\lambda(X+Y))U$$

for some matrix $U \in \mathbb{O}^n$.

(a) Prove $\lambda(X)^\top \lambda(X+Y) = \langle U^\top \operatorname{diag}\lambda(X)U, X+Y\rangle$.

Proof.

$$\mathrm{Tr}(U^\top \operatorname{diag}\lambda(X)U(X+Y)) = \mathrm{Tr}(U^\top \operatorname{diag}\lambda(X)UU^\top \operatorname{diag}\lambda(X+Y)U) = \lambda(X)^\top\lambda(X+Y).$$

□

(b) Apply Fan's inequality to the two inner products

$$\langle X, X+Y\rangle \quad \text{and} \quad \langle U^\top \operatorname{diag}\lambda(X)U, Y\rangle$$

to deduce $X = U^\top \operatorname{diag}\lambda(X)U$.

Proof. By Fan's inequality:

$$\begin{aligned}\langle X, X+Y\rangle &\le \lambda(X)^\top\lambda(X+Y) = \langle U^\top \operatorname{diag}(\lambda(X))U, X+Y\rangle \\ &\le \langle U^\top \operatorname{diag}\lambda(X)U, X\rangle + \lambda(X)^\top\lambda(Y).\end{aligned}$$

By assumption, $\langle X, Y\rangle = \lambda(X)^\top\lambda(Y)$. Thus, $\langle X, X\rangle \le \langle U^\top \operatorname{diag}\lambda(X)U, X\rangle$. Notice that $\|U^\top \operatorname{diag}\lambda(X)U\| = \|X\|$. By Cauchy-Schwarz, $\langle U^\top \operatorname{diag}\lambda(X)U, X\rangle < \langle X, X\rangle$ unless $U^\top \operatorname{diag}\lambda(X)U = X$. Therefore, $X = U^\top \operatorname{diag}\lambda(X)U$. □

(c) Deduce Fan's theorem.

Proof. By symmetry of the argument for Y, we have $Y = U^\top \operatorname{diag}(Y)U$ as well. We have shown that equality in Fan's inequality implies a simultaneous ordered spectral decomposition for X and Y. If the decomposition exists, then

$$\mathrm{Tr}(XY) = \mathrm{Tr}(U^\top \operatorname{diag}\lambda(X)UU^\top \operatorname{diag}\lambda(Y)U) = \mathrm{Tr}(\operatorname{diag}\lambda(X)\operatorname{diag}\lambda(Y)) = \lambda(X)^\top\lambda(Y).$$

This proves the iff part of Fan's theorem. □

16 ** (Generalizing Theobald's condition). Consider a set of matrices $X^1, X^2, \ldots, X^m$ in $\mathbb{S}^n$ satisfying the conditions

$$\mathrm{Tr}(X^iX^j) = \lambda(X^i)^\top\lambda(X^j) \quad \text{for all } i \text{ and } j.$$

Generalize the argument of exercise 15 to prove the entire set of matrices $\{X^1, X^2, \ldots, X^m\}$ has a simultaneous ordered spectral decomposition.

Proof. Let $U \operatorname{diag}\lambda(X^1 + \ldots + X^m)U^\top$ be a spectral decomposition of $\sum_{i=1}^m X^i$. Following the same general argument,

$$\begin{aligned}\mathrm{Tr}(X^iX^i) + \sum_{j\neq i}\lambda(X^i)^\top\lambda(X^j) &= \langle X^i, \sum_{j=1}^m X^j\rangle \le \langle U \operatorname{diag}\lambda(X^i)U^\top, \sum_{j=1}^m X^j\rangle \\ &\le \langle U \operatorname{diag}\lambda(X^i)U^\top, X^i\rangle + \sum_{j\neq i}\lambda(X^i)^\top\lambda(X^j).\end{aligned}$$

Thus, like before, $\mathrm{Tr}(X^iX^i) \le \langle U \operatorname{diag}\lambda(X^i)U^\top, X^i\rangle$, implying $X^i = U \operatorname{diag}\lambda(X^i)U^\top$. Since i was arbitrary, this implies the existence of a simultaneous ordered spectral decomposition for $\{X^1, \ldots, X^m\}$, i.e. for every $i \in [m]$, $X^i = U \operatorname{diag}\lambda(X^i)U^\top$. □

17 ** (Singular values and von Neumann's lemma). Let $\mathbb{M}^n$ denote the vector space of $n \times n$ real matrices. For a matrix $A \in \mathbb{M}^n$ we define the *singular values* of A by $\sigma_i(A) = \sqrt{\lambda_i(A^\top A)}$ for $i = 1, 2, \ldots, n$, and hence define a map $\sigma : \mathbb{M}^n \to \mathbb{R}^n$. (Notice zero may be a singular value.)

(a) Prove

$$\lambda\left(\begin{bmatrix} 0 & A^\top \\ A & 0 \end{bmatrix}\right) = \begin{bmatrix} \sigma(A) \\ [-\sigma(A)] \end{bmatrix}.$$

Proof. Suppose v is an eigenvector of $A^\top A$ with eigenvalue λ. Then, it is an eigenvector of $\sqrt{A^\top A}$ with eigenvalue $\sigma = \sqrt{\lambda}$. Let $\operatorname{sgn} \in \{-1, 1\}$. Then,

$$\begin{bmatrix} 0 & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} \sigma v \\ \operatorname{sgn} Av \end{bmatrix} = \begin{bmatrix} \operatorname{sgn} A^\top Av \\ \sigma Av \end{bmatrix} = \operatorname{sgn} \sigma \begin{bmatrix} \sigma v \\ \operatorname{sgn} Av \end{bmatrix}.$$

In other words, $[\sigma v \ \ \operatorname{sgn} Av]$ is an eigenvector of the matrix with eigenvalue $\operatorname{sgn} \sigma$, if $\sigma \neq 0$. If $\sigma = 0$, then $v \in \operatorname{null}(A)$. Because $\operatorname{null}(A) = \operatorname{range}(A^\top)^\perp$, we have by rank-nullity $\dim \operatorname{null}(A) = n - \dim \operatorname{range}(A^\top) = \dim \operatorname{null}(A^\top)$. Therefore, we can pick an element u of $\operatorname{null}(A^\top)$ (orthogonal to previous choices) to obtain the eigenvector $[v \ \ \operatorname{sgn} u]$ with eigenvalue $\sigma = 0$.
Given n orthogonal eigenvectors of $A^\top A$ with eigenvalues σ_i^2, we have exhibited $2n$ orthogonal eigenvectors of the above matrix with eigenvalues $\pm\sigma_i$. This proves the desired statement. □

(b) For any other matrix $B \in \mathbb{M}^n$, use part (a) and Fan's inequality to prove

$$\operatorname{Tr}(A^\top B) \leq \sigma(A)^\top \sigma(B).$$

Proof. We apply Fan's inequality to the matrices below.

$$2\operatorname{Tr}(A^\top B) = \operatorname{Tr}\left(\begin{bmatrix} 0 & A^\top \\ A & 0 \end{bmatrix}\begin{bmatrix} 0 & B^\top \\ B & 0 \end{bmatrix}\right) \leq \begin{bmatrix} \sigma(A) \\ [-\sigma(A)] \end{bmatrix}^\top \begin{bmatrix} \sigma(B) \\ [-\sigma(B)] \end{bmatrix} = 2\sigma(A)^\top \sigma(B).$$

□

(c) If A lies in $\mathbb{S}^n_+$, prove $\lambda(A) = \sigma(A)$.

Proof. Since $A^\top A = A^2$, we have $\lambda(A^\top A) = \lambda(A)^2$ where the square is entrywise. Furthermore, $\lambda(A) \geq 0$. Thus, $\lambda(A) = \sqrt{\lambda(A^\top A)} = \sigma(A)$ where the $\sqrt{\cdot}$ is entrywise. □

(d) By considering matrices of the form $A + \alpha I$ and $B + \beta I$, deduce Fan's inequality from von Neumann's lemma (part (b)).

Proof. By taking $\alpha, \beta > 0$ large enough, $A+\alpha I$ and $B+\beta I$ are PSD. Thus, $\lambda(A+\alpha I) = \sigma(A+\alpha I)$ and the same for $B+\beta I$. Then, part (b) says

$$\begin{aligned}
\operatorname{Tr} AB + \beta \operatorname{Tr} A + \alpha \operatorname{Tr} B + \alpha\beta n &= \operatorname{Tr}\left[(A+\alpha I)(B+\beta I)\right] \\
&\leq \sigma(A+\alpha I)^\top \sigma(B+\beta I) = \lambda(A+\alpha I)^\top \lambda(B+\beta I) \\
&= (\lambda(A) + \alpha \mathbf{1})^\top (\lambda(B) + \beta \mathbf{1}) \\
&= \lambda(A)^\top \lambda(B) + \beta \operatorname{Tr} A + \alpha \operatorname{Tr} B + \alpha\beta n,
\end{aligned}$$

from which Fan's inequality follows by subtracting quantities that are on both sides. $\square$

2 Chapter 2: Inequality constraints

2.1 2.1 Optimality Conditions

1. The *directional derivative* of a function f at $\bar{x}$ in a direction $d \in \mathbb{E}$ is

$$f'(\bar{x}; d) = \lim_{t \downarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t},$$

   when this limit exists.

2. If exists $a \in \mathbb{E}$ where $f'(\bar{x}; d) = \langle a, d \rangle$ for all $d \in \mathbb{E}$, then we say f is (Gateaux) differentiable at $\bar{x}$ with (Gateaux) derivative $\nabla f(\bar{x}) = a$.

3. The *normal cone* to C at $\bar{x} \in C$ is the convex cone $\{d \in \mathbb{E} : \langle d, x - \bar{x} \rangle \leq 0\}$.

Proposition 2.1 (2.1.1 (First order necessary condition))**.** *Suppose C is a convex set in $\mathbb{E}$ and $\bar{x}$ is a local minimizer of function $f : C \to \mathbb{R}$. Then for any point $x \in C$, the directional derivative, if it exists, satisfies $f'(\bar{x}; x - \bar{x}) \geq 0$. In particular, if f is differentiable at $\bar{x}$, then $-\nabla f(\bar{x}) \in N_C(\bar{x})$.*

Proposition 2.2 ((First order sufficient condition))**.** *Suppose that the set $C \subset \mathbb{E}$ is convex and that the function $f : C \to \mathbb{R}$ is convex. Then for any $\bar{x}, x \in C$, $f'(\bar{x}; x - \bar{x})$ exists in $[-\infty, +\infty)$. If $f'(\bar{x}; x - \bar{x}) \geq 0$ holds for all $x \in C$, or $-\nabla f(\bar{x}) \in N_C(\bar{x})$, then $\bar{x}$ is a global minimizer of f on C.*

Corollary 2.3 (2.1.3., (First order conditions for linear constraints))**.** *For a convex set $C \subset \mathbb{E}$, a function $f : C \to \mathbb{R}$, a linear map $A : \mathbb{E} \to \mathbb{Y}$ (where $\mathbb{Y}$ is a Euclidean space) and a point b in $\mathbb{Y}$, consider the optimization problem*

$$\inf\{f(x) \mid x \in C, Ax = b\}.$$

Suppose $\bar{x} \in \operatorname{int} C$ satisfies $A\bar{x} = b$.

(a) If $\bar{x}$ is a local minimizer for the problem and f is differentiable at $\bar{x}$ then $\nabla f(\bar{x}) \in A^*\mathbb{Y}$.

(b) Conversely, if $\nabla f(\bar{x}) \in A^*\mathbb{Y}$ and f is convex then $\bar{x}$ is a global minimizer for the problem.

Theorem 2.4 (2.1.5 (Second order conditions))**.** *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable and has a critical point $\bar{x}$. If $\bar{x}$ is a local minimizer then the Hessian $\nabla^2 f(\bar{x})$ is PSD. Conversely, if the Hessian is PD then $\bar{x}$ is a local minimizer.*

Theorem 2.5 (2.1.6 (Basic separation))**.** *If $C \in \mathbb{E}$ is convex and closed and $y \notin C$, then there exists $a \in \mathbb{E}$ and $b \in \mathbb{R}$ such that for all $x \in C$, $\langle a, y \rangle > b \geq \langle a, x \rangle$.*

Proposition 2.6 (2.1.7)**.** *If $f : \mathbb{E} \to \mathbb{R}$ is differentiable and bounded below then there are points where f has arbitrarily small derivative.*

2.2 Exercises for 2.1

1. Prove the normal cone is a closed convex cone.

Proof. If $\langle d, x - \bar{x} \rangle \leq 0$ for all $x \in C$, then $\langle rd, x - \bar{x} \rangle \leq 0$ for all $x \in C$ too, for any $r \in \mathbb{R}_+$. Thus $N_C(\bar{x})$ is a cone. Furthermore, if $\langle d_i, x - \bar{x} \rangle \leq 0$ for $i \in \{1, 2\}$, then for any $\lambda \in [0, 1]$,

$$\langle \lambda d_1 + (1 - \lambda) d_2, x - \bar{x} \rangle = \lambda \langle d_1, x - \bar{x} \rangle + (1 - \lambda) \langle d_2, x - \bar{x} \rangle \leq 0.$$

Therefore $N_C(\bar{x})$ is convex. Now let $d^i \to d$ be a convergent sequence in $N_C(\bar{x})$ and let $x \in C$ be arbitrary.

$$\langle d^i, x - \bar{x} \rangle \to \langle d, x - \bar{x} \rangle,$$

and since each element of the sequence is ≤ 0, the limit, $\langle d, x - \bar{x} \rangle \leq 0$. Since x was arbitrary, $d \in N_C(\bar{x})$. Therefore, $N_C(\bar{x})$ is closed. $\square$

2 (Examples of normal cones). For the following sets $C \subset \mathbb{E}$, check C is convex and compute the normal cone $N_C(\bar{x})$ for points $\bar{x}$ in C:

(a) C a closed interval in $\mathbb{R}$.

$$a \leq x, y \leq b \implies a \leq \lambda x + (1 - \lambda) y \leq b.$$

If $\bar{x}$ is in the interior of C, then if d is nonzero, then $d \notin N_C(\bar{x})$ because $\bar{x} + \epsilon d \in C$ for some $\epsilon > 0$. Thus, only the normal cones $N_C(a)$ and $N_C(b)$ (if the interval is $[a, b]$) are nontrivial. By inspection, $N_C(a) = -\mathbb{R}_+$ and $N_C(b) = \mathbb{R}_+$.

(b) $C = B$, the unit ball.

$$\|x\|, \|y\| \leq 1 \implies \|\lambda x + (1 - \lambda) y\| \leq \lambda \|x\| + (1 - \lambda) \|y\| \leq 1.$$

As proved in (a), for $\|\bar{x}\| < 1$, $N_C(\bar{x}) = \{0\}$. Now suppose $\|\bar{x}\| = 1$. We have $N_C(\bar{x}) = \mathbb{R}_+ \bar{x}$. If $r \geq 0$ then for $y \in B$,

$$\langle r\bar{x}, y - \bar{x} \rangle = r \langle \bar{x}, y \rangle - r \leq r - r = 0$$

by Cauchy-Schwarz and $\|y\| \leq 1$. Now if $d \notin \mathbb{R}_+ \bar{x}$, then $\langle d, d\|d\|^{-1} - \bar{x} \rangle > \|d\| - \|d\| \|\bar{x}\| = 0$.

(c) C a subspace.
Observe that for any set $C \in \mathbb{E}$, $N_C(\bar{x}) = N_{C - \bar{x}}(0)$. If C is a subspace, then $C - \bar{x} = C$. Therefore, the normal cone is equal at any point of the subspace. Now let $0 \in C$ and $d \in N_C(0)$. If $\exists x \in C, \langle d, x \rangle < 0$, then $-x \in C, d \in N_C(0) \implies \langle d, -x \rangle \leq 0$, a contradiction. Therefore, $N_C(0) = \{d \in \mathbb{E} : \langle d, x \rangle = 0 \; \forall x \in C\} = C^\perp$, the orthogonal subspace to C. By our earlier comment about the normal cones being equal, this is true of $N_C(\bar{x})$ for any $\bar{x} \in C$.

(d) C a closed halfspace: $\{x \mid \langle a, x \rangle \leq b\}$ where $0 \neq a \in \mathbb{E}$ and $b \in \mathbb{R}$.
Noninterior points consist of the points $\delta C = \{x : \langle a, x \rangle = b\}$. Let $\bar{x} \in \delta C$. We have $N_C(\bar{x}) = \mathbb{R}_+ a$. By definition of C, $\langle a, x - \bar{x} \rangle = \langle a, x \rangle - b \leq 0$, so $r \langle a, x - \bar{x} \rangle \leq 0$ for every $r \in \mathbb{R}_+$. Furthermore, if z has a component orthogonal to a, we have $\langle z, x - \bar{x} \rangle > 0$ by traveling in the direction of the orthogonal component on the hyperplane δC.

(e) $C = \{x \in \mathbb{R}^n \mid x_j \geq 0 \text{ for all } j \in J\}$ (for $J \subset \{1, 2, \ldots, n\}$).
Noninterior points are those for which some constraint is tight. Let $T \subset J$ be the set of indices of tight constraints for the point $\bar{x}$. We have $N_C(\bar{x}) = \{d \in \mathbb{E} : d_j \leq 0 \forall j \in T, d_j = 0 \forall j \notin T\}$. To see this, take $x \in C$. For each $j \in T$, since $\bar{x}_j = 0$, $(x - \bar{x})_j \geq 0$. This implies $\langle d, x - \bar{x} \rangle \leq 0$ for every $d \in N_C(\bar{x})$. If $d \in \mathbb{E}$ is such that $d_j > 0$ for some $j \in T$, then we can take $x = \bar{x} + \epsilon e_j$ for any $\epsilon > 0$ to show d cannot be in the normal cone. If $d_j \neq 0$ for some $j \notin T$, Then we can take $x = \bar{x} + \operatorname{sgn}(d_j)\epsilon e_j$ for small enough $\epsilon > 0$ for x to be contained in C (if $j \in J$, we are relying on $\bar{x}_j > 0$ not being tight).

3 (Self-dual cones). Prove each of the following cones K satisfy the relationship $N_K(0) = -K$.
The dual cone to C is defined as $C^* = \{d \in \mathbb{E} : \langle d, x \rangle \geq 0 \, \forall x \in C\}$. We see that $C^* = -N_C(0)$. $N_C(0)$ is the set of $d \in \mathbb{E}$ where $\langle d, x \rangle \leq 0$ for all $x \in C$, which is precisely the negated elements $d \in \mathbb{E}$ for which $\langle d, x \rangle \geq 0$.

(a) $\mathbb{R}^n_+$.
Let us compute $(\mathbb{R}^n_+)^*$. Suppose $d \in (\mathbb{R}^n_+)^*$. Then, it must have nonnegative coordinates; otherwise if $d_j < 0$, then $e_j \in \mathbb{R}^n_+$ would give $\langle d, e_j \rangle = d_j < 0$. Moreover, if $d \in \mathbb{E}$ has all nonnegative coordinates, then $\langle d, x \rangle$ for $x \in \mathbb{R}^n_+$ is a sum of nonnegative numbers, which is nonnegative. Therefore, $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$.

(b) $\mathbb{S}^n_+$.
Suppose $A \in (\mathbb{S}^n_+)^*$. Then, A must be PSD. Otherwise, it has an eigenvector u with negative eigenvalue $\lambda < 0$, and $uu^\top \in \mathbb{S}^n_+$ but $\langle A, uu^\top \rangle = u^\top A u = \lambda \|u\|^2 < 0$. Now suppose $A \in \mathbb{S}_n$ is PSD. Then, $\langle A, X \rangle \geq]\lambda(A)[^\top \lambda(X) \geq 0$ (Lower bound implied by Fan's inequality,][denotes rearrangement of a vector into nondecreasing order) for every $X \in \mathbb{S}^n_+$. Therefore, $(\mathbb{S}^n_+)^* = \mathbb{S}^n_+$. Note the first part of this proof, i.e. $(\mathbb{S}^n_+)^*$ only contains PSD matrices, would not be possible without the assumption A is symmetric (i.e. restricting to the universe of symmetric matrices). Note that

$$x^\top \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} x = x_1^2 + x_2^2 \geq 0$$

for all $x \in \mathbb{R}^n$, but the matrix is not PSD.

(c) $C = \{x \in \mathbb{R}^n \mid x_1 \geq 0, x_1^2 \geq x_2^2 + \ldots + x_n^2\}$.
Suppose $d \in C^*$. Clearly $d_1 \geq 0$, because $\mathbb{R}_+ e_1 \in C$. Now suppose $d_1^2 < \sum_{j=2}^n d_j^2$. Consider the element $x \in \mathbb{E}$ defined by $x_1 = d_1$, and $x_i = -d_i |d_1| / \sqrt{\sum_{j=2}^n d_j^2}$ for $i = 2, \ldots, n$. Then

$$\sum_{j=2}^n x_j^2 = \frac{d_1^2}{\sum_{j=2}^n d_j^2} \sum_{j=2}^n d_j^2 = d_1^2 = x_1^2 \implies x \in C.$$

Now we compute

$$\langle d, x \rangle = x_1 d_1 + \sum_{j=2}^n x_j d_j = d_1^2 - \frac{|d_1|}{\sqrt{\sum_{j=2}^n d_j^2}} \sum_{j=2}^n d_j^2 = d_1^2 - |d_1| \sqrt{\sum_{j=2}^n d_j^2} < 0,$$

because $|d_1| < \sqrt{\sum_{j=2}^{n} d_j^2}$. Therefore, $d \notin C^*$. Now suppose $d \in C$. For any $x \in \mathbb{E}$, denote $x_{2:}$ as the vector in $\mathbb{R}^{n-1}$ of the 2nd through nth entries of x. Let $x \in C$. Then

$$\langle d, x\rangle \geq d_1 x_1 - \|d_{2:}\|\|x_{2:}\| \geq 0,$$

because $d_1 \geq \|d_{2:}\|$ and $x_1 \geq \|x_{2:}\|$.

4 (Normals to affine sets). Given a linear map $A : \mathbb{E} \to \mathbb{Y}$ (where $\mathbb{Y}$ is a Euclidean space) and a point $b \in \mathbb{Y}$, prove the normal cone to the set $\{x \in \mathbb{E} \mid Ax = b\}$ at any point in it is $A^*\mathbb{Y}$. Hence deduce Corollary 2.1.3.

Proof. Denote $C = \{x \in \mathbb{E} \mid Ax = b\}$. Given $\bar{x} \in C$, we have $N_C(\bar{x}) = N_{C-\bar{x}}(0)$. $C - \bar{x}$ is the subspace $\{x \in \mathbb{E} \mid Ax = 0\}$. By problem 2, part (c), $N_{C-\bar{x}}(0)$ is the perpendicular subspace to $\{x \in \mathbb{E} \mid Ax = 0\} = \mathrm{null}(A)$. By the fundamental theorem of linear algebra, $\mathrm{null}(A)^\perp = \mathrm{range}(A^*)$. Therefore, $N_C(\bar{x}) = \mathrm{range}(A^*) = A^*\mathbb{Y}$.
Corollary 2.1.3 states that a differentiable f achieving a local minimum at $\bar{x} \in \operatorname{int} C$ subject to $Ax = b$ implies that $-\nabla f(\bar{x}) \in A^*\mathbb{Y}$, and that this is a sufficient condition for convex functions.
By Proposition 2.1.1, $-\nabla f(\bar{x}) \in N_{C'}(\bar{x})$, where $C' = C \cap \{x \in \mathbb{E} : Ax = b\}$. Since $x \in \operatorname{int} C$, one can move in any direction from x a small amount and stay in C. Therefore, for each direction that an element of $\{x \in \mathbb{E} : Ax = b\}$ is away from $\bar{x}$, $C \cap \{x \in \mathbb{E} : Ax = b\}$ contains an element that direction away from $\bar{x}$. Thus, $N_{C'}(\bar{x}) \subset N_{\{x:Ax=b\}}(\bar{x}) = A^*\mathbb{Y}$. □

5. Prove that the differentiable function $x_1^2 + x_2^2(1 - x_1)^3$ has a unique critical point in $\mathbb{R}^2$, which is a local minimizer, but has no global minimizer. Can this happen on $\mathbb{R}$?

Proof. The derivative of this function is

$$\nabla f(x) = \begin{bmatrix} 2x_1 - 3x_2^2(1 - x_1)^2 \\ 2x_2(1 - x_1)^3. \end{bmatrix}$$

If x is a critical point, then $\nabla f(x) = 0$. This means either $x_2 = 0$ or $x_1 = 1$. If $x_1 = 1$, then the first coordinate equals 2, which is a contradiction. Thus, $x_2 = 0$, which implies that $x_1 = 0$. Thus, $x = 0$ is the only critical point.
To see it is a local minimizer, note that in a neighborhood of 0, $|x_1| < 1$, which implies the second term is nonnegative, and the first term is always nonnegative. And $f(0) = 0$.
This cannot happen on $\mathbb{R}$. If x is a local minimizer and the only critical point and f is differentiable, suppose x is not a global minimum, i.e. there exists y where $f(y) < f(x)$. This implies there exists z between y and x where $f(z) = f(x)$ by the intermediate value theorem. Then by the mean value theorem, there exists w between z and x where $f'(w) = 0$. This is a contradiction, since we assumed x was the only critical point. □

6 (The Rayleigh quotient).

(a) Let the function $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ be continuous, satisfying $f(\lambda x) = f(x)$ for all $\lambda > 0$ in $\mathbb{R}$ and nonzero $x \in \mathbb{R}^n$. Prove f has a minimizer.
The minimum of f over the unit circle $\{x \in \mathbb{R}^n : \|x\| = 1\}$, which exists by compactness of the unit circle and continuity of f, must be the minimum of f. To see why, if x^* is the minimizer over the unit circle, given $x \in \mathbb{R}^n \setminus \{0\}$, we have $f(x) = f(x/\|x\|) \geq f(x^*)$.

(b) Given a matrix $A \in \mathbb{S}^n$, define a function $g(x) = x^\top Ax/\|x\|^2$ for nonzero $x \in \mathbb{R}^n$. Prove g has a minimizer.

First of all, g is continuous. Also, it satisfies $f(x) = f(\lambda x)$ for every $x \in \mathbb{R}^n \backslash \{0\}, \lambda \in \mathbb{R}$:

$$\frac{(\lambda x)^\top A(\lambda x)}{\|\lambda x\|^2} = \frac{\lambda^2 x^\top Ax}{\lambda^2\|x\|^2} = \frac{x^\top Ax}{\|x\|^2}.$$

Then by part (a), g has a minimizer.

(c) Calculate $\nabla g(x)$ for nonzero x.

We compute the directional derivative directly. Let $d \in \mathbb{E}, t > 0$.

$$\begin{aligned} g(x+td) - g(x) &= \frac{(x+td)^\top A(x+td)}{\|x+td\|^2} - \frac{x^\top Ax}{\|x\|^2} = \frac{x^\top Ax + 2td^\top Ax + t^2\|d\|^2}{\|x+td\|^2} - \frac{x^\top Ax}{\|x\|^2} \\ &= \frac{\|x\|^2 x^\top Ax + 2t\|x\|^2 d^\top Ax + t^2\|x\|^2\|d\|^2}{\|x\|^2\|x+td\|^2} - \frac{x^\top Ax\|x+td\|^2}{\|x\|^2\|x+td\|^2} \\ &= \frac{\|x\|^2 x^\top Ax + 2t\|x\|^2 d^\top Ax + t^2\|x\|^2\|d\|^2}{\|x\|^2\|x+td\|^2} - \frac{x^\top Ax(\|x\|^2 + 2td^\top x + t^2\|d\|^2)}{\|x\|^2\|x+td\|^2} \\ &= \frac{2td^\top Ax}{\|x+td\|^2} - \frac{2tx^\top Axd^\top x}{\|x\|^2\|x+td\|^2} + \frac{\overbrace{t^2\|x\|^2\|d\|^2 - t^2\|d\|^2(x^\top Ax)}^{(\cdot)/t \xrightarrow{t\downarrow 0} 0}}{\|x\|^2\|x+td\|^2}. \end{aligned}$$

This implies that the directional derivative is

$$g'(x; d) = \frac{2d^\top Ax}{\|x\|^2} - \frac{2(x^\top Ax)d^\top x}{\|x\|^4}.$$

This implies that the gradient is $\nabla g(x) = \frac{2Ax}{\|x\|^2} - \frac{2x^\top Axx}{\|x\|^4}$.

(d) Deduce that minimizers of g must be eigenvectors, and calculate the minimum value.

By the first order necessary condition for local minima, for a minimizer x we have $\nabla g(x) = 0$, in other words

$$\frac{2Ax}{\|x\|^2} = \frac{2x^\top Axx}{\|x\|^4} \implies Ax = \frac{x^\top Ax}{\|x\|^2}x,$$

i.e. x has eigenvalue $x^\top Ax/\|x\|^2$. Thus, the minimizer is the eigenvector with the minimum eigenvalue of A.

(e) Find an alternative proof of part (d) by using a spectral decomposition of A.

Let $A = \sum_{i=1}^n \lambda_i u_i u_i^\top$ with $\{u_1, \ldots, u_n\}$ an orthonormal basis. Any vector x can be expressed as $x = \sum_{i=1}^n c_i u_i, c_i \in \mathbb{R}$. If $\|x\| = 1$ then by orthogonality $\sum_{i=1}^n c_i^2 = 1$. We have

$$\begin{aligned} g(x) = x^\top Ax &= \left(\sum_{i=1}^n c_i u_i\right)^\top \left(\sum_{i=1}^n \lambda_i u_i u_i^\top\right)\left(\sum_{i=1}^n c_i u_i\right) = \left(\sum_{i=1}^n c_i u_i\right)\sum_{i=1}^n \lambda_i c_i u_i \\ &= \sum_{i=1}^n \lambda_i c_i^2 \geq \sum_{i=1}^n \lambda_{\min} c_i^2 = \lambda_{\min}. \end{aligned}$$

This implies that the minimum value is $\lambda_{\min}$, the minimum eigenvalue, and is only achieved when all the weight of c is on the i for which $\lambda_{\min} = \lambda_i$.

7. Suppose a convex function $g : [0,1] \to \mathbb{R}$ satisfies $g(0) = 0$. Prove the function $t \in (0,1] \mapsto g(t)/t$ is nondecreasing. Hence prove that for a convex function $f : C \to \mathbb{R}$ and points $\bar{x}, x \in C \subset \mathbb{E}$, the quotient $(f(\bar{x} + t(x - \bar{x})) - f(\bar{x}))/t$ is nondecreasing as a function of $t \in (0,1]$, and complete the proof of Proposition 2.1.2.

Proof. Suppose $g(0) = 0$ and is convex. Let $0 < t_1 < t_2$. By convexity,

$$g(t_1) \le \frac{t_1}{t_2} g(t_2) + \left(1 - \frac{t_1}{t_2}\right) g(0) = \frac{t_1}{t_2} g(t_2).$$

By dividing both sides by t_1, we get $g(t_1)/t_1 \le g(t_2)/t_2$, i.e. the function $g(t)/t$ is nondecreasing.

Now given a convex $f : C \to \mathbb{R}$ and points $\bar{x}, x \in C$, let us examine the function $(0,1] \to \mathbb{R}$, $g := t \mapsto (f(\bar{x} + t(x - \bar{x})) - f(\bar{x}))/t$. First of all, $f(\bar{x} + t(x - \bar{x})) - f(\bar{x})$ as a function of t is 0 at 0 and convex, because it is a composition of a convex function, f, with an affine function, $\bar{x} + t(x - \bar{x})$. Therefore, $g(t)$ is nondecreasing in t. Now let $s = \inf\{t \in (0,1] : g(t)\}$ (if the set is unbounded below let $s = -\infty$). If $s \in (-\infty, \infty)$, then consider $\epsilon > 0$. There must exist $t_0 \in (0,1]$ such that $g(t_0) < s + \epsilon$. Because g is nondecreasing, we have $t \in (0, t_0] \implies g(t) < s + \epsilon$. We have shown that for any ϵ, if t is close enough to 0 then $g(t)$ is ϵ close to s. This proves that $\lim_{t \downarrow 0} g(t) = s$. If $s = -\infty$, replace ϵ close to s with less than an arbitrary $M \in \mathbb{R}$. We have shown that $f'(\bar{x}; x - \bar{x})$ exists and is in $[-\infty, \infty)$.

Now to complete the proof of Proposition 2.1.2, we show that $f'(\bar{x}; x - \bar{x}) \ge 0$ for all $x \in C$ implies that $\bar{x}$ is a global minimizer. Let $x \in C$ be arbitrary.

$$\lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \ge 0, \quad \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \text{ nondecreasing for } t \in (0,1]$$

implies, by setting $t = 1$, that $f(x) - f(\bar{x}) \ge 0$. This proves that $\bar{x}$ is a global minimizer of f on C. □

8 * (Nearest points).

(a) Prove that if $f : C \to \mathbb{R}$ is strictly convex then it has at most one global minimizer on C.

Suppose x and y are both global minimizers. Then

$$f\left(\frac{x+y}{2}\right) < \frac{1}{2} f(x) + \frac{1}{2} f(y) = f(x).$$

This contradicts the fact that x and y are minimizers.

(b) Prove $f(x) = \|x - y\|^2/2$ is strictly convex on $\mathbb{E}$ for any point $y \in \mathbb{E}$.

We prove the function $\|x\|^2$ is strictly convex. Let $x, y \in \mathbb{E}$ and $\lambda \in (0,1)$.

$$\begin{aligned}
\|\lambda x + (1-\lambda) z\|^2 &\le \lambda^2 \|x\|^2 + 2\lambda(1-\lambda)\langle x, z\rangle + (1-\lambda)^2 \|z\|^2 \\
&\le \lambda^2 \|x\|^2 + 2\lambda(1-\lambda)\|x\|\|z\| + (1-\lambda)^2 \|z\|^2 \\
&\le \lambda^2 \|x\|^2 + \lambda(1-\lambda)\left(|x\|^2 + \|z\|^2\right) + (1-\lambda)^2 \|z\|^2 = \lambda \|x\|^2 + (1-\lambda)\|z\|^2.
\end{aligned}$$

The third inequality is strict unless $\|x\| = \|z\|$, and the second inequality is strict unless z is a multiple of x. Therefore, we have strict inequality unless $x = z$. This shows strict convexity of $\|x\|^2$. To conclude for $\|x-y\|^2/2$, note that for $\lambda \in (0,1)$,

$$\|\lambda x + (1-\lambda)z - y\|^2 = \|\lambda(x-y) + (1-\lambda)(z-y)\|^2 \leq \lambda\|x-y\| + (1-\lambda)\|z-y\|$$

with strict inequality unless $x - y = z - y$, i.e. $x = z$.

(c) Suppose C is a nonempty, closed convex subset of $\mathbb{E}$.

(i) If y is any point in $\mathbb{E}$, prove there is a unique nearest point $P_C(y)$ to y in C, characterized by

$$\langle y - P_C(y), x - P_C(y)\rangle \leq 0 \text{ for all } x \in C.$$

By definition, $P_C(y) = \min_{x\in C} \|x-y\|^2/2$. A minimizer exists, because we can always consider a ball around y sufficiently large to include some point of C intersected with C, which is compact. Furthermore, it is unique by part (a).
The gradient of $\|x-y\|^2/2$ is $x - y$. We know that $P_C(y)$ is a global minimizer of this function over C, and thus by Proposition 2.1.1 satisfies the first order condition. Furthermore, by Proposition 2.1.2, any element satisfying the first order condition is a global minimizer. Since the minimizer is unique, it is precisely characterized by the first order condition:

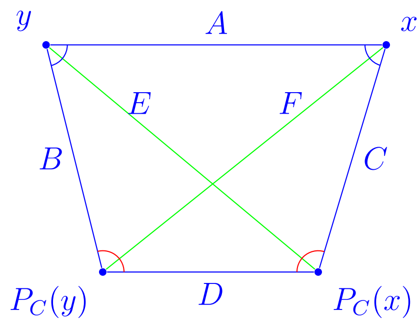
$$\langle P_C(y) - y, x - P_C(y)\rangle \geq 0 \; \forall x \in C.$$

(ii) For any $\bar{x} \in C$, deduce that $d \in N_C(\bar{x})$ holds if and only if $\bar{x}$ is the nearest point in C to $\bar{x} + d$.
$d \in N_C(\bar{x})$ holds iff $\langle d, x - \bar{x}\rangle \leq 0$ for every $x \in C$. In other words, $\langle \bar{x} + d - \bar{x}, x - \bar{x}\rangle \leq 0$ for every $x \in C$, i.e. $\bar{x}$ satisfies the nearest point condition to $y = \bar{x} + d$.

(iii) Deduce, furthermore, that any points y and z in $\mathbb{E}$ satisfy

$$\|P_C(y) - P_C(z)\| \leq \|y - z\|,$$

so in particular the *projection* $P_C : \mathbb{E} \to C$ is continuous.
The characterization of the nearest point says that the line segment from y to $P(y)$ and the line segment from $P(y)$ to x for any $x \in C$ forms an obtuse angle. Because $P(x), P(y) \in C$, it suffices to prove essentially that for any quadrilateral (multidimensional, so all four points may not be in the same plane), if a side forms obtuse angles with its neighboring sides then it is shorter than the side across it. One can loosely relate this to triangles; the side across an obtuse angle of a triangle is the longest. For quadrilaterals, the side across a pair of obtuse angles is longer than the side between the obtuse angles. Here is a picture for clarity.



Step 1: Show that $E^2 \geq B^2 + D^2$ and $F^2 \geq D^2 + C^2$, because the red angles are obtuse.
Step 2: Show that $A^2+B^2+A^2+C^2 \geq E^2+F^2$ because the red angles are obtuse. (Perhaps this can be interpreted as being because the sum of the blue angles is less than 180).
Step 3: Then add the equations to get $D \leq A$ as desired.

Recall that the nearest point characterization implies

$$\langle y - P_C(y), P_C(x) - P_C(y)\rangle \leq 0 \tag{2.2.1}$$

$$\langle x - P_C(x), P_C(y) - P_C(x)\rangle \leq 0 \tag{2.2.2}$$

We perform step 1 (for brevity $P := P_C$):

$$\begin{aligned} E^2 &= \|y - P(x)\|^2 = \|y - P(y) - (P(x) - P(y))\|^2 \\ &= \|y - P(y)\|^2 - 2\langle y - P(y), P(x) - P(y)\rangle + \|P(x) - P(y)\|^2 \\ &\overset{(2.2.1)}{\overbrace{\geq}} \|y - P(y)\|^2 + \|P(x) - P(y)\|^2 = B^2 + D^2. \end{aligned}$$

The proof for $F^2 \geq C^2 + D^2$ using (2.2.2) is essentially the same.
Now we perform step 2.

$$\begin{aligned} E^2 + F^2 &= \|y - P(x)\|^2 + \|x - P(y)\|^2 = \|y - x - (P(x) - x)\|^2 + \|x - y - (P(y) - y)\|^2 \\ &= 2\|x - y\|^2 + \|P(x) - x\|^2 + \|P(y) - y\|^2 - 2\langle y - x, P(x) - x\rangle - 2\langle x - y, P(y) - y\rangle \\ &= 2\|x - y\|^2 + \|P(x) - x\|^2 + \|P(y) - y\|^2 + 2\langle x - y, P(x) - P(y) - (x - y)\rangle. \end{aligned}$$

Now by adding (2.2.1) and (2.2.2), we obtain

$$\langle P(x) - P(y) - (x - y), P(x) - P(y)\rangle \leq 0 \implies -\langle P(x) - P(y), P(x) - P(y) - (x - y)\rangle \geq 0.$$

Thus, adding this to the end of the previous derivation,

$$\begin{aligned} E^2 + F^2 &\leq 2\|x - y\|^2 + \|P(x) - x\|^2 + \|P(y) - y\|^2 - \|P(x) - P(y) - (x - y\|^2 \\ &\leq 2\|x - y\|^2 + \|P(x) - x\|^2 + \|P(y) - y\|^2 = 2A^2 + B^2 + C^2. \end{aligned}$$

Now for step three:

$$\begin{gathered} E^2 \geq B^2 + D^2, F^2 \geq C^2 + D^2, \quad 2A^2 + B^2 + C^2 \geq E^2 + F^2 \\ \implies 2A^2 + B^2 + C^2 \geq B^2 + D^2 + C^2 + D^2 \implies 2A^2 \geq 2D^2 \\ \implies A \geq D, \text{ i.e. } \|x - y\| \geq \|P(x) - P(y)\|. \end{gathered}$$

(d) Given a nonzero element a of $\mathbb{E}$, calculate the nearest point in the subspace $\{x \in \mathbb{E} \mid \langle a, x\rangle = 0\}$ to the point $y \in \mathbb{E}$.
Since $\mathbb{E} = \operatorname{Span}\{a\} \oplus \operatorname{Span}\{a\}^\perp$, $y = ca + a_\perp$ for some $a_\perp \in \operatorname{Span}\{a\}^\perp$ and $c \in \mathbb{R}$. Further, $\{x \in \mathbb{E} \mid \langle a, x\rangle = 0\} = \operatorname{Span}\{a\}^\perp$. $\langle y - a_\perp, a'_\perp - a_\perp\rangle = 0$ for every $a'_\perp \in \operatorname{Span}\{a\}^\perp$, so $a_\perp$ is the nearest point in $\{x \in \mathbb{E} : \langle a, x\rangle = 0\}$ to y. A way to see this without the nearest point characterization is, $\|y - a'_\perp\|^2 = c^2\|a\|^2 + \|a_\perp - a'_\perp\|^2 \geq c^2\|a\|^2 = \|y - a_\perp\|^2$. To compute $a_\perp$, realize that

$$\langle a, y\rangle = c\|a\|^2 \implies c = \frac{\langle a, y\rangle}{\|a\|^2}.$$

Therefore,

$$a_\perp = y - ca = y - \frac{\langle a, y\rangle}{\|a\|^2}a.$$

(e) **(Projection on $\mathbb{R}^n_+$ and $\mathbb{S}^n_+$).** Prove the nearest point in $\mathbb{R}^n_+$ to a vector $y \in \mathbb{R}^n$ is y^+, where $y_i^+ = \max\{y_i, 0\}$ for each i. For a matrix $U \in \mathbb{O}^n$ and $y \in \mathbb{R}^n$, prove that the nearest PSD matrix to $U \operatorname{diag} y U^\top$ is $U \operatorname{diag} y^+ U^\top$.
Let $x \in \mathbb{R}^n_+$. Suppose $I = \{i : y_i \leq 0\}$.

$$\|y - x\|^2 = \sum_{i=1}^n (y_i - x_i)^2 \geq \sum_{i \in I} (y_i - x_i)^2 \geq \sum_{i \in I} y_i^2,$$

because $y_i \leq 0$ and $x_i \geq 0$, so $|y_i - x_i| \geq |y_i|$. Now observe that $\|y - y^+\|^2 = \sum_{i\in I} y_i^2$, achieving the minimum. Therefore y^+ is the nearest point to y in $\mathbb{R}^n_+$. Now let $X \in \mathbb{S}_n^+$ with eigendecomposition $X = V \operatorname{diag} x V^\top$, with $x = \lambda(X)$. First we assume $y = [y]$, i.e. y's entries are sorted in nonincreasing order.

$$\begin{aligned}\|X - U \operatorname{diag} y U^\top\|^2 &= \|X\|^2 + \|U \operatorname{diag} y U^\top\|^2 - 2\langle X, U \operatorname{diag} y U^\top\rangle \\ &= \|x\|^2 + \|y\|^2 - 2\langle X, U \operatorname{diag} y U^\top\rangle \geq \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle = \|x - y\|^2.\end{aligned}$$

Since $x \in \mathbb{R}^n_+$, by the above part, the RHS is minimized by $x = y^+$. By Fan's Theorem, the inequality is an equality iff $X = U \operatorname{diag} x U^\top$. Therefore, if $y = [y]$, we have shown that the nearest PSD matrix is $U \operatorname{diag} y^+ U^\top$. If $y \neq [y]$, we can always rearrange the columns of U to get a matrix $U' \operatorname{diag}[y](U')^\top = U \operatorname{diag} y U^\top$. We have the nearest PSD matrix is $U' \operatorname{diag}[y^+](U')^\top$. But this equals $U \operatorname{diag} y^+ U^\top$.

9 * (Coercivity). Suppose that the function $f : \mathbb{E} \to \mathbb{R}$ is differentiable and satisfies the growth condition $\lim_{\|x\|\to\infty} f(x)/\|x\| = \infty$. Prove that the gradient map ∇f has range $\mathbb{E}$. (Hint: Minimize the function $f(\cdot) - \langle a, \cdot\rangle$ for elements a of $\mathbb{E}$.)

Proof. Let $a \in \mathbb{E}$ be arbitrary. If there exists x a minimizer of $f(\cdot) - \langle a, \cdot\rangle$, then the first order necessary condition says that $\nabla f(x) - a = 0$, i.e. $x = a$. So all we have to do is show a minimizer exists. Since

$$\lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|} - \frac{\langle a, x\rangle}{\|x\|} \geq \lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|} - \|a\| \to \infty,$$

There exists an $r \geq 1$ where $\|x\| \geq r$ implies $f(x)/\|x\| - \langle a, x\rangle/\|x\| \geq f(0) \implies f(x) - \langle a, x\rangle \geq f(0)r \geq f(0)$. Therefore, we can restrict $f(\cdot) - \langle a, \cdot\rangle$ to rB, a compact set, over which a minimizer exists. This minimizer x^* is also a global minimizer over $\mathbb{E}$ because

$$f(x^*) - \langle a, x^*\rangle \leq f(0) \leq f(x) - \langle a, x\rangle \; \forall x \in \mathbb{E}.$$

□

10.

(a) Prove $f : \mathbb{S}^n_{++} \to \mathbb{R}$ defined by $f(X) = \operatorname{Tr} X^{-1}$ is differentiable on $\mathbb{S}^n_{++}$. (Hint: Expand the expression $(X + tY)^{-1}$ as a power series.)
First observe that $(X+tY)^{-1} = X^{-1}(I+tYX^{-1})^{-1}$, since $(X+tY)X^{-1} = I+tYX^{-1}$. Now recall that if $|x| < 1$,

$$\frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n.$$

We can apply this to compute a power series for $(I + tYX^{-1})^{-1}$ assuming t is small enough, because YX^{-1} is diagonalizable: if $UDU^\top$ is an eigendecomposition of the symmetric matrix $X^{-1/2}YX^{-1/2}$ (we are using PDness of X)

$$YX^{-1} = X^{1/2}UDU^\top X^{-1/2} = MDM^{-1}$$

with $M = X^{1/2}U$. Because I is simultaneously diagonalizable with any matrix, we can apply the scalar power series to each diagonal entry of diagonalized YX^{-1} to obtain

$$(X + tY)^{-1} = X^{-1}(I + tYX^{-1}) = X^{-1}\sum_{n=0}^{\infty}(-1)^n(tYX^{-1})^n \quad \text{if } \|YX^{-1}\| < 1/t.$$

As t approaches 0, this power series becomes valid. Now we have

$$\begin{aligned}
\operatorname{Tr}(X + tY)^{-1} - \operatorname{Tr} X^{-1} &= \operatorname{Tr}\left(X^{-1} + X^{-1}\sum_{n=1}^{\infty}(-1)^n(tYX^{-1})^n\right) - \operatorname{Tr} X^{-1} \\
&= \operatorname{Tr}\left(X^{-1}\sum_{n=1}^{\infty}(-1)^n(tYX^{-1})^n\right) \\
&= \operatorname{Tr}\left(-tX^{-1}YX^{-1} + \sum_{n=2}^{\infty}(-1)^n(tYX^{-1})^n\right).
\end{aligned}$$

Therefore,

$$\lim_{t\downarrow 0}\frac{\operatorname{Tr}(X + tY)^{-1} - \operatorname{Tr} X^{-1}}{t} = -\operatorname{Tr}(X^{-1}YX^{-1}) = -\operatorname{Tr}(X^{-2}Y) = -\langle X^{-2}, Y\rangle.$$

Since $Y \in \mathbb{S}_n$ was arbitrary, we have $\nabla \operatorname{Tr} X^{-1} = -X^{-2}$.

(b) Define a function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ by $f(X) = \log\det X$. Prove $\nabla f(I) = I$. Deduce $\nabla f(X) = X^{-1}$ for any $X \in \mathbb{S}^n_{++}$.
Given a matrix $X \in \mathbb{S}^n_{++}$, notice that $\log\det X = \log\prod_{i=1}^n \lambda(X)_i = \sum_{i=1}^n \log\lambda(X)_i$. We can do something similar to the previous part, noting that $\log(1+x)$ has the series

$$\log(1+x) = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}, \quad |x| < 1.$$

Now consider that for $X \in \mathbb{S}_n$ satisfying $I + tX \in \mathbb{S}^n_{++}$,

$$\begin{aligned}\log\det(I+tX) &= \sum_{i=1}^{n}\log(1+t\lambda(X)_i) = \operatorname{Tr}\log(I+tX)\\ &= \operatorname{Tr}\left[\sum_{i=1}^{\infty}(-1)^{i+1}\frac{(tX)^i}{i}\right].\end{aligned}$$

Further noting that $\log\det I = 0$, we have

$$\lim_{t\downarrow 0}\frac{\log\det(I+tX)-\log\det I}{t} = \lim_{t\downarrow 0}\operatorname{Tr}\left(X + \sum_{i=2}^{\infty}(-1)^{i+1}t^{i-1}\frac{X^i}{n}\right) = \operatorname{Tr}(X) = \langle X, I\rangle.$$

Thus, $\nabla f(I) = I$.
Now for $X \in \mathbb{S}^n_{++}$ and $Y \in \mathbb{S}^n$, $t > 0$ such that $X + tY \in \mathbb{S}^n_{++}$,

$$\log\det(X+tY) = \log\det(X) + \log\det(I + tX^{-1}Y).$$

So

$$\lim_{t\downarrow 0}\frac{\log\det(X+tY)-\log\det(X)}{t} = \lim_{t\downarrow 0}\frac{\log\det(I+tX^{-1}Y)}{t} = \langle I, X^{-1}Y\rangle = \langle X^{-1}, Y\rangle.$$

Therefore, $\nabla f(X) = X^{-1}$.

11 ** (Kirchhoff's law). Consider a finite, undirected, connected graph with vertex set V and edge set E. Suppose that α and β in V are distinct vertices and that each edge ij in E has an associated "resistance" $r_{ij} > 0$ in $\mathbb{R}$. We consider the effect of applying a unit "potential different" between the vertices α and β. Let $V_0 = V\setminus\{\alpha,\beta\}$, and for "potentials" $x \in \mathbb{R}^{V_0}$ we define the "power" $p : \mathbb{R}^{V_0} \to \mathbb{R}$ by

$$p(x) = \sum_{ij\in E}\frac{(x_i - x_j)^2}{2r_{ij}},$$

where we set $x_\alpha = 0$ and $x_\beta = 1$.

(a) Prove the power function p has compact level sets.
For an arbitrary $M > 0$, suppose that there exists an $i \in V$ with $|x_i| > M$. Since

the graph is connected, there is a path from β to i, $\beta = v_1, v_2, \ldots, v_k = i$. Letting $r = \max_{ij \in E} r_{ij}$, We have

$$\begin{aligned}
\sum_{j=1}^{k-1} \frac{(x_{v_j} - x_{v_{j+1}})^2}{r_{v_j v_{j+1}}} &\geq \sum_{j=1}^{k-1} \frac{(|x_{v_j}| - |x_{v_{j+1}}|)^2}{r} \\
&\geq \frac{k-1}{r} \left(\sum_{j=1}^{k-1} \frac{|x_{v_j}| - |x_{v_{j+1}}|}{k-1} \right)^2 = \frac{k-1}{r} \left(\frac{|x_i|}{k-1} \right)^2 \\
&\geq \frac{M}{r(k-1)} \geq \frac{M}{r(|V|-1)}.
\end{aligned}$$

Therefore, $\|x\|_\infty > M \implies p(x) > \frac{M}{r(|V|-1)}$. In other words, $p(x) \leq \alpha$ for $\alpha > 0$ implies that $\|x\|_\infty \leq \alpha r(|V|-1)$. Therefore, p has bounded level sets. Since p is continuous with closed domain, the level sets are compact.

(b) Deduce the existence of a solution to the following equations (describing "conservation of current"):

$$\begin{aligned}
\sum_{j : ij \in E} \frac{x_i - x_j}{r_{ij}} &= 0 \quad \text{for } i \text{ in } V_0 \\
x_\alpha &= 0 \\
x_\beta &= 1.
\end{aligned}$$

$$p'(x; \pm e_i) = \frac{\partial p}{\partial x_i} = 2 \sum_{j \,:\, ij \in E} \frac{x_i - x_j}{r} = 0 \;\forall i \in V_0$$

implies the aforementioned solution. There is an $x \in \mathbb{R}^{V_0}$ satisfying the above because we showed p has a minimizer (due to $\mathbb{R}^{V_0}$ being nonempty and closed and p's level sets being bounded), and the above is a necessary condition of a minimizer.

(c) Prove the power function p is strictly convex.
Let $x \neq y \in \mathbb{R}^{V_0}$. Suppose that $x_k - x_l \neq y_k - y_l$ for some $kl \in E$. Then

$$\begin{aligned}
p(\lambda x + (1-\lambda y)) &= \sum_{ij \in E} \frac{(\lambda(x_i - x_j) + (1-\lambda)(y_i - y_j))^2}{r_{ij}} \\
&< \frac{\lambda(x_k - x_l)^2 + (1-\lambda)(y_k - y_l)^2}{r_{kl}} + \sum_{ij \neq kl \in E} \frac{\lambda(x_i - x_j)^2 + (1-\lambda)(y_i - y_j)^2}{r_{ij}} \\
&= \lambda p(x) + (1-\lambda) p(y),
\end{aligned}$$

using convexity of $(\cdot)^2$ (Exercise 8 (b)) for $ij \neq kl$ and strict convexity for $x_k - x_l \neq y_k - y_l$. Therefore, we just need to check when $x_i - x_j = y_i = y_j$ for all $ij \in E$. In this case, $x = y$, because the graph is connected; by starting from vertex β, whose value is shared in both graphs, the value of neighboring vertices must equal between the graphs, and so on.

(d) Use part (a) of Exercise 8 to show that the conservation of current equations in part (b) have a unique solution.
Since p is strictly convex, it has at most one minimizer. Any solution to the conservation of current equations is a minimizer, because it satisfies the first order sufficient condition for convex functions. And we have shown that a solution exists.

12 ** (Matrix completion). For a set $\delta \subset \{(i,j) \mid 1 \leq i \leq j \leq n\}$, suppose the subspace $L \subset \mathbb{S}^n$ of matrices with (i,j)th entry of zero for all (i,j) in δ satisfies $L \cap S^n_{++} \neq \emptyset$. By considering the problem (for $C \in \mathbb{S}^n_{++}$)

$$\inf\{\langle C, X\rangle - \log\det X \mid X \in L \cap \mathbb{S}^n_{++}\},$$

use Section 1.2, Exercise 14 and Corollary 2.1.3 to prove there exists a matrix X in $L \cap \mathbb{S}^n_{++}$ with $C - X^{-1}$ having (i,j)th entry of zero for all (i,j) not in δ.
This is saying that there is a matrix whose inverse matches C up to a set of entires not on the diagonal?

Proof. By Section 1.2, exercise 14, $\langle C, X\rangle - \log\det X$ has compact level sets, whose intersection with L (a closed set) is another compact set. Therefore, it has a minimizer. Using Exercise 10b, the gradient is $C - X^{-1}$, and if X is a minimizer, then for all $Y \in L \cap \mathbb{S}^n_{++}$ we have $\langle C - X^{-1}, Y - X\rangle \geq 0$. If $C - X^{-1}$ had a nonzero entry c for some $(i,j) \notin \delta$, then we could pick $Y = X - c\epsilon(e_i e_j^\top + e_j e_i^\top)$ for a small enough ϵ (because $\mathbb{S}^n_{++}$ is open) to obtain $\langle C - X^{-1}, Y - X\rangle = -2c^2\epsilon < 0$, a contradiction. Therefore, for all $(i,j) \notin \delta$, $(C - X^{-1})_{ij} = 0$. □

13 ** (BFGS update, cf 80). Given a matrix $C \in \mathbb{S}^n_{++}$ and vectors $s, y \in \mathbb{R}^n$ satisfying $s^\top y > 0$, consider the problem

$$\inf\{\langle C, X\rangle - \log\det X \mid Xs = y, X \in \mathbb{S}^n_{++}\}.$$

(a) Prove that for the problem above, the point

$$X = \frac{(y - \delta s)(y - \delta s)^\top}{s^\top(y - \delta s)} + \delta I$$

is feasible for small $\delta > 0$.
For small enough δ, $s^\top y > \delta\|s\|^2$, so $(y - \delta s)^\top s > 0$. Furthermore, if this holds then X is positive definite because the left term is PSD and the right is PD. Then

$$Xs = \frac{(y - \delta s)(y - \delta s)^\top s}{(y - \delta s)^\top s} + \delta s = y - \delta s + \delta s = y.$$

(b) Prove the problem has an optimal solution using Section 1.2, Exercise 14.
We know the function $\langle C, X\rangle - \log\det X$ has compact level sets in $\mathbb{S}^n_{++}$. We can intersect these with the closed set $\{X \in \mathbb{S}^n : Xs = y\}$ ($\cdot s$ is continuous) to obtain more compact level sets restricted to $Xs = y$. By part (a) some level set is nonempty. Thus there exists a minimizer.

(c) Use Corollary 2.1.3 to find the solution. (The solution is called the BFGS update of C^{-1} under the secant condition $Xs = y$.)

By Corollary 2.1.3, X is optimal if and only if for all $Y \in \mathbb{S}^n_{++}$ such that $Ys = y$,

$$\langle C - X^{-1}, Y - X \rangle \geq 0.$$

Note that $(Y - X)s = 0$, i.e. $Y - X \in \mathrm{null}(\cdot s)$. And in fact, since $\mathbb{S}^n_{++}$ is open, there exists $\epsilon > 0$ such that for all $N \in \mathrm{null}(\cdot s)$, $X + \epsilon N \in \mathbb{S}^n_{++}$. In other words, we must have

$$\langle C - X^{-1}, N \rangle = 0 \quad \forall\, N \in \mathrm{null}(\cdot s).$$

We have equality because $\mathrm{null}(\cdot s)$ is a subspace. Consider elements in $\mathbb{S}^n$ of the form $zs^\top + sz^\top$ for any $z \in \mathbb{R}^n$. We have

$$\langle zs^\top + sz^\top, N \rangle = \mathrm{Tr}(zs^\top N) + \mathrm{Tr}(Nsz^\top) = 0,$$

because $Ns = 0$ and $s^\top N = (Ns)^\top = 0$. Therefore, if $C - X^{-1} = zs^\top + sz^\top$ for some z and $X \in \mathbb{S}^n_{++}$, $Xs = y$, then X is optimal. Let's now check that X is feasible if

$$z = \frac{1}{s^\top y}\left[Cy - \left(\frac{1}{2} + \frac{y^\top C y}{2s^\top y}\right) s\right].$$

First let's check that $X^{-1} \in \mathbb{S}^n_{++}$ (which justifies the inverse notation). Let $x \in \mathbb{R}^n$.

$$\begin{aligned}
x^\top X^{-1} x &= x^\top C x - x^\top (sz^\top + zs^\top) x \\
&= x^\top C x - 2(x^\top s)(x^\top z) \\
&= x^\top C x - 2(x^\top s)\frac{1}{s^\top y}\left[x^\top C y - \left(\frac{1}{2} + \frac{y^\top C y}{2s^\top y}\right) x^\top s\right] \\
&= x^\top C x - 2\frac{s^\top x}{s^\top y} x^\top C y + \left(\frac{s^\top x}{s^\top y}\right)^2 y^\top C y + \frac{(x^\top s)^2}{s^\top y} \\
&= \left(x - \frac{s^\top x}{s^\top y} y\right)^\top C \left(x - \frac{s^\top x}{s^\top y} y\right) + \frac{(x^\top s)^2}{s^\top y}.
\end{aligned}$$

If $x \not\perp s$, then since C is PD and $(x^\top s)^2 > 0$, the whole expression is > 0. If $x \neq 0$ but $x \perp s$, then then the first term becomes $x^\top C x > 0$. Thus, $X^{-1} \in \mathbb{S}^n_{++}$.

Now we show that $Xs = y$, which since X is invertible is equivalent to $X^{-1} y = s$.

$$\begin{aligned}
X^{-1} y &= (C - sz^\top - zs^\top) y \\
&= Cy - s\left[\frac{y^\top C y}{s^\top y} - \left(\frac{1}{2} + \frac{y^\top C y}{2s^\top y}\right)\right] - \left[Cy - \left(\frac{1}{2} + \frac{y^\top C y}{2s^\top y}\right) s\right] \\
&= -s\frac{y^\top C y}{s^\top y} + \left(1 + \frac{y^\top C y}{s^\top y}\right) s = s.
\end{aligned}$$

Therefore, X defined by $X^{-1} = C - sz^\top - zs^\top$ is optimal. More explicitly,

$$X = \left\{C + \frac{1}{s^\top y}\left[\left(1 + \frac{y^\top C y}{s^\top y}\right) ss^\top - sy^\top C - Cys^\top\right]\right\}^{-1}.$$

This solution also happens to be unique, though it would take more effort to show.

14. ** Suppose intervals $I_1, I_2, \ldots, I_n \subset \mathbb{R}$ are nonempty and closed and the function $f : I_1 \times I_2 \times \ldots \times I_n \to \mathbb{R}$ is differentiable and bounded below. Use the idea of the proof of Proposition 2.1.7 to prove that for any $\epsilon > 0$ there exists a point $x^\epsilon \in I_1 \times I_2 \times \ldots \times I_n$ satisfying

$$(-\nabla f(x^\epsilon))_j \in N_{I_j}(x_j^\epsilon) + [-\epsilon, \epsilon] \quad (j = 1, 2, \ldots, n).$$

Proof. Consider the function $f(x) + \epsilon\|x\|_1$. This function has compact levels sets because f is lower bounded and its domain is closed. Therefore there exists a minimizer x^ϵ. Now suppose there exists a $j \in [n]$ where

$$(-\nabla f(x^\epsilon))_j \notin N_{I_j}(x_j^\epsilon) + [-\epsilon, \epsilon].$$

This implies that for some $x_j \in I_j$, $\operatorname{sgn}(x_j - x_j^\epsilon)\nabla f(x^\epsilon)_j < -\epsilon$. In other words,

$$\lim_{t \downarrow 0} \frac{f(x^\epsilon + te_j \operatorname{sgn}(x_j - x_j^\epsilon)) - f(x^\epsilon)}{t} < -\epsilon.$$

Then for small enough t,

$$\begin{aligned}
&f(x^\epsilon + te_j \operatorname{sgn}(x_j - x_j^\epsilon)) + \epsilon\|x^\epsilon + te_j \operatorname{sgn}(x_j - x_j^\epsilon)\|_1 - f(x^\epsilon) - \epsilon\|x^\epsilon\|_1 \\
&< -t\epsilon + \epsilon\|x^\epsilon - te_j \operatorname{sgn}(x_j - x_j^\epsilon)\| - \epsilon\|x^\epsilon\|_1 \\
&\leq -t\epsilon + \epsilon\|x^\epsilon - te_j \operatorname{sgn}(x_j - x_j^\epsilon) - x^\epsilon\|_1 = t\epsilon - t\epsilon = 0,
\end{aligned}$$

contradicting minimality of x^ϵ. (looks like ℓ_1 norm wasn't necessary?) □

15 * (Nearest polynomial with a given root). Consider the Euclidean space of copmlex polynomials of degree no more than n, with inner product

$$\left\langle \sum_{j=0}^{n} x_j z^j, \sum_{j=0}^{n} y_j z^j \right\rangle = \sum_{j=0}^{n} \bar{x}_j y_j.$$

Given a polynomial p in this space, calculate the nearest polynomial with a given complex root α, and prove the distance to this polynomial is $(\sum_{j=0}^{n} |\alpha|^{2j})^{-1/2}|p(\alpha)|$.

Proof. If $p(\alpha) = q(\alpha) = 0$, then $ap(\alpha) + bq(\alpha) = 0$ for any $a, b \in \mathbb{R}$. Therefore, the set $C = \{p : p(\alpha) = 0\}$ is a subspace. In fact, $C = \operatorname{null}(p \mapsto p(\alpha))$, and $p \mapsto p(\alpha)$ is a linear functional, so in fact $\dim C = n$. Actually, even more,

$$p(\alpha) = \langle \bar{\mathbf{a}}, p \rangle, \quad \bar{\mathbf{a}}_j := \bar{\alpha}^j, \; j = 0, \ldots, n.$$

From exercise 8 part (d), we know the distance from q to the nearest point in $C = \{p : \langle \bar{\mathbf{a}}, p \rangle = 0\}$ is

$$\frac{\langle \bar{\mathbf{a}}, q \rangle}{\|\bar{\mathbf{a}}\|} = \left(\sum_{j=0}^{n} |\alpha^{2j}| \right)^{-1/2} q(\alpha).$$

□

2.3 Theorems of the Alternative

Theorem 2.7 (2.2.1 (Gordan))**.** *Given $a^0, a^1, \ldots, a^m \in \mathbb{E}$, exactly one of the two systems has a solution:*

$$\sum_{i=0}^{m} \lambda_i a^i = 0, \quad 0 \le \lambda^0, \ldots, \lambda^m \in \mathbb{R}, \ \sum_{i=0}^{m} \lambda_i = 1 \tag{2.3.1}$$

$$\langle a^i, x \rangle < 0 \quad \forall i = 0, \ldots, m, \ x \in \mathbb{E}. \tag{2.3.2}$$

Define the function

$$f(x) = \log\left(\sum_{i=0}^{m} \exp(\langle a^i, x \rangle)\right).$$

Theorem 2.8 (2.2.6)**.** *The following statements are equivalent:*

(i) $f(x)$ *is bounded below;*

(ii) (2.3.1) *is solvable;*

(iii) (2.3.2) *is unsolvable.*

Lemma 2.9 (2.2.7 (Farkas))**.** *Let $c, a^1, \ldots, a^m \in \mathbb{E}$. Then exactly one of the following systems has a solution:*

$$\sum_{i=1}^{m} \mu_i a^i = c, \quad 0 \le \mu_1, \ldots, \mu_m \in \mathbb{R}, \tag{2.3.3}$$

$$\langle a^i, x \rangle \le 0 \ \forall i = 1, \ldots, m, \ \langle c, x \rangle > 0, \ x \in \mathbb{E}. \tag{2.3.4}$$

Farkas's lemma has the interpretation of the finitely generated cone

$$C = \{\mu_1 a^1 + \ldots + \mu_m a^m : 0 \le \mu_1, \ldots, \mu_m \in \mathbb{R}\}$$

being linearly separable from any point $c \notin C$.

2.4 Exercises for 2.2

1. Prove the implications (ii) $\implies$ (iii) $\implies$ (i) in Theorem 2.2.6.

Proof. (ii) says there is a set of $0 \le \lambda^0, \ldots, \lambda^m \in \mathbb{R}$ summing to 1 where $\sum_{i=0}^m \lambda_i a^i = 0$. If this is the case, then if any $x \in \mathbb{E}$ satisfies $\langle a^i, x \rangle < 0$ for all $i = 0, \ldots, m$,

$$0 = \langle 0, x \rangle = \sum_{i=0}^{m} \lambda_i \langle a^i, x \rangle < 0,$$

a contradiction, since some λ_i is nonzero. Therefore, (ii) implies (iii).
If (iii), i.e. there does not exist $x \in \mathbb{E}$ where $\langle a^i, x \rangle < 0$ for all $i = 0, \ldots, m$, then for any $x \in \mathbb{E}$, $\langle a^{i^*}, x \rangle \ge 0$ for some $i^* \in \{0, \ldots, m\}$. By nonnegativity of exp, and increasingness of log,

$$f(x) = \log\left(\sum_{i=0}^{m} \exp(\langle a^i, x \rangle)\right) \ge \log \exp(\langle a^{i^*}, x \rangle) = \langle a^{i^*}, x \rangle \ge 0.$$

Therefore, f is bounded below, i.e., (i). $\square$

2.

(a) Prove the orthogonal projection $P_{\mathbb{Y}} : \mathbb{E} \to \mathbb{Y}$ is a linear map.
It may be easier to do this by letting U's columns be an orthogonal basis for $\mathbb{Y}$ and showing that $UU^\top = P_{\mathbb{Y}}$, but in the spirit of convex analysis here's a more "direct" proof.
If we prove that $P_{\mathbb{Y}}(\alpha x + y) = \alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y)$ for any $\alpha \in \mathbb{R}, x, y \in \mathbb{E}$, we will have proved $P_{\mathbb{Y}}$ is linear. Let us directly show $\alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y)$ is the nearest point in $\mathbb{Y}$ to $\alpha x + y$. Let $z \in \mathbb{Y}$ be arbitrary.

$$\begin{aligned}
&\|\alpha x + y - z\|^2 = \|\alpha x + y - \alpha P_{\mathbb{Y}}(x) - P_{\mathbb{Y}}(y) + \alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y) - z\|^2 \\
&= \|\alpha x + y - \alpha P_{\mathbb{Y}}(x) - P_{\mathbb{Y}}(y)\|^2 + 2\langle \alpha x + y - \alpha P_{\mathbb{Y}}(x) - P_{\mathbb{Y}}(y), \alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y) - z\rangle \\
&\qquad\qquad\qquad\qquad\qquad + \|\alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y) - z\|^2 \\
&= \|\alpha x + y - \alpha P_{\mathbb{Y}}(x) - P_{\mathbb{Y}}(y)\|^2 + \|\alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y) - z\|^2 \\
&\geq \|\alpha x + y - \alpha P_{\mathbb{Y}}(x) - P_{\mathbb{Y}}(y)\|^2.
\end{aligned}$$

To see that the inner product is 0, note that $\alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y) - z \in \mathbb{Y}$, and for any $v \in \mathbb{Y}$,

$$\langle \alpha x - \alpha P_{\mathbb{Y}}(x), v\rangle = \alpha\langle x - P_{\mathbb{Y}}(x), v\rangle = \langle y - P_{\mathbb{Y}}, v\rangle = 0$$

because of the nearest point characterization for subspaces (for any $\bar{x} \in \mathbb{Y}$, $\mathbb{Y} - \bar{x} = \mathbb{Y} = -\mathbb{Y}$.) We have shown $P_{\mathbb{Y}}(\alpha x + y) = \alpha P_{\mathbb{Y}}(x) + P_{\mathbb{Y}}(y)$, so this completes the proof.

(b) Give a direct proof of the Farkas lemma for the case $m = 1$.
The $m = 1$ case says that for two elements $a, c \in \mathbb{E}$, one of the following holds:

$$\begin{aligned}
&1.\ \exists\, \lambda \geq 0,\ \lambda a = c, \\
&2.\ \exists\, x \in \mathbb{E},\ \langle a, x\rangle \leq 0, \quad \langle c, x\rangle > 0.
\end{aligned}$$

Define $\mathbb{Y} = \operatorname{Span}\{a\}$, $u = c - P_{\mathbb{Y}}(c)$. Our comment at the end of part (b) on the nearest point characterization for subspaces shows that $u \perp \mathbb{Y}$. Thus, if $u \neq 0$,

$$0 < \|u\|^2 = \langle c - P_{\mathbb{Y}}(c), c - P_{\mathbb{Y}}(c)\rangle = \langle c, c - P_{\mathbb{Y}}(c)\rangle = \langle c, u\rangle.$$

In other words, $\langle c, u\rangle > 0$. At the same time, $\langle a, u\rangle = 0$ because $a \in \mathbb{Y}$. This gives a solution to 2.
The other case where 1. is not solvable is when $\lambda a = c$ for a $\lambda < 0$, $a \neq 0$. In that case, $\langle a, c\rangle = \lambda\|a\|^2 < 0$, and $\langle c, c\rangle > 0$. Therefore, 1. being unsolvable implies 2. being solvable. We have shown in Exercise 1 that 1. being solvable implies 2. being unsolvable in the general case, so we are done.

3. Use the Basic separation theorem (2.1.6) to give another proof of Gordan's theorem.

Proof. First let us show that the set

$$C = \left\{ \sum_{i=0}^m \lambda_i a^i \,\middle|\, 0 \leq \lambda_0, \ldots, \lambda_m \in \mathbb{R},\ \sum_{i=0}^m \lambda_i = 1 \right\}$$

is convex and closed. By Chapter 1.1 Exercise 2(c), $C = \operatorname{conv}\{a^0, \ldots, a^m\}$ is convex. To see it is closed, for any sequence

$$\lambda_0^k a^0 + \lambda_1^k a^1 + \ldots + \lambda_m^k a^n \xrightarrow{k\to\infty} a,$$

We can take a convergent subsequence of each λ_i^k, $i = 0, \ldots, m$ and use the closedness of $[0,1]$ to produce an element of C equal to a. Now basic separation implies that for any point $c \notin C$, there exists $x \in \mathbb{E}$ where $\langle a, x\rangle < \langle c, x\rangle$ for every $a \in C$. By convexity of C, this is equivalent to if $\langle a^i, x\rangle < \langle c, x\rangle$ for each $i = 0, \ldots, m$.
Now the first system in Gordan's lemma being unsolvable is equivalent to $0 \notin C$. Thus, there exists $x \in \mathbb{E}$ where $\langle a^i, x\rangle < \langle 0, x\rangle = 0$ for each $i = 0, \ldots, m$, i.e. the second system is solvable. And we already know the first system being solvable implies the second system is unsolvable. This proves Gordan's theorem. □

4. * Deduce Gordan's theorem from the Farkas lemma. (Hint: Consider the elements $(a^i, 1)$ of the space $\mathbb{E} \times \mathbb{R}$.)

Proof. Denote $\tilde{a}^i = (a^i, 1)$. Notice that

$$\exists\, 0 \le \lambda_0, \ldots, \lambda_m \in \mathbb{R},\ \sum_{i=0}^m \lambda_i = 1 : \qquad \sum_{i=0}^m \lambda_i a^i = 0 \tag{2.4.1}$$

$$\iff \exists\, 0 \le \lambda_0, \ldots, \lambda_m \in \mathbb{R} : \qquad \sum_{i=0}^m \lambda_i \tilde{a}^i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{2.4.2}$$

Thus, if (2.4.1) does not hold, then (2.4.2) does not hold. If (2.4.2) does not hold, then by Farkas's Lemma, there exists $x \in \mathbb{E} \times \mathbb{R}$ where $\langle \tilde{a}^i, x\rangle \le 0$ for all $i = 0, \ldots, m$ and $\langle (0,1), x\rangle > 0$. If $d \in \mathbb{R}$ is the last coordinate of x and $x' \in \mathbb{E}$ are the rest, we see that $d > 0$ and $\langle a^i, x'\rangle \le -d < 0$ for all $i = 0, \ldots, m$. In other words, we have found a solution to the second system of Gordan's theorem. Proving that (2.4.1) being solvable implies the second system was done in Exercise 1, so we are done. □

5 * (Caratheodory's theorem 52). Suppose $\{a^i \mid i \in I\}$ is a finite set of points in $\mathbb{E}$. For any subset J of I, define the cone

$$C_J = \left\{ \sum_{i \in J} \mu_i a^i \,\middle|\, 0 \le \mu_i \in \mathbb{R},\ (i \in J) \right\}.$$

(a) Prove the cone C_I is the union of those cones C_J for which the set $\{a^i \mid i \in J\}$ is linearly independent. Furthermore, prove directly that any such cone C_J is closed.
We prove the first part by showing that any element of the cone can be expressed as a conic combination of a linearly independent set of the a^i's. Let $x = \sum_{i\in I} \mu_i a^i \in C$, and let N be the indices of nonzero weights, $N = \{i \in I : \mu_i > 0\}$. Now if $A = \{a^i : i \in N\}$ is not linearly independent, there exist $c_i \in \mathbb{R}$, $i \in N$ not all 0

$$\sum_{i \in N} c_i a^i = 0.$$

Let $r = \min_{i \in N} \mu_i / |c_i|$, returning ∞ if $c_i = 0$, and let i^* be any index achieving the minimum. Indeed at least one $c_i \neq 0$. Then

$$x = x - 0 = \sum_{i \in I} (\mu_i - r \operatorname{sgn}(c_{i^*}) c_i) a^i.$$

We have $\mu_{i^*} - r \operatorname{sgn}(c_{i^*}) c_{i^*} = \mu_{i^*} - \mu_{i^*} = 0$. Furthermore, for all i, $\mu_i - r \operatorname{sgn}(c_{i^*}) c_i \geq \mu_i - |rc_i| \geq \mu_i - (\mu_i/|c_i|)|c_i| = 0$, so we still have a conic combination of elements. Yet, we have reduced the number of nonzero coefficients by 1. By repeating this process as long as the set of elements with nonzero coefficients is linearly dependent, we can represent x as a conic combination of linearly dependent elements a^i.
Now we prove each C_J is closed, where $\{a^i : i \in J\}$ is linearly independent. Let A be the matrix with $a^i, i \in J$ for its columns. Then $A^\top A$ is invertible. Then if Ax^n is a sequence in C_J converging to a, then by continuity of $(A^\top A)^{-1} A^\top$,

$$(A^\top A)^{-1} A^\top (Ax^n) = x^n \to (A^\top A)^{-1} A^\top a.$$

In other words, x^n converges to a limit x^*. Then $Ax^* = a$, because A is continuous.

(b) Deduce that any finitely generated cone is closed.
Any finitely generated cone is a union of a finite number of closed cones, by the above part. A finite union of closed sets is closed.

(c) If the point x lies in $\operatorname{conv}\{a^i \mid i \in I\}$, prove that in fact there is a subset $J \subset I$ of size at most $1 + \dim \mathbb{E}$ such that x lies in $\operatorname{conv}\{a^i \mid i \in J\}$. (Hint: Apply part (a) to the vectors $(a^i, 1)$ in $\mathbb{E} \times \mathbb{R}$.)
Define $\tilde{a}^i = (a^i, 1)$. Suppose

$$\operatorname{conv}\{a^i : i \in I\} \ni x = \sum_{i \in I} \lambda_i a^i, \quad 0 \leq \lambda_i \in \mathbb{R}, \sum_{i \in I} \lambda_i = 0.$$

Then

$$\tilde{x} := \sum_{i \in I} \lambda_i \tilde{a}^i \in \operatorname{conv}\{\tilde{a}^i : i \in I\} \subset \operatorname{cone}\{\tilde{a}^i : i \in I\},$$

where $\operatorname{cone}\{\tilde{a}^i : i \in I\} := \{\sum_i \mu_i a^i : \mu_i \geq 0\}$. Thus, there is a $J \subset I$ such that $\tilde{x} = \sum_{i \in J} \nu_i \tilde{a}^i$, $\nu_i \geq 0$ for all $i \in J$. However, the last coordinate of $\tilde{x}$ also constrains $\sum_{i \in J} \nu_i = 1$. Therefore, $\sum_{i \in J} \nu_i a^i$ is a convex combination equaling x, using $|J| \leq \dim(\mathbb{E} \times \mathbb{R}) = \dim \mathbb{E} + 1$ elements of $\{a^i : i \in I\}$.

(d) Use part (c) to prove that if a subset of $\mathbb{E}$ is compact then so is its convex hull.
Let the subset be denoted S. By part (c), for any sequence x^n in $\operatorname{conv} S$, each x^n can be expressed as a convex combination of $\dim \mathbb{E} + 1$ points in $\mathbb{E}$; just apply (c) to the finite set of points which x^n was previously expressed as a convex combination of. Note that the $\dim \mathbb{E} + 1$ points may change from element to element of the sequence. Thus

$$\forall\, n = 1, 2, \ldots, \quad x^n = \sum_{i=1}^{1+\dim \mathbb{E}} \lambda_i^n (a^i)^n, \qquad \forall i, n,\ (a^i)^n \in S$$

But, for each i, since S is compact, we can take a convergent subsequence of $(a^i)^n$. Furthermore, we can take a convergent subsequence for each λ_i^n. Then if $\lambda_i^*, (a^i)^*$ denote the limits, $\sum_{i=1}^{\dim \mathbb{E}+1} \lambda_i^*(a^i)^* = x$ which is in the convex hull, because since S is closed, each $(a^i)^* \in S$. Furthermore, $\{(\lambda_1, \ldots, \lambda_n) : \forall\, i \in [n],\, \lambda_i \geq 0,\, \sum_{i=1}^n \lambda_i = 1\}$ is a closed set.

6. * Give another proof of the Farkas lemma by applying the Basic separation theorem (2.1.6) to the set C defined by (2.2.11) (shown below) and using the fact that any finitely generated cone is closed.

$$C = \left\{ \sum_{i=1}^m \mu_i a^i \,\middle|\, 0 \leq \mu_1, \ldots, \mu_m \in \mathbb{R} \right\}.$$

Proof. The previous exercise proves that any finitely generated cone, e.g. C, is closed. By basic separation, for any $c \notin C$, there exists $a \in \mathbb{E}$ and $b \in \mathbb{R}$ such that $\langle a, c\rangle > b \geq \langle a, x\rangle$ for all $x \in C$. Because $0 \in C$, we have $b \geq 0$. We will now show that in fact, $\langle a, x\rangle \leq 0$ for every $x \in C$. This is because if $\langle a, x\rangle > 0$ for some x, then there exists $r > 0$ such that $\langle a, rx\rangle > b$, say $r = 2b/\langle a, x\rangle$. Thus, $\langle a, c\rangle > b \geq 0 \geq \langle a, x\rangle$ for all $x \in C$. We can conclude (2.3.4) is solvable by taking $x = a^1, \ldots, a^m$. We have shown that (2.3.3) being unsolvable implies (2.3.4) is solvable. The other direction is easy. This concludes the proof of Farkas's Lemma. □

7 ** (Ville's theorem). With the function f defined by (2.2.5) ($f = \log \sum_i \exp\langle a^i, \cdot\rangle$) (with $\mathbb{E} = \mathbb{R}^n$), consider the optimization problem

$$\inf\{f(x) \mid x \geq 0\} \tag{2.4.3}$$

and its relationship with the two systems

$$\sum_{i=0}^m \lambda_i a^i \geq 0, \quad \sum_{i=0}^m \lambda_i = 1, \quad 0 \leq \lambda_0, \ldots, \lambda_m \in \mathbb{R} \tag{2.4.4}$$

and

$$\langle a^i, x\rangle < 0 \;\text{ for }\; i = 0, 1, \ldots, m, \quad x \in \mathbb{R}^n_+. \tag{2.4.5}$$

Imitate the proof of Gordan's theorem (using Section 2.1, Exercise 14) to prove the following are equivalent:

(i) Problem (2.4.3) is bounded below.

(ii) System (2.4.4) is solvable.

(iii) System (2.4.5) is unsolvable.

Generalize by considering the problem $\inf\{f(x) \mid x_j \geq 0 (j \in J)\}$.

Proof. **(i) $\Longrightarrow$ (ii):**
By section 2.1, Exercise 14, for each ϵ, there exists a $\bar{x}$ where

$$\forall j = 1, \ldots, n, \quad (\nabla f(\bar{x}))_j = \left(\sum_{i=0}^{m} \lambda_{i,\bar{x}} a^i \right)_j \geq -\epsilon,$$

where $\lambda_{i,\bar{x}} = \exp(\langle a^i, \bar{x} \rangle) / \sum_{k=0}^{m} \exp(\langle a^k, \bar{x} \rangle) > 0$. This is because $N_{[0,\infty)}(x) = \{0\}$ if $x > 0$ and $N_{[0,\infty)}(x) = [-\infty, 0]$ if $x = 0$. Either way, $-x \in N_{[0,\infty)} + [-\epsilon, \epsilon] \implies x \geq -\epsilon$.
Now we can take a sequence $\bar{x}^k$ such that $\lim_{k\to\infty} \nabla f(\bar{x}^k) \geq 0$. Or to be more precise, for each j, $(\nabla f(\bar{x}^k))_j$ eventually is greater than an arbitrarily small negative number. Thus, if we take a convergent subsequence of $(\lambda_{i,\bar{x}^k})_{i=1}^n$, the limit $(\lambda_i^*)_{i=1}^n$ satisfies $\sum_{i=1}^n \lambda_i^* a^i \geq 0$, and $\lambda_i^* \geq 0 \; \forall i \in [n]$, $\sum_{i=1}^n \lambda_i = 1$.
(ii) $\Longrightarrow$ (iii):
Suppose (2.4.4) and (2.4.5) are both solvable. Let $x \in \mathbb{R}^n_+$ be the solution to (2.4.5) satisfying $\langle a^i, x \rangle < 0$ for every $i = 0, \ldots, m$. Then for the solution to (2.4.4), $\lambda_0, \ldots, \lambda_m \geq 0$, $\sum_i \lambda_i = 1$,

$$0 \leq \left\langle \sum_{i=0}^{n} \lambda_i a^i, x \right\rangle = \sum_{i=0}^{n} \lambda_i \langle a^i, x \rangle < 0,$$

a contradiction. The first inequality holds because $\sum_{i=0}^n \lambda_i a^i \geq 0$ and $x \in \mathbb{R}^n_+$.
(iii) $\Longrightarrow$ (i)
If (2.4.5) is unsolvable, then for every $x \in \mathbb{R}^n_+$, there exists an $i \in \{0, \ldots, m\}$ such that $\langle a^i, x \rangle \geq 0$. Thus, $f(x) = \log\left(\sum_{k=0}^m \exp(\langle a^k, x \rangle)\right) \geq \langle a^i, x \rangle \geq 0$, implying the problem (2.4.3) is lower bounded.
To generalize to $\inf\{f(x) \mid x_j \geq 0 \; \forall j \in J\}$, in (2.4.4) we change $\sum_i \lambda_i a^i \geq 0$ to $\left(\sum_i \lambda_i a^i\right)_j \geq 0 \; \forall j \in J$ and $\left(\sum_i \lambda_i a^i\right)_j = 0 \; \forall j \notin J$. In (2.4.5) we change $x \in \mathbb{R}^n_+$ to $x \in \mathbb{R}$, $x_j \geq 0 \; \forall j \in J$. □

8 ** (Stiemke's theorem). Consider the optimization problem (2.2.4) and its relationship with the two systems

$$\sum_{i=0}^{m} \lambda_i a^i = 0, \; 0 < \lambda_0, \lambda_1, \ldots, \lambda_m \in \mathbb{R} \tag{2.4.6}$$

and

$$\langle a^i, x \rangle \leq 0 \;\text{ for } i = 0, 1, \ldots, m, \;\text{ not all } 0, \quad x \in \mathbb{E}. \tag{2.4.7}$$

Prove the following are equivalent:

(i) Problem (2.2.4) has an optimal solution.

(ii) System (2.4.6) is solvable.

(iii) System (2.4.7) is unsolvable.

Proof. **(i) $\implies$ (ii):**
If there is an optimal solution x^* to problem 2.2.4, then by the first order necessary condition (Prop 2.1.1),

$$\nabla f(x^*) = \sum_{i=0}^{m} \frac{\exp(\langle a^i, x^* \rangle)}{\sum_{j=0}^{m} \exp(\langle a^j, x^* \rangle)} a^i = 0.$$

By setting $\lambda_i = \exp(\langle a^i, x^* \rangle) / \sum_{j=0}^{m} \exp(\langle a^j, x^* \rangle) > 0$ we obtain a solution to (2.4.6).
(ii) $\implies$ (iii):
Suppose that (2.4.6) and (2.4.7) are both solvable. Let $(\lambda_i^*)_{i=0}^m$, x^* be the respective solutions. By (2.4.7), there must be some $i^* \in \{0, \ldots, m\}$ for which $\langle a^{i^*}, x^* \rangle < 0$. But since every $i \in \{0, \ldots, m\}$ satisfies $\langle a^i, x^* \rangle \leq 0$, this contradicts (2.4.6), which implies $\sum_{i=0}^m \lambda_i \langle a^i, x^* \rangle = 0$, as $\lambda_{i^*} \langle a^{i^*}, x^* \rangle < 0$.
(iii) $\implies$ (i):
Define the linear subspace $A = \operatorname{Span}\{a^0, a^1, \ldots, a^m\} \subset \mathbb{E}$. Let $A^\perp$ be the orthogonal subspace. If (iii), i.e. (2.4.7) is unsolvable, then every $x \in \mathbb{E}$ either satisfies $\langle a^i, x \rangle = 0$ for every $i \in \{0, \ldots, m\}$ or $\langle a^i, x \rangle > 0$ for some $i \in \{0, \ldots, m\}$. I.e., either $x \in A^\perp$, or $\langle a^i, x \rangle > 0$ for some i.
Now define $S = \operatorname{bd} B \cap A$ as the unit circle intersected with A. For any $x \in S$, $x \notin A^\perp$, because the only element in both A and $A^\perp$ is 0. Therefore, $\langle a^k, x \rangle > 0$ for some $k \in \{0, \ldots, m\}$. Define the function $g(x) = \max_i \langle a^i, x \rangle$. The maximum of continuous functions is continuous. Thus, g is continuous, and over the compact set S attains a minimum $m > 0$. Now for any $x \in A$, $\|x\| > r > 0$, with $k = \arg\max_i \langle a^i, x \rangle$, we have $g(x) = \langle a^k, x \rangle = \|x\| \langle a^k, x/\|x\| \rangle > rm$, and

$$f(x) = \log \left(\sum_{i=0}^{m} \exp(\langle a^i, x \rangle) \right) \geq \log \exp(\langle a^k, x \rangle) = g(x) > rm.$$

Therefore, on A, f has bounded sublevel sets. Since A is closed, by Proposition 1.1.3, f has a minimizer a^* over A. And since any $x \in \mathbb{E}$ can be expressed as $x = a + a_\perp$, $a \in A, a_\perp \in A^\perp$, and $f(x) = f(a + a_\perp) = f(a) \geq f(a^*)$, a^* is in fact a global minimizer of f on $\mathbb{E}$. In other words, problem (2.2.4) has an optimal solution.
To generalize with the problem $\inf\{f(x) \mid x_j \geq 0 \ (j \in J)\}$, replace $\sum_i \lambda_i a^i = 0$ in (2.4.7) with $\left(\sum_i \lambda_i a^i\right)_j \geq 0$ for $j \in J$ and $\left(\sum_i \lambda_i a^i\right)_j = 0$ for $j \notin J$. In (2.4.7), add the constraint $x_j \geq 0$ for all $j \in J$. $\square$

9 ** (Schur-convexity). The *dual cone* of the cone $\mathbb{R}^n_\geq$ is defined by

$$\left(\mathbb{R}^n_\geq\right)^+ = \{y \in \mathbb{R}^n \mid \langle x, y \rangle \geq 0 \text{ for all } x \text{ in } \mathbb{R}^n_\geq\}.$$

(a) Prove a vector y lies in $\left(\mathbb{R}^n_\geq\right)^+$ iff

$$\sum_{i=1}^{j} y_i \geq 0 \ \text{ for } j = 1, 2, \ldots, n-1, \quad \sum_{i=1}^{n} y_i = 0.$$

Suppose $y \in \left(\mathbb{R}^n_\geq\right)^+$. We must have $\sum_i y_i = 0$, or else $\langle \mathbf{1}, y \rangle > 0$ or $\langle -\mathbf{1}, y \rangle > 0$ and both $\mathbf{1}, -\mathbf{1} \in \mathbb{R}^n_\geq$. We also must have $\sum_{i=1}^{j} y_i \geq 0$ for every $j = 1, \ldots, n-1$, because

otherwise the vector with j ones at the start and the rest zeros is in $\mathbb{R}^n_\geq$ and has a negative inner product with y.
Now suppose $y \in \mathbb{R}^n$ satisfies $\sum_{i=1}^n y_i = 0$ and $\sum_{i=1}^j y_i \geq 0$ for each $j = 1, \ldots, n$. Take $x \in \mathbb{R}^n_\geq$. We have $\langle x, y \rangle =$

$$\sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i \left(\sum_{j=1}^i y_j - \sum_{j=1}^{i-1} y_j \right) = \sum_{i=1}^{n-1} (x_i - x_{i+1}) \sum_{j=1}^i y_j + x_n \overbrace{\sum_{j=1}^n y_j}^{0} \geq 0.$$

(b) By writing $\sum_{i=1}^j [x]_i = \max_k \langle a^i, x \rangle$ for some suitable set of vector a^k, prove that the function $x \mapsto \sum_{i=1}^j [x]_i$ is convex. (Hint: Use Section 1.1, Exercise 7.)
By Section 1.1., Exercise 7, a maximum of inner product functions is convex. We let $\{a^k\}_k$ be the set of zero one vectors with exactly j ones.

(c) Deduce that the function $x \mapsto [x]$ is $(\mathbb{R}^n_\geq)^+$-*convex*, that is:

$$z := \lambda[x] + (1 - \lambda)[y] - [\lambda x + (1 - \lambda) y] \in \left(\mathbb{R}^n_\geq\right)^+ \quad \text{for } 0 \leq \lambda \leq 1.$$

$$\sum_{i=1}^n z_i = \lambda \langle \mathbf{1}, x \rangle + (1 - \lambda) \langle \mathbf{1}, y \rangle - \langle \mathbf{1}, \lambda x + (1 - \lambda y \rangle = 0,$$

and defining i_x, i_y s.t. $[\lambda x + (1 - \lambda) y]_i = \lambda x_{i_x} + (1 - \lambda) y_{i_x}$,

$$\sum_{i=1}^j \lambda [x]_j + (1 - \lambda)[y]_j \geq \sum_{i=1}^j \lambda [x]_{i_x} + (1 - \lambda)[y]_{i_y}, \quad \forall j = 1, \ldots, n$$

which shows that $z \in (\mathbb{R}_\geq)^+$.

(d) Use Gordan's theorem and Proposition 1.2.4 (rearrangement inequality) to deduce that for any x and y in $\mathbb{R}^n_\geq$, if $y - x$ lies in $\left(\mathbb{R}^n_\geq\right)^+$ then x lies in $\operatorname{conv}(\mathbb{P}^n y)$.
Notice that for each $z \in \mathbb{R}^n_\geq$,

$$0 \leq \langle z, y - x \rangle = z^\top y - z^\top x, \qquad [y] = y, \ [x] = x. \tag{2.4.8}$$

Therefore, if we define $a^\Pi = \Pi y - x$ for each permutation matrix $\Pi \in \mathbb{P}^n$, then the system

$$z \in \mathbb{R}^n, \quad \langle a^\Pi, z \rangle < 0 \qquad \forall \Pi \in \mathbb{P}^n$$

is unsolvable, because by (2.4.8) for any $z \in \mathbb{R}^n$ there exists Π where $\langle a^\Pi, z \rangle = [z]^\top [y] - z^\top x \geq [z]^\top y - [z]^\top x \geq 0$ also using the rearrangement inequality. Thus, the system

$$\exists \lambda_\Pi \geq 0 \ (\Pi \in \mathbb{P}^n), \ \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi = 0 : \quad \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi (\Pi y - x) = 0$$

is solvable. This is equivalent to stating that $x \in \operatorname{conv}(\mathbb{P}^n y)$.

(e) A functon $f : \mathbb{R}^n_\geq \to \mathbb{R}$ is *Schur-convex* if

$$x, y \in \mathbb{R}^n_\geq,\ y - x \in \left(\mathbb{R}^n_\geq\right)^+ \implies f(x) \leq f(y).$$

Prove that if f is convex, then it is Schur-convex if and only if it is the restriction to $\mathbb{R}^n_\geq$ of a *symmetric* convex function $g : \mathbb{R}^n \to \mathbb{R}$ (where by symmetric we mean $g(x) = g(\Pi x)$ for any $x \in \mathbb{R}^n$ and any permutation matrix Π).

First we prove the if direction, i.e. assume f is the restriction of some symmetric convex g. Using part (d), if $x, y \in \mathbb{R}^n_\geq$ and $y - x \in (\mathbb{R}^n_\geq)^+$ then $x = \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi \Pi y$ for nonnegative λ_Π summing to 1. Thus,

$$f(x) = g\left(\sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi \Pi y\right) \leq \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi g(\Pi y) = \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi g(y) = g(y) = f(y).$$

Therefore, f is Schur-convex.

Now we prove the only if direction. Assume f is convex and Schur-convex. Define a function $g : \mathbb{R}^n \to \mathbb{R}$ by $g(x) = f([x])$. By definition, g is symmetric. Now for $\lambda \in [0, 1]$ and $x, y \in \mathbb{R}^n$,

$$\begin{aligned} g(\lambda x + (1-\lambda) y) &= f([\lambda x + (1-\lambda) y]) \leq f(\lambda [x] + (1-\lambda)[y]) \\ &\leq \lambda f([x]) + (1-\lambda) f([y]) = \lambda g(x) + (1-\lambda) g(y), \end{aligned}$$

using part (c) and the fact that f is Schur-convex, and convex. Thus g is convex.

2.5 2.3 Max-functions

First order necessary conditions for an optimization problem with constraints involving differentiable functions will be obtained by considering the *max-function*

$$g(x) = \max_{i=0,1,\dots,m} g_i(x). \tag{2.5.1}$$

Proposition 2.10 (2.3.2 (Directional derivatives of max-functions))**.** *Let $\bar{x} \in \operatorname{int} C \subset \mathbb{E}$. Suppose $g_0, \dots, g_m : C \to \mathbb{R}$ are continuous and differentiable at $\bar{x}$. Furthermore, denote $K = \{i \mid g_i(\bar{x}) = g(\bar{x})\}$. Then for all directions $d \in \mathbb{E}$, the directional derivative of the max-function g in* (2.5.1) *is*

$$g'(\bar{x}; d) = \max_{i \in K} \langle \nabla g_i(\bar{x}), d \rangle. \tag{2.5.2}$$

Most of this book considers optimization problems of the form

$$\begin{array}{ll} \inf & f(x) \\ \text{subject to} & g_i(x) \leq 0 \quad \forall i \in I \\ & h_j(x) = 0 \quad \forall j \in J \\ & x \in C \end{array} \tag{2.5.3}$$

The set of all feasible x is called the *feasible region*. If the feasible region is empty, the problem is called *inconsistent*. We say a feasible point $\bar{x}$ is a *local minimizer* if for all **feasible** x close to $\bar{x}$, $f(x) \geq f(\bar{x})$.
We begin with the differentiable inequality constrained problem

$$\begin{array}{ll} \inf & f(x) \\ \text{subject to} & g_i(x) \leq 0 \quad \forall i = 1, \dots, m \\ & x \in C. \end{array} \tag{2.5.4}$$

For a feasible point $\bar{x}$ we define the *active set* $I(\bar{x}) = \{i : g_i(\bar{x}) = 0\}$. If $\bar{x} \in \operatorname{int} C$, we call $\lambda \in \mathbb{R}^m_+$ a *Lagrange multiplier vector* for $\bar{x}$ if $\bar{x}$ is a critical point of the *Lagrangian* $\mathcal{L}(x; \lambda) = f(x) + \sum_{i=1}^m \lambda g_i(x)$, and *complementary slackness* holds: $\lambda_i = 0$ for all $i \notin I(\bar{x})$.

Theorem 2.11 (2.3.6 (Fritz John conditions))**.** *Suppose problem* (2.5.4) *has a local minimizer $\bar{x} \in \operatorname{int} C$. If f, g_i $(i \in I(\bar{x}))$ are differentiable, then there exist λ_0, λ_i $(i \in I(\bar{x})) \in \mathbb{R}_+$ not all 0 such that*

$$\lambda_0 \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0,$$

Assumption 2.1 (2.3.7 (The Mangasarian-Fromovitz constraint qualification))**.** There is a direction $d \in \mathbb{E}$ satisfying $\langle \nabla g_i(\bar{x}), d \rangle < 0$ for all $i \in I(\bar{x})$.

Theorem 2.12 (2.3.8 (Karush-Kuhn-Tucker conditions))**.** *Suppose problem* (2.5.4) *has a local minimizer $\bar{x} \in \operatorname{int} C$. If f, g_i $(i \in I(\bar{x}))$ are differentiable and the Mangasarian-Fromovitz constraint qualification* (2.1) *holds, then there exists a Lagrange multiplier vector $\lambda \in \mathbb{R}^m_+$ for $\bar{x}$.*

2.6 Exercises for 2.3

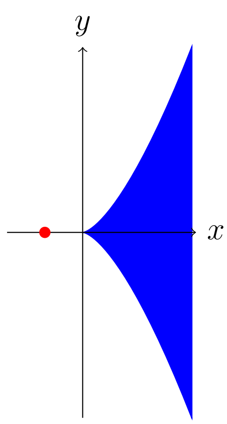
1. Prove by induction that if the functions $g_0, g_1, \dots, g_m : \mathbb{E} \to \mathbb{R}$ are all continuous at the point $\bar{x}$ then so is the max-function $g(x) = \max_i g_i(x)$.

Proof. Let's begin with two functions, g_0 and g_1. For the remainder of the proof, define the following quantities. Given $\epsilon > 0$, for $i \in \{0, 1\}$, $\exists \delta_i > 0$ such that $|x - \bar{x}| < \delta_i \implies |g_i(x) - g_i(\bar{x})| < \epsilon$. Define $\delta = \min\{\delta_0, \delta_1\}$.
If $g_0(\bar{x}) > g_1(\bar{x})$, then if $|x - \bar{x}| < \delta$, then $g(x) - g(\bar{x}) = g(x) - g_0(\bar{x}) < g_0(\bar{x}) + \epsilon - g_0(\bar{x}) = \epsilon$. Furthermore, $g(\bar{x}) - g(x) = g_0(\bar{x}) - g(x) < g_0(\bar{x}) - g_1(\bar{x}) + \epsilon < \epsilon$. Thus, $|g(x) - g(\bar{x})| < \epsilon$. The case where $g_1(\bar{x}) > g_0(\bar{x})$ is essentially identical.
Now suppose $g_0(\bar{x}) = g_1(\bar{x})$. Thus, $|x - \bar{x}| < \delta \implies |g(x) - g(\bar{x})|$, which is either $|g_0(x) - g_0(\bar{x})|$ or $|g_1(x) - g_1(\bar{x})|$, is $< \epsilon$. This proves that $g = \max\{g_0, g_1\}$ is continuous.
We may finish the proof by replacing g_0, g_1 with the continuous function $\max\{g_0, g_1\}$, taking the max of this and g_2 to show $\max\{g_0, g_1, g_2\}$ is continuous, and etc. $\square$

2 (Failure of Karush-Kuhn-Tucker). Consider the following problem:

$$\begin{array}{ll} \inf & (x_1 + 1)^2 + x_2^2 \\ \text{subject to} & -x_1^3 + x_2^2 \leq 0 \\ & x \in \mathbb{R}^2. \end{array}$$

(a) Sketch the feasible region and hence solve the problem.



The feasible region is shaded in blue, and the objective function can be written as $\|x - [-1, 0]\|^2$, i.e. the distance to the point $[-1, 0]$ shown as a red dot. Clearly, the minimizer over the feasible region is $x^* = [0, 0]$.

(b) Find multipliers λ_0 and λ satisfying the Fritz John conditions (2.11).
Seeing as the only inequality constraint is tight at $x^* = 0$, the Fritz John conditions say that there exist $\lambda_0, \lambda \geq 0$ not all 0 such that

$$\lambda_0 \begin{bmatrix} 2x_1 + 2 \\ 2x_2 \end{bmatrix} + \lambda \begin{bmatrix} -3x_1^2 \\ 2x_2 \end{bmatrix} = 0$$

with $x = 0$ plugged in. Indeed, we may take $\lambda_0 = 0$ and any $\lambda > 0$ because the second term is 0.

(c) Prove there exists no Lagrange multiplier vector for the optimal solution. Explain why not.

More explicitly plugging in $x^* = 0$ to the expression for the gradient of the Lagrangian,

$$\begin{bmatrix} 2\lambda_0 \\ 0 \end{bmatrix} = 0.$$

Therefore, $\lambda_0 = 0$. Clearly, we cannot have $\lambda_0 = 1$, as required for a Lagrange multiplier vector. This can be explained by the fact that the Mangasarian-Fromovitz constraint qualification does not hold: for any $d \in \mathbb{E}$, $\langle \nabla g(x^*), d \rangle = 0$, meaning there cannot be $d \in \mathbb{E}$ such that $\langle \nabla g(x^*), d \rangle < 0$.

3 (Linear independence implies Mangasarian-Fromovitz). If the set of vectors $\{a^1, a^2, \ldots, a^m\}$ in $\mathbb{E}$ is linearly independent, prove directly there exists a direction $d \in \mathbb{E}$ satisfying $\langle a^i, d \rangle < 0$ for $i = 1, 2, \ldots, m$.

Proof. If the set of vectors is linearly independent, then by definition there do not exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ not all 0 such that

$$\sum_{i=1}^{m} \lambda_i a^i = 0,$$

let alone $\lambda_1, \ldots, \lambda_m \geq 0$ summing to 1. Therefore, the first system of Gordan's theorem is unsolvable. Therefore, the second system is, i.e. $\exists d \in \mathbb{E}$ such that $\langle a^i, d \rangle < 0$ for all $i = 1, \ldots, m$. □

4. For each of the following problems, explain why there must exist an optimal solution, and find it by using the Karush-Kuhn-Tucker conditions.

(a)

$$\begin{aligned} \inf \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & -2x_1 - x_2 + 19 \leq 0 \\ & -x_1 \leq 0. \end{aligned}$$

The optimization problem can be rephrased as the distance of the nearest point in the intersection of the closed halfspaces $\{x \in \mathbb{R}^2 : -2x_1 - x_2 + 19 \leq 0\}$ and $\{x \in \mathbb{R}^2 : -x_1 \leq 0\}$ to 0. The intersection is convex, closed, and nonempty ($[10, 0]$ is inside). Therefore, there exists a unique optimal solution by Exercise 8 of Section 2.1.
The KKT conditions (note that they apply here for any feasible point because $[1, 0]$ achieves a negative product with the gradients of both inequality constraints) imply for an optimal x, there exists $\lambda_1, \lambda_2 \geq 0$ such that

$$2\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_1 - 2\lambda_1 - \lambda_2 \\ 2x_2 - \lambda_1 \end{bmatrix} = 0. \tag{2.6.1}$$

Furthermore, by complementary slackness,

$$\begin{aligned} &\lambda_1(-2x_1 - x_2 + 19) = 0 \\ &-\lambda_2 x_1 = 0. \end{aligned}$$

Notice that $x_1 = 0$ implies that $x_2 \geq 19$. Furthermore, $x_1 = 0$ implies that $\lambda_1 = \lambda_2 = 0$ by the first row of (2.6.1). This implies $2x_2 = 0$, a contradiction. Therefore, $\lambda_2 = 0$

to satisfy complementary slackness. By (2.6.1) this implies $x_1 = \lambda_1$ and $\lambda_1 = 2x_2$, i.e. $x_1 = 2x_2$. We can't have $\lambda_1 = 0$, because then $x_1 = 0$ which we ruled out. Therefore, by complementary slackness $-2x_1 - x_2 + 19 = -5x_2 + 19 = 0$, i.e. $x_2 = 19/5$ and $x_1 = 38/5$.

(b)

$$\begin{array}{ll} \inf & 5x_1^2 + 6x_2^2 \\ \text{subject to} & x_1 - 4 \le 0 \\ & 25 - x_1^2 - x_2^2 \le 0. \end{array}$$

Since the inequality constraint functions are continuous, the feasible region is closed. Because the objective function has bounded sublevel sets, it has a minimizer over the feasible region, i.e. an optimal solution (note the feasible region is nonempty because $[-5, 0]$ is in it).
If $\lambda_1, \lambda_2 \in \mathbb{R}_+$ compose a Lagrange multiplier vector for a feasible point x, then

$$\begin{bmatrix} 10x_1 \\ 12x_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2\lambda_2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10x_1 + \lambda_1 - 2\lambda_2 x_1 \\ 12x_2 - 2\lambda_2 x_2 \end{bmatrix} = 0. \tag{2.6.2}$$

Furthermore, by complementary slackness,

$$\lambda_1(x_1 - 4) = 0$$
$$\lambda_2(25 - x_1^2 - x_2^2) = 0.$$

If $\lambda_1 = \lambda_2 = 0$, then by (2.6.2) implies that $[x_1, x_2] = 0$, a contradiction because this point is infeasible. Therefore, at least one constraint is tight. Suppose the first one is, i.e. $x_1 = 4$. Then, the first row of (2.6.2) implies $40 + \lambda_1 - 8\lambda_2 = 0$. The second equation (note $|x_2| \ge 3$, because $x_2^2 \ge 25 - x_1^2 = 9$) implies $\lambda_2 = 6$. However, this implies $\lambda_1 = 8$. We also have $x_1^2 + x_2^2 = 25$ because $\lambda_2 \ne 0$, and complementary slackness. Therefore $x_2 = \pm 3$. Thus, assuming $x_1 = 4$, the only minimizers are $[4, \pm 3]$.
If $x_1 < 4$, then $\lambda_1 = 0$ by complementary slackness. This implies $10x_1 - 2\lambda_2 x_1 = 0$. If $x_1 = 0$, then $x_2 \ne 0$ and thus by $12x_2 - 2\lambda_2 x_2 = 0$, we have $\lambda_2 = 6$. Then $x_2 = \pm 5$. If $x_1 \ne 0$, then $\lambda_2 = 5$, implying $x_2 = 0$, which implies $x_1 = -5$.
Thus, the only candidate optimal solutions are $[4, \pm 3]$, $[0, \pm 5]$, and $[-5, 0]$. They evaluate to 134, 150, and 125 respectively. Therefore, the optimal solution is $[-5, 0]$ with objective value 125. (Would have been easier to note $5x_1^2 + 6x_2^2 \ge 5(x_1^2 + x_2^2) \ge 5(25) = 125$.)

5 (Cauchy-Schwarz and steepest descent). For a nonzero vector y in $\mathbb{E}$, use the Karush-Kuhn-Tucker conditions to solve the problem

$$\inf\{\langle y, x\rangle : \|x\|^2 \le 1\}.$$

Note that any nonzero x satisfies the Mangasarian-Fromovitz constraint qualification condition, and there clearly exists an x for which the objective is negative. Therefore, any optimal

point must satisfy the KKT conditions.
The KKT conditions imply that

$$\begin{aligned} y + 2\lambda x &= 0 \\ \lambda(\|x\|^2 - 1) &= 0. \end{aligned}$$

If $\lambda = 0$, then $y = 0$, a contradiction since y was assumed nonzero. Thus, $\|x\|^2 = 1$ and $y = -2\lambda x$. $\|y\| = 2\lambda\|x\| = 2\lambda$ implies $\lambda = \|y\|/2$. This implies $x = -y/\|y\|$. This is thus the (unique) optimal solution, with objective value $-\|y\|^2$.

6 * (Hölder's inequality). For real $p > 1$, define q by $p^{-1} + q^{-1} = 1$, and for $x \in \mathbb{R}^n$ define

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_1|^p\right)^{1/p}.$$

For a nonzero vector $y \in \mathbb{R}^n$, consider the optimization problem

$$\inf\{\langle y, x\rangle : \|x\|_p^p \le 1\}. \tag{2.6.3}$$

(a) Prove $\frac{d}{du}|u|^p/p = u|u|^{p-2}$ for all real u.
If $u > 0$, then $\frac{d}{du}|u|^p/p = \frac{d}{du}u^p/p = u^{p-1} = u|u|^{p-2}$. If $u < 0$, then $\frac{d}{du}|u|^p/p = \frac{d}{du}(-u)^p/p = -(-u)^{p-1} = u|u|^{p-2}$.
To compute the derivative at 0, note $\lim_{u\to 0} |u|^p/pu = \lim_{u\to 0} \operatorname{sgn}(u)|u|^{p-1}/p = 0$ since $p > 1$.

(b) Prove reals u and v satisfy $v = u|u|^{p-2}$ if and only if $u = v|v|^{q-2}$.
If $u = 0$ then clearly $v = 0$. Now assume $u > 0$. Thus $v = u^{p-1}$. Thus $u = v^{1/(p-1)} = v^{q-1}$, since $1/q + 1/p = 1 \implies q = p/(p-1) = 1 + 1/(p-1)$. Now if $u < 0$, we have $|u| = |v|^{q-1}$. Therefore, $u = -|v|^{q-1} = -|v||v|^{q-2} = v|v|^{q-2}$, as $v = u|u|^{p-2}$ implies $v < 0$.

(c) Prove problem (2.6.3) has a nonzero optimal solution.
We have the minimization of a continuous function over a compact set: the constraint is equivalent to $\|x\|_p \le 1$, $\|\cdot\|_p$ is continuous, and satisfies $\|x\|_\infty > \tau \implies \|x\|_p > \tau$. Thus $\|x\|_p \le \tau \implies \|x\| \le \sqrt{n}\|x\|_\infty \le \sqrt{n}\tau$, i.e. bounded sublevel sets.
Furthermore, since y is nonzero, we can obtain a negative objective by setting one of the entries of x to 1 or -1 where y is nonzero.

(d) Use the Karush-Kuhn-Tucker conditions to find the unique optimal solution.
The gradient of the constraint is $px|x|^{p-2}$, where operations on the vector are done entrywise. If x is nonzero, then clearly we can find a $d \in \mathbb{E}$ such that $\langle px|x|^{p-2}, d\rangle < 0$, thus any optimal solution satisfies the KKT conditions.
The KKT conditions require $\lambda \ge 0$,

$$\begin{aligned} y + \lambda px|x|^{p-2} &= 0 \\ \lambda(\|x\|_p^p - 1) &= 0. \end{aligned}$$

If $\lambda = 0$ then we have $y = 0$. Thus, we must have $\|x\|_p^p = 1$ and $\lambda \neq 0$. We have $-x|x|^{p-2} = y/\lambda p$. By part (b), we have $x = -(y/\lambda p)|y/\lambda p|^{q-2}$.

$$\|x\|_p = \left\| \left| \frac{y}{\lambda p} \right|^{q-1} \right\|_p = \frac{\| |y|^{q-1} \|_p}{\lambda^{q-1} p^{q-1}} = 1 \implies \lambda^{q-1} = \frac{\| |y|^{q-1} \|_p}{p^{q-1}}.$$

Use the fact that $q - 1 = 1/(p-1)$:

$$\lambda^{1/(p-1)} = \frac{1}{p^{q-1}} \left(\sum_{i=1}^n |y_i|^{p/(p-1)} \right)^{1/p} = \frac{1}{p^{q-1}} \left(\sum_{i=1}^n |y_i|^q \right)^{1/p} = \frac{1}{p^{1/(p-1)}} \|y\|_q^{q/p}.$$

Use the fact that $(p-1)/p = 1 - 1/p = 1/q$ to raise both sides to the $(p-1)$ power:

$$\lambda = \frac{\|y\|_q}{p}, \implies x = -\frac{y}{\|y\|_q} \left| \frac{y}{\|y\|_q} \right|^{q-2}$$

(e) Deduce that any vectors x and y in $\mathbb{R}^n$ satisfy $\langle y, x \rangle \leq \|y\|_q \|x\|_p$.
$\sup \langle y, x \rangle : \|x\|_p \leq 1 = \inf \langle -y, x \rangle : \|x\|_p \leq 1$, implying by part 1 that $x = y/\|y\|_q |y/\|y\|_q|^{q-2}$ maximizes $\langle y, x \rangle$ over $\|x\|_p^p = 1$. Then for any nonzero x,

$$\begin{aligned} \langle y, x \rangle = \|x\|_p \langle y, x/\|x\|_p \rangle &\leq \|x\|_p \left\langle y, \frac{y}{\|y\|_q} \left| \frac{y}{\|y\|_q} \right|^{q-2} \right\rangle \\ = \|x\|_p \frac{(\sum_{i=1}^n |y|^q)}{\|y\|_q^{q-1}} &= \|x\|_p \frac{\|y\|_q^q}{\|y\|_q^{q-1}} = \|x\|_p \|y\|_q. \end{aligned}$$

7. * Consider a matrix $A \in \mathbb{S}_{++}^n$ and a real $b > 0$.

(a) Assuming the problem

$$\inf\{-\log \det X : \operatorname{Tr} AX \leq b, X \in \mathbb{S}_{++}^n\}$$

has a solution, find it.
Let us first compute the gradient of $\operatorname{Tr} AX$. Supposing that the gradient of a function $f(X)$ exists, the directional derivative of $f(AX)$ is

$$\lim_{t \downarrow 0} \frac{f(A(X+tY)) - f(AX)}{t} = f'(AX; AY) = \langle (\nabla f)(AX), AY \rangle = \langle A(\nabla f)(AX), Y \rangle,$$

using the fact that A is symmetric. Therefore, $\nabla f(AX) = A(\nabla f)(AX)$. Thus, the gradient of the constraint is A. This implies that an optimal solution must satisfy the KKT conditions, because clearly there exists $B \in \mathbb{S}^n$ such that $\langle A, B \rangle < 0$. The gradient of the objective is $-X^{-1}$. Now the KKT conditions state there exists $\lambda \geq 0$ such that

$$\begin{gathered} -X^{-1} + \lambda A = 0 \\ \lambda(\operatorname{Tr} AX - b) = 0. \end{gathered}$$

If $\lambda = 0$, we obtain a contradiction as X^{-1} cannot equal 0. Therefore, $\operatorname{Tr} AX - b = 0$. Then the first equation, after right multiplying by X and taking the trace, implies $\lambda = n/b$. Thus $X = \frac{b}{n}A^{-1}$.

(b) Repeat using the objective function $\operatorname{Tr} X^{-1}$.
Now the KKT conditions state

$$\begin{aligned} -X^{-2} + \lambda A &= 0 \\ \lambda(\operatorname{Tr} AX - b) &= 0. \end{aligned}$$

Again, we cannot have $\lambda = 0$ because $-X^{-2} \neq 0$. We have $X^{-2} = \lambda A$, i.e. $X^2 = A^{-1}/\lambda \implies X = \sqrt{A^{-1}/\lambda}$. Plugging this into $\operatorname{Tr} AX = b$, we get $\lambda^{-1/2} \operatorname{Tr} \sqrt{A} = b \implies \lambda = (\operatorname{Tr} \sqrt{A}/b)^2$. Plugging this into $X = \sqrt{A^{-1}/\lambda}$, we get $X = \frac{b\sqrt{A^{-1}}}{\operatorname{Tr} \sqrt{A}}$.

(c) Prove the problems in parts (a) and (b) have optimal solutions. (Hint: Section 1.2, Exercise 14.)
Section 1.2, Exercise 14 says that for any positive definite matrix C, the function $\langle C, X\rangle - \log\det X$ has compact sublevel sets. Notice that $\operatorname{Tr} AX \le b$ implies $\langle A, X\rangle - \log\det X \le b - \log\det X$. We have the feasible point $X = bI/n \operatorname{Tr} A$. Therefore, we can restrict the domain of the infimum to $\{X : \operatorname{Tr} AX \le b\} \cap \{X : \langle A, X\rangle - \log\det X \le b - \log\det(bI/n \operatorname{Tr} A)\}$. The first set is closed (in fact, compact) due to $\operatorname{Tr} AX$ being continuous and the second set is compact, so the domain is compact, implying the existence of a minimizer.
For (b), note that $\operatorname{Tr} X^{-1} \le \alpha$ implies that the minimum eigenvalue of X, $\lambda_n(X)$ is $\ge 1/\alpha$. Denoting $a > 0$ as the minimum eigenvalue of A, $\operatorname{Tr} AX \le b$ implies $\lambda_1(X) \le b/a$. Therefore, $\{X : \operatorname{Tr} X^{-1} \le \alpha\} \cap \{X : \operatorname{Tr} AX \le b\}$ is compact, because it is the intersection of two sets that are closed in $\mathbb{S}^n_{++}$, and the intersection is contained in $\operatorname{int} \mathbb{S}^n_{++}$ and bounded because $\frac{1}{\alpha} \le \lambda_n(X) \le \lambda_1(X) \le \frac{b}{a}$.

8 ** (Minimum volume ellipsoid).

(a) For a point $y \in \mathbb{R}^n$ and the function $g : \mathbb{S}^n \to \mathbb{R}$ defined by $g(X) = \|Xy\|^2$, prove $\nabla g(X) = Xyy^\top + yy^\top X$ for all matrices $X \in \mathbb{S}^n$.

$$\begin{aligned} &\lim_{t\downarrow 0} \frac{\|(X+tY)y\|^2 - \|Xy\|^2}{t} = \lim_{t\downarrow 0} \frac{y^\top (X+tY)^2 y - yX^2 y}{t} \\ &= \lim_{t\downarrow 0} \frac{y^\top X^2 y + ty^\top XYy + ty^\top YXy + t^2 y^\top Yy - yX^2 y}{t} \\ &= \operatorname{Tr} y^\top XYy + \operatorname{Tr} y^\top YXy = \operatorname{Tr} yy^\top XY + \operatorname{Tr} Xyy^\top Y = \langle yy^\top X + Xyy^\top, Y\rangle, \end{aligned}$$

after which it follows that $\nabla g(X) = Xyy^\top + yy^\top X$.

(b) Consider a set $\{y^1, y^2, \dots, y^m\} \subset \mathbb{R}^n$. Prove this set spans $\mathbb{R}^n$ if and only if the matrix $\sum_i y^i (y^i)^\top$ is positive definite.
If it does not span $\mathbb{R}^n$, then there exists nonzero $x \in \mathbb{R}^n$ orthogonal to $y^1, \dots, y^m$. Then $x^\top \left(\sum_i y^i (y^i)^\top\right) x = 0$.

If it does span $\mathbb{R}^n$, then suppose some $x \in \mathbb{R}^n$ satisfies $x^\top \left(\sum_i y^i(y^i)^\top\right) x = 0$. Since each $x^\top y^i(y^i)^\top x = (x^\top y^i)^2 \geq 0$, this implies that each term equals 0. In other words, x is orthogonal to $\{y^1, \ldots, y^m\}$. But this contradicts $\{y^1, \ldots, y^m\}$ spanning $\mathbb{R}^n$. (If $x = \sum_i c_i y^i$, then $\|x\|^2 = x^\top \sum_i c_i y^i = 0$, a contradiction.)

(c) Now suppose the vectors $y^1, \ldots, y^m$ span $\mathbb{R}^n$. Prove the problem

$$\begin{array}{ll} \inf & -\log\det X \\ \text{subject to} & \|Xy^i\|^2 - 1 \leq 0 \quad \forall i = 1, \ldots, m \\ & X \in \mathbb{S}^n_{++} \end{array}$$

has an optimal solution. (Hint: Use part (b) and Section 1.2, Exercise 14.)
Since $\mathrm{Tr}(\sum_i y^i(y^i)^\top)X - \log\det X$ has compact level sets, and we know that $\mathrm{Tr}(\sum_i y^i(y^i)^\top)X \leq n$ from the constraints, we can take the objective value at some feasible X, add it to n, and know that the optimal X^* is contained in the compact level set corresponding to that objective value. After intersecting this level set with the closed sets $\{X : \|Xy^i\|^2 \leq 1\}$, we obtain a compact subset of the feasible region containing X^*.

(d) Now suppose $\bar{X}$ is an optimal solution for the problem in part (c). (In this case the set $\{y \in \mathbb{R}^n : \|\bar{X}y\| \leq 1\}$ is a minimum volume ellipsoid centered at the origin containing the vectors $y^1, y^2, \ldots, y^m$.) Show the Mangasarian-Fromovitz constraint qualification holds at $\bar{X}$ by considering the direction $d = -\bar{X}$.
We need to show that for every $i = 1, \ldots, m$, $\langle \bar{X}yy^\top + yy^\top\bar{X}, -\bar{X}\rangle < 0$. Evaluating the inner product gives $-2y^\top\bar{X}^2y < 0$.

(e) Write down the Karush-Kuhn-Tucker conditions that $\bar{X}$ must satisfy.
The KKT conditions are:

$$-\bar{X}^{-1} + \sum_{i=1}^m \bar{X}y^i(y^i)^\top + \bar{y}^i(y^i)^\top X = 0$$
$$\forall i \in [m],\ \lambda_i(\|Xy^i\|^2 - 1) = 0.$$

(f) When $\{y^1, y^2, \ldots, y^m\}$ is the standard basis of $\mathbb{R}^n$, the optimal solution of the problem in part (c) is $\bar{X} = I$. Find the corresponding Lagrange multiplier vector.
We have

$$-I + 2\sum_{i=1}^n \lambda_i e_i e_i^\top = \sum_{i=1}^n (2\lambda_i - 1)e_i e_i^\top = 0$$

implying $\lambda_i = \frac{1}{2}$ for each $i = 1, \ldots, n$. Note each constraint $\|Xy^i\| = \|e_i\|^2 = 1$ is tight.

3 Chapter 3: Fenchel Duality

3.1 3.1 Subgradients and Convex Functions

Time to allow convex functions to be nondifferentiable, and take on the value $+\infty$.

- The *domain* of a function $f : \mathbb{E} \to (-\infty, +\infty]$ is the set $\operatorname{dom} f = \{x \in \mathbb{E} : x < +\infty\}$.
- A function is *proper* if its domain is nonempty.
- A convex function f is *sublinear* if

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) \;\text{ for all } x, y \in \mathbb{E},\; \lambda, \mu \in \mathbb{R}_+.$$

- If $f(\lambda x) = \lambda f(x)$ for all $x \in \mathbb{E}$ and $\lambda \in \mathbb{R}_+$, then f is *positively homogeneous.*
- If $f(x + y) \leq f(x) + f(y)$ for all $x, y \in \mathbb{E}$ then f is subadditive.
- If f is sublinear, then $f((\lambda - \mu)x) \leq \lambda f(x) + \mu f(-x)$ applied to $\lambda = 1$, $\mu = 2$ shows $-f(x) \leq f(-x)$ for all $x \in \mathbb{E}$.
- The *lineality space* of a sublinear function f is the set

$$\operatorname{lin} f = \{x \in \mathbb{E} \mid -f(x) = f(-x)\}.$$

- The *core* of a set C, $\operatorname{core} C$ is the set of $x \in C$ for each $d \in \mathbb{E}$, $x + td \in C$ for all small t.
- The set of subgradients is called the *subdifferential*, denoted by $\partial f(\bar{x})$, defining $\partial f(\bar{x}) = \emptyset$ for $\bar{x}$ not in $\operatorname{dom} f$. Its *domain* is defined $\operatorname{dom} \partial f = \{x \in \mathbb{E} \mid \partial f(x) \neq \emptyset\}$.
- f is *essentially strictly convex* if it is strictly convex on any convex subset of $\operatorname{dom} \partial f$.

Proposition 3.1 (3.1.1 (Sublinearity))**.** *A function $f : \mathbb{E} \to (-\infty, +\infty]$ is sublinear if and only if it is positively homogeneous and subadditive. For a sublinear function f, the lineality space* $\operatorname{lin} f$ *is the largest subspace of* $\mathbb{E}$ *on which f is linear.*

Proposition 3.2 (3.1.2 (Sublinearity of the directional derivative))**.** *If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex then, for any point $\bar{x} \in \operatorname{core} \operatorname{dom} f$, the directional derivative $f'(\bar{x}; \cdot)$ is everywhere finite and sublinear.*

Proposition 3.3 (3.1.5 (Subgradients at optimality))**.** *For any proper function $f : \mathbb{E} \to (-\infty, +\infty]$, the point $\bar{x}$ is a (global) minimizer of f if and only if the condition $0 \in \partial f(\bar{x})$ holds.*

Proposition 3.4 (3.1.6 (Subgradients and directional derivatives))**.** *If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex and the point $\bar{x}$ lies in $\operatorname{dom} f$, then an element ϕ of $\mathbb{E}$ is a subgradient of f at $\bar{x}$ if and only if it satisfies $\langle \phi, \cdot \rangle \leq f'(\bar{x}; \cdot)$.*

Lemma 3.5 (3.1.7)**.** *Suppose that the function $p : \mathbb{E} \to (-\infty, +\infty]$ is sublinear and that the point $\bar{x}$ lies in* $\operatorname{core}(\operatorname{dom} p)$*. Then the function $q(\cdot) = p'(\bar{x}; \cdot)$ satisfies the conditions*

(i) $q(\lambda \bar{x}) = \lambda p(\bar{x})$ *for all real* λ*,*

(ii) $q \leq p$*, and*

(iii) $\operatorname{lin} q \supset \operatorname{lin} p + \operatorname{Span}\{\bar{x}\}$.

$$+\infty > g(d;s) \geq g(d;t) \downarrow f'(\bar{x};d) \geq g(d;-s) > -\infty. \tag{3.1.1}$$

$$\langle \phi, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{E}. \tag{3.1.2}$$

Theorem 3.6 (3.1.8 (Max formula))**.** *If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex then any point $\bar{x}$ in* $\operatorname{core} \operatorname{dom} f$ *and any direction $d \in \mathbb{E}$ satisfy*

$$f'(\bar{x};d) = \max\{\langle \phi, d \rangle \mid \phi \in \partial f(\bar{x})\}. \tag{3.1.3}$$

In particular, the subdifferential $\partial f(\bar{x})$ is nonempty.

Corollary 3.7 (3.1.10 (Differentiability of convex functions))**.** *Suppose the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex and the point $\bar{x}$ lies in* $\operatorname{core} \operatorname{dom} f$*. Then f is Gateaux differentiable at $\bar{x}$ exactly when f has a unique subgradient at $\bar{x}$ (in which case this subgradient is the derivative).*

We say the convex function f is *essentially smooth* if it is Gateaux differentiable on $\operatorname{dom} \partial f$.

Theorem 3.8 (3.1.11 (Hessian characterization of convexity))**.** *Given an open convex set $S \subset \mathbb{R}^n$ suppose the continuous function $f : \operatorname{cl} S \to \mathbb{R}$ is twice continuously differentiable on S. Then f is convex if and only if its Hessian matrix is positive semidefinite everywhere on S.*

3.2 Exercises for 3.1

1. Prove Proposition 3.1.1 (Sublinearity).

Proof. Suppose f is sublinear. Then for any $x \in \mathbb{E}$, $\lambda > 0$, $f(\lambda x + 0x) \leq \lambda f(x) + 0f(x) = \lambda f(x)$, and moreover $f((\lambda x)/\lambda) \leq f(\lambda x)/\lambda$, implying $\lambda f(x) \leq f(\lambda x)$. Thus, $f(\lambda x) = \lambda f(x)$. If $\lambda = 0$, then $f(\lambda x) = f(\lambda x + \lambda x) \leq 2\lambda f(x) = 0$, and $f(\lambda x) = f(\lambda x + \lambda x) \leq 2f(\lambda x)$ implying $f(\lambda x) \geq 0$. Thus $f(\lambda x) = 0$. Subadditivity obviously follows from sublinearity (plug in $\lambda, \mu = 1$).

Now suppose f is subadditive and positively homogeneous. Then for any $\lambda, \mu \geq 0$ and $x, y \in \mathbb{E}$,

$$f(\lambda x + \mu y) \leq f(\lambda x) + f(\mu y) = \lambda f(x) + \mu f(y).$$

Let $\mathbb{Y}$ be a subspace on which f is linear. Then, clearly, $\mathbb{Y} \subset \operatorname{lin} f$, because $x \in \mathbb{Y}$ implies $-f(x) = f(-x)$.

Now let $\mu, \lambda \in \mathbb{R}$ and $x, y \in \operatorname{lin} f$ be arbitrary. Then by sublinearity and the definition of $\operatorname{lin} f$,

$$f(\mu x + \lambda y) \leq \operatorname{sgn}(\mu)\mu f(\operatorname{sgn}(\mu)x) + \operatorname{sgn}(\lambda)\lambda f(\operatorname{sgn}(\lambda)y) = \mu f(x) + \lambda f(y),$$

as we can pull out the sgns and cancel out the resulting sgn^2s. Furthermore, since sublinear functions satisfy $-f(z) \leq f(-z)$ for every $z \in \mathbb{E}$,

$$-f(\mu x + \lambda y) \leq f(-\mu x - \lambda y) \leq -\mu f(x) - \lambda f(y),$$

following the same argument as above. Thus, $f(\mu x + \lambda y) \geq \mu f(x) + \lambda f(y)$. Thus, $f(\mu x + \lambda y) = \mu f(x) + \lambda f(y)$. Note this shows that $\operatorname{lin} f$ is a subspace. Therefore, f is linear on $\operatorname{lin} f$, which is itself a linear subspace. $\square$

2 (Core versus interior). Consider the set in $\mathbb{R}^2$

$$D = \{(x, y) \mid y = 0 \text{ or } |y| \geq x^2\}.$$

Prove $0 \in \operatorname{core}(D) \setminus \operatorname{int}(D)$.

Proof. First we prove that 0 is not in $\operatorname{int}(D)$. For any $\epsilon > 0$, consider the ball of radius ϵ centered at 0, ϵB. Let $C = \sqrt{4 + \epsilon}$. Consider the element $x = [2\sqrt{\epsilon}, \epsilon]/C$. We have

$$\|x\|^2 = \frac{4\epsilon + \epsilon^2}{C^2} = \frac{4\epsilon + \epsilon^2}{4 + \epsilon} = \epsilon.$$

i.e. $x \in \epsilon B$. However, $x \notin D$, since $x_1^2 = 4\epsilon > \epsilon = |x_2|$. Since ϵ was arbitrary, we have $0 \notin \operatorname{int}(D)$.

Now we show that $0 \in \operatorname{core}(D)$. Let $d = [x, y]$. If $y = 0$, then $d \in D$, and in fact for every $t \in \mathbb{R}$, $td \in D$. If $x = 0$, then clearly $|ty| \geq 0 = x$ for every $t \in \mathbb{R}$, i.e. $td \in D$. If $y \neq 0, x \neq 0$, let $0 < t \leq |y/x^2|$. We have

$$\frac{(td)_1^2}{|(td)_2|} = \frac{t^2x^2}{|ty|} = \frac{tx^2}{|y|} \leq \frac{|y|}{|y|} = 1,$$

i.e. the ratio of the x value of td squared to the magnitude of the y value is ≤ 1. Thus, $td \in \mathbb{E}$. Having shown that for all $d \in \mathbb{E}$, there is an $s > 0$ such that $0 < t < s$ implies $td \in D$, we have shown $0 \in \operatorname{core}(D)$. $\square$

3. Prove the subdifferential is a closed convex set.

Proof. First we show $\partial f(\bar{x})$ is convex. Let $\phi_1, \phi_2 \in \partial f$. For any $\lambda \in [0,1]$ and $x \in \mathbb{E}$,

$$\begin{aligned}\langle \lambda\phi_1 + (1-\lambda)\phi_2, x - \bar{x}\rangle &= \lambda\langle \phi_1, x-\bar{x}\rangle + (1-\lambda)\langle \phi_2, x - \bar{x}\rangle \\ &\leq \lambda(f(x) - f(\bar{x})) + (1-\lambda)(f(x) - f(\bar{x})) = f(x) - f(\bar{x}).\end{aligned}$$

Now we show $\partial f(\bar{x})$ is closed. One approach is to note $\partial f(\bar{x}) = \bigcap_{x\in\mathbb{E}}\{\phi \in \mathbb{E} : \langle \phi, x - \bar{x}\rangle \leq f(x) - f(\bar{x})\}$, an intersection of closed sets. Another is more direct; if a sequence ϕ^i satisfies $\langle \phi^i, x - \bar{x}\rangle \leq 0$, then so does the limit ϕ, i.e. $\langle \phi, x - \bar{x}\rangle \leq 0$. □

4 (Subgradients and normal cones). If a point $\bar{x}$ lies in a set $C \subset \mathbb{E}$, prove $\partial\delta_C(\bar{x}) = N_C(\bar{x})$.

Proof. Given $x \in \mathbb{E}$, either $\delta_C(x) - \delta_C(\bar{x}) = 0$ or $\delta_C(x) - \delta_C(\bar{x}) = +\infty$ depending on whether $x \in C$ or $x \notin C$ respectively. Either way, $f(x) - f(\bar{x}) \geq 0$. For $d \in N_C(\bar{x})$, we have for any $x \in \mathbb{E}$

$$\langle d, x - \bar{x}\rangle \leq 0 \leq f(x) - f(\bar{x}).$$

Thus, $N_C(\bar{x}) \subset \partial\delta_C(\bar{x})$.
Now consider $\phi \in \partial\delta_C(\bar{x})$. For every $x \in C$,

$$\langle \phi, x - \bar{x}\rangle \leq \delta_C(x) - \delta_C(\bar{x}) = 0.$$

Therefore, $\partial\delta_C(\bar{x}) \subset N_C(\bar{x})$. □

5. Prove the following functions $x \in \mathbb{R} \mapsto f(x)$ are convex and calculate ∂f:

(a) $|x|$
Let $\lambda \in [0,1]$. $|\lambda x + (1-\lambda)y| \leq \lambda|x| + (1-\lambda)|y|$ by the triangle inequality. Thus $|x|$ is convex. For the subgradient $\partial f(\bar{x})$ at $\bar{x}$, we need $\phi \in \mathbb{R}$ such that $\phi(x - \bar{x}) \leq |x| - |\bar{x}|$ for every $x \in \mathbb{R}$. If $\bar{x} < 0$, and $\phi > -1$, $\bar{x}/2 - \bar{x} > 0$ and $(-1)(\bar{x}/2 - \bar{x}) = |\bar{x}/2| - |\bar{x}|$, implying $\phi(\bar{x}/2 - \bar{x}) > |\bar{x}/2| - |\bar{x}|$. If $\phi < -1$, then $(-1)(\bar{x} - 2\bar{x}) = |\bar{x}| - |2\bar{x}|$ and $\bar{x} - 2\bar{x} > 0$ implies $\phi(2\bar{x} - \bar{x}) > |2\bar{x}| - |\bar{x}|$. Finally, if $\phi = -1$, $\phi(x - \bar{x}) = \phi x - |\bar{x}| \leq |x| - |\bar{x}|$. Therefore, $\partial|\bar{x}| = \{-1\}$. The case for $x > 0$ can be handled similarly to show $\partial|x| = \{1\}$.
Now let us compute $\partial|0|$. If $|\phi| \leq 1$, then $\phi x \leq |\phi||x| \leq |x|$, so $\phi \in \partial|0|$. If $|\phi| > 1$, then $\phi^2 = |\phi||\phi| > |\phi|$. Thus, $\partial|0| = \{\phi \in \mathbb{R} : |\phi| \leq 1\}$.

(b) $\delta_{\mathbb{R}_+}$
Denote $f = \delta_{\mathbb{R}_+}$. Let $x, y \in \mathbb{R}$ and $0 < \lambda < 1$. If one of x or y is not in $\mathbb{R}_+$, then clearly

$$f(\lambda x + (1-\lambda)y) \leq +\infty = \lambda f(x) + (1-\lambda)f(y).$$

If both $x, y \in \mathbb{R}_+$, then by convexity of $\mathbb{R}_+$, $\lambda x + (1-\lambda)y \in \mathbb{R}_+$ and thus $f(\lambda x + (1-\lambda)y) = 0 = \lambda f(x) + (1-\lambda)f(y)$.
For $\bar{x} \notin \operatorname{dom} f = \mathbb{R}_+$, $\partial f(\bar{x}) = \emptyset$. For $\bar{x} \in \operatorname{dom} f$, by Exercise 4, $\partial f(\bar{x}) = N_{\mathbb{R}_+}(\bar{x})$. That is, if $\bar{x} > 0$, $\partial f(\bar{x}) = \{0\}$ and if $\bar{x} = 0$, $\partial f(\bar{x}) = (-\infty, 0]$.

(c) $-\sqrt{x}$ if $x \geq 0$, $+\infty$ otherwise.
We wish to show that for $\lambda \in [0,1]$ and $x, y \geq 0$, $\sqrt{\lambda x + (1-\lambda)y} \geq \lambda\sqrt{x} + (1-\lambda)\sqrt{y}$. Notice that

$$\begin{aligned}(\lambda\sqrt{x} + (1-\lambda)\sqrt{y})^2 &= \lambda^2 x + (1-\lambda)^2 y + 2\lambda(1-\lambda)\sqrt{xy} \\ &= \lambda^2 x + (1-\lambda)^2 y + \lambda(1-\lambda)(x + y - (\sqrt{x} - \sqrt{y})^2) \\ &= \lambda x + (1-\lambda)y - (1-\lambda)\lambda(\sqrt{x} - \sqrt{y})^2 \leq \lambda x + (1-\lambda)y.\end{aligned}$$

Taking the square root of both sides gives us the desired result. Now if either $x, y < 0$, clearly the convexity condition holds as we have $+\infty$ on the RHS.
Let us now compute the subgradient. At 0, if $\phi > -\infty$, then for small enough $x > 0$, $\phi x > -\sqrt{x}$, as is clearly the case when $\phi \geq 0$ and shown by $-\sqrt{x}/x = -1/\sqrt{x} \to -\infty$ as $x \to 0$ for $-\infty < \phi < 0$. Thus $\partial\sqrt{0} = \emptyset$. For $\bar{x} \geq 0$, we'll use calculus and Corollary 3.1.10 (3.7): as $\bar{x} \in \operatorname{int}(\operatorname{dom} -\sqrt{\cdot})$ and $-\sqrt{\cdot}$ is convex, the derivative $-1/2\sqrt{\bar{x}}$ is equal to the unique subgradient at $\bar{x}$. Thus $\partial f(\bar{x}) = \{-1/2\sqrt{\bar{x}}\}$.

(d)

$$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

If neither x nor y equals 0, the function agrees with $\delta_{(-\infty,0]}$, implying that the convexity condition holds (refer to part (b), having $\mathbb{R}_+$ being the choice of convex set doesn't matter). Now let $\lambda \in (0,1)$ and WLOG suppose $x = 0$. We do indeed have

$$f(\lambda x + (1-\lambda)y) = f((1-\lambda)y) \leq (1-\lambda)f(y) \leq \lambda f(x) + (1-\lambda)f(y),$$

where the first inequality holds because if $(1-\lambda)y > 0$ then $y > 0$, so $f((1-\lambda)y) = +\infty = (1-\lambda)f(y)$, and otherwise $f((1-\lambda)y) = 0 \leq (1-\lambda)f(y)$.
For $x < 0$, $\partial f(x) = \{0\}$ since if $\phi \neq 0$ we can choose a small shift in either direction from x to obtain a product $\phi(y - x) > 0$.
Now for $\partial f(0)$, consider $\phi \in \mathbb{R}$. For small enough $x > 0$, $\phi(-x) > -1 = f(x) - f(0)$, implying $\phi \notin \partial f(0)$. Thus, $\partial f(0) = \emptyset$.

6. Prove Proposition 3.1.6 (3.4) (Subgradients and directional derivatives).

Proof. Let f be a convex function and $\bar{x} \in \operatorname{dom} f$. When $\bar{x} + td \notin \operatorname{dom} f$ for every $t > 0$, $f(\bar{x}; d) = +\infty$. Now suppose that $\phi \in \mathbb{E}$ satisfies $\langle \phi, \cdot \rangle \leq f'(\bar{x}; \cdot)$. Then, for every $x \in \mathbb{E}$, we have $\langle \phi, x - \bar{x} \rangle \leq f'(\bar{x}, x - \bar{x}) \leq f(x) - f(\bar{x})$ (refer to section 2.1, exercise 7 for last inequality), implying ϕ is a subgradient.
Now suppose $\phi \in \mathbb{E}$. If $\langle \phi, d \rangle > f'(\bar{x}; d)$ for some $d \in \mathbb{E}$, then for $t > 0$ small enough,

$$\langle \phi, d \rangle > \frac{f(\bar{x} + td) - f(\bar{x})}{t} \implies \langle \phi, \bar{x} + td - \bar{x} \rangle > f(\bar{x} + td) - f(\bar{x}),$$

i.e. ϕ is not a subgradient. ∎

7. Prove Lemma 3.1.7 (3.5).

Proof. Let $\bar{x} \in \operatorname{core}(\operatorname{dom} p)$. Note by Proposition 3.1.2, $p'(\bar{x};\cdot) = q(\cdot)$ is finite and sublinear. Now let $\lambda \in \mathbb{R}$. We have

$$p'(\bar{x};\lambda\bar{x}) = \lim_{t\downarrow 0} \frac{p(\bar{x}+t\lambda\bar{x}) - p(\bar{x})}{t} = \lim_{t\downarrow 0} \frac{(1+t\lambda)p(\bar{x}) - p(\bar{x})}{t} = \lambda p(\bar{x}),$$

where we used positive homogeneity of p, the fact that $1 + t\lambda \geq 0$ for small enough t, and finiteness of $p(\bar{x})$. Thus, $q(\lambda\bar{x}) = \lambda p(\bar{x})$. This proves (i).
For (ii), let $d \in \mathbb{E}$.

$$q(d) = \lim_{t\downarrow 0} \frac{p(\bar{x}+td) - p(\bar{x})}{t} \leq \lim_{t\downarrow 0} \frac{p(\bar{x}) + tp(d)}{t} = p(d),$$

using sublinearity of p and finiteness of $p(\bar{x})$.
Using (i), $q(-\lambda\bar{x}) = -\lambda p(\bar{x}) = -q(\lambda\bar{x})$, implying $\operatorname{Span}\{\bar{x}\} \subset \operatorname{lin} q$. Now take $x \in \operatorname{lin} p$. We have $q(x) \leq p(x)$ by (ii), and by sublinearity of q (now using Prop 3.1.2 and the fact $\bar{x} \in \operatorname{core}(\operatorname{dom} p)$) $-q(x) \leq q(-x) \leq p(-x) = -p(x)$, implying $q(x) \geq p(x)$ and thus $q(x) = p(x)$. Therefore, $q(-x) = p(-x) = -p(x) = -q(x)$, i.e. $x \in \operatorname{lin} q$. Since $\operatorname{lin} q$ is a linear subspace by Prop. 3.1.1, $\operatorname{lin} q \supset \operatorname{lin} p + \operatorname{Span}\{\bar{x}\}$. □

8 (Subgradients of norm). Calculate $\partial\|\cdot\|$. Generalize your result to an arbitrary sublinear function.

Proof. Let $\phi \in \mathbb{E}$ and suppose $\bar{x} \in \mathbb{E}$ is nonzero. Notice that

$$\left\langle \phi, \frac{\|\bar{x}\|}{\|\phi\|}\phi - \bar{x} \right\rangle = \|\bar{x}\|\|\phi\| - \langle \phi, \bar{x}\rangle \geq 0 = \left\| \frac{\|\bar{x}\|\phi}{\|\phi\|} \right\| - \|\bar{x}\| = 0, \tag{3.2.1}$$

with equality if and only if ϕ is a positive multiple of $\bar{x}$. Therefore, $\partial\|\bar{x}\|$ can only contain nonnegative multiples of $\bar{x}$. Suppose $\phi = c\bar{x}$ for some $c \geq 0$. Then if $c > 1/\|\bar{x}\|$,

$$\langle c\bar{x}, 2\bar{x} - \bar{x}\rangle = c\|\bar{x}\|^2 > \|\bar{x}\| = \|2\bar{x}\| - \|\bar{x}\|.$$

If $c < 1/\|\bar{x}\|$, then

$$\langle c\bar{x}, 0 - \bar{x}\rangle = -c\|\bar{x}\|^2 > -\|\bar{x}\| = 0 - \|\bar{x}\|.$$

Therefore, $\phi \in \partial\|\bar{x}\|$ implies $c = 1/\|\bar{x}\|$. To see this is sufficient,

$$\left\langle \frac{\bar{x}}{\|\bar{x}\|}, x - \bar{x} \right\rangle = \frac{\langle \bar{x}, x\rangle}{\|\bar{x}\|} - \|\bar{x}\| \leq \|x\| - \|\bar{x}\|.$$

Thus, $\partial\|\bar{x}\| = \{\bar{x}/\|\bar{x}\|\}$.
If $\bar{x} = 0$, then for any $\phi \in \mathbb{E}$ with $\|\phi\| \leq 1$, we have

$$\langle \phi, x\rangle \leq \|x\|,$$

implying $\phi \in \partial\|0\|$. Now if $\|\phi\| > 1$, we see $\langle \phi, \phi\rangle = \|\phi\|^2 > \|\phi\|$, so that $\partial\|0\| = B$.

□

9 (Subgradients of maximum eigenvalue). Prove

$$\partial\lambda_1(0) = \{Y \in \mathbb{S}^n_+ \mid \operatorname{Tr} Y = 1\}.$$

Proof. We first show the reverse inclusion. Given $Y \in \mathbb{S}^n_+$ such that $\operatorname{Tr} Y = 1$, by Fan's inequality, for any $X \in \mathbb{S}^n$,

$$\langle Y, X\rangle \leq \lambda(Y)^\top \lambda(X) = \sum_{i=1}^n \lambda_i(Y)\lambda_i(X) \leq \sum_{i=1}^n \lambda_i(Y)\lambda_1(X) = \lambda_1(X),$$

using the fact that $\lambda_i(Y) \geq 0$ for every $i \in [n]$ and that $\sum_{i=1}^n \lambda_i(Y) = \operatorname{Tr} Y = 1$.
Now suppose that $\phi \in \partial\lambda_1(0)$. Let v be a unit eigenvector of ϕ with eigenvalue λ. Then,

$$\langle \phi, -vv^\top\rangle = -\lambda \leq 0 \implies \lambda \geq 0$$

as $\lambda_1(-vv^\top) = 0$. Thus, $\phi \in \mathbb{S}^n_+$. Moreover,

$$\operatorname{Tr}\phi = \langle \phi, I\rangle \leq \lambda_1(I) = 1,$$

showing the forward inclusion. This proves the desired statement. $\square$

10. ** For any vector μ in the cone $\mathbb{R}^n_\geq$, prove

$$\partial\langle \mu, [\cdot]\rangle(0) = \operatorname{conv}(\mathbb{P}^n\mu).$$

(See Section 2.2, Exercise 9).

Proof. We'll prove that $\phi \in \partial\langle \mu, [\cdot]\rangle(0)$ if and only if $\mu - [\phi] \in (\mathbb{R}_\geq)^+$. Suppose that $\mu - [\phi] \in (\mathbb{R}_\geq)^+$. Then for any $x \in \mathbb{R}^n$, we have

$$\langle \phi, x\rangle \leq \langle [\phi], [x]\rangle = \langle [\phi], [x]\rangle - \langle [\phi] - \mu, [x]\rangle + \langle \mu, [x]\rangle \leq \langle \mu, [x]\rangle$$

because $\langle [\phi] - \mu, [x]\rangle \leq 0$. Therefore, $\phi \in \partial\langle \mu, [\cdot]\rangle(0)$.
Now otherwise, there exists $x' \in \mathbb{R}^n_\geq$ where $\langle \mu - [\phi], x'\rangle < 0$. By picking $\tilde{x}$ whose entries are obtained by the appropriate permutation applied to the entries of x',

$$\langle \phi, \tilde{x}\rangle = \langle [\phi], x'\rangle = \langle [\phi] - \mu, x'\rangle + \langle \mu, x'\rangle > \langle \mu, x'\rangle = \langle \mu, [x']\rangle,$$

implying that $\phi \notin \partial\langle \mu, [\cdot]\rangle(0)$.
By section 2.2, Exercise 9 part (d), $\mu - [\phi] \in (\mathbb{R}^n_\geq)^+$ implies $[\phi] \in \operatorname{conv}(\mathbb{P}^n\mu)$. Since $\phi = P[\phi]$ for some $P \in \mathbb{P}^n$ and permutations are closed under composition, we have $\phi \in \operatorname{conv}(\mathbb{P}^n\mu)$.
Now suppose $\phi \in \operatorname{conv}(\mathbb{P}^n\mu)$. By part (c) of Exercise 9, $[\cdot]$ is $(\mathbb{R}^n_\geq)^+$ convex, meaning that for any $x^1, \ldots, x^m \in \mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_m \geq 0$ summing to 1,

$$\sum_{i=1}^m \lambda[x^i] - \left[\sum_{i=1}^m \lambda x^i\right] \in (\mathbb{R}^n_\geq)^+.$$

Since $\phi \in \operatorname{conv}(\mathbb{P}^n \mu)$, there exist λ_Π, $(\Pi \in \mathbb{P}^n) \geq 0$ summing to 1 such that $\phi = \sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi \Pi \mu$. Thus,

$$\sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi [\mu] - \left[\sum_{\Pi \in \mathbb{P}^n} \lambda_\Pi \mu \right] \in (\mathbb{R}^n_{\geq})^+.$$

The left term is μ and the right term is $[\phi]$, so we have $\mu - [\phi] \in (\mathbb{R}^n_{\geq})^+$.
We have proven that $\mu - [\phi] \in (\mathbb{R}^n_{\geq})^+$ if and only if $\phi \in \operatorname{conv}(\mathbb{P}^n \mu)$, which completes the proof. $\square$

11. * Define a function $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x_1, x_2, \ldots, x_n) = \max_j \{x_j\}$, let $\bar{x} = 0$ and $d = (1, 1, \ldots, 1)^\top$ and let $e_k = (1, 1, \ldots, 1, 0, \ldots, 0)^\top$ (ending in $k-1$ zeroes). Calculate the functions p_k defined in the proof of Theorem 3.1.8 (Max formula), using Proposition 2.3.2 (Directional derivatives of max functions).

Proof.

$$p_0(x) = f'(0; x) = \lim_{t \downarrow 0} \frac{t \max_j x_j}{t} = \max_j \{x_j\}.$$

Interestingly, $p_0 = f$. Now

$$p_1(x) = p_0'(d; x) = \lim_{t \downarrow 0} \frac{\max_j \{1 + t x_j\} - 1}{t} = \max_j \{x_j\}.$$

Yet again, $p_1 = f$. For p_2,

$$p_2(x) = p_1'(e_2, x) = \lim_{t \downarrow 0} \frac{\max_j \{\mathbb{1}[j \leq n-1] + t x_j\} - 1}{t} = \max_{1 \leq j \leq n-1} \{x_j\}.$$

For p_3,

$$p_3(x) = p_2'(e_3, x) = \lim_{t \downarrow 0} \frac{\max_{1 \leq j \leq n-1} \{\mathbb{1}\{j \leq n-2\} + t x_j\} - 1}{t} = \max_{1 \leq j \leq n-2} \{x_j\}.$$

Continuing in this fashion, we see that $p_k(x) = \max_{1 \leq j \leq n-k+1} \{x_j\}$. Finally, $p_1(x) = x_1 = \langle \mathbf{e}_1, x \rangle$, where $\mathbf{e}_1$ is the first coordinate vector. This is the vector constructed in Theorem 3.1.8 (3.6) so that $f'(0, d) = p_n(d)$. Indeed, $f'(0, d) = 1 = p_n(d)$. $\square$

12 * (Recognizing convex functions). Suppose the set $S \subset \mathbb{R}^n$ is open and convex, and consider a function $f : S \to \mathbb{R}$. For points $x \notin S$, define $f(x) = +\infty$.

(a) Prove $\partial f(x)$ is nonempty for all $x \in S$ if and only if f is convex. (Hint: For points $u, v \in S$ and $\lambda \in [0, 1]$, use the subgradient inequality (3.1.2) at the points $\bar{x} = \lambda u + (1-\lambda) v$ and $x = u, v$ to check the definition of convexity.)
First suppose that $\partial f(x)$ is not empty for all $x \in S$. Then take $u, v \in S$. Take any $\lambda \in [0, 1]$. By convexity of S, $z = \lambda u + (1-\lambda) v \in S$. Let $\phi \in \partial f(z)$. Then

$$(1-\lambda)\langle \phi, u - v \rangle = \langle \phi, u - z \rangle \leq f(u) - f(z)$$
$$\lambda \langle \phi, v - u \rangle = \langle \phi, v - z \rangle \leq f(v) - f(z).$$

Multiplying the first equation by λ and the second by $(1-\lambda)$ and adding them, we obtain $\lambda f(u)+(1-\lambda)f(v)\geq f(z)$, as desired. For $w\notin S$, it's clear that $f(\lambda u+(1-\lambda)w)\leq \lambda f(u)+(1-\lambda)f(w)$, because the RHS is $+\infty$ if $\lambda\neq 1$. Thus, f is convex.
Now suppose that f is convex. For any $x\in S$, we have $x\in \operatorname{core}(\operatorname{dom} f)$, so by Theorem 3.1.8 (3.6) the directional derivative is equal to $f'(x;\cdot)=\max\{\langle\phi,\cdot\rangle \mid \phi\in\partial f(x)\}$. In particular, $\partial f(x)$ is nonempty, for each $x\in S$.

(b) Prove that if $I\subset\mathbb{R}$ is an open interval and $g:I\to\mathbb{R}$ is differentiable then g is convex if and only if g' is nondecreasing on I, and g is strictly convex if and only if g' is strictly increasing on I. Deduce that if g is twice differentiable then g is convex if and only if g'' is nonnegative on I, and g is strictly convex if g'' is strictly positive on I.
Suppose g is convex. Let $x,y\in I$ with $x<y$. Let $x<t<y<s$, $t,s\in S$. We will prove that

$$\frac{f(t)-f(x)}{t-x}\leq\frac{f(s)-f(y)}{s-y} \tag{3.2.2}$$

with strict inequality if f is strictly convex. By convexity, we have the following two equations:

$$f(s)\frac{y-x}{s-x}+f(x)\frac{s-y}{s-x}\geq f(y) \tag{3.2.3}$$
$$f(s)\frac{t-x}{s-x}+f(x)\frac{s-t}{s-x}\geq f(t). \tag{3.2.4}$$

Rearranging (3.2.3) gives

$$-f(x)\frac{s-y}{s-x}\leq f(s)\frac{s-x+y-s}{s-x}-f(y) \implies \frac{f(s)-f(x)}{s-x}\leq\frac{f(s)-f(y)}{s-y}.$$

Rearranging (3.2.4) gives

$$f(t)\leq f(s)\frac{t-x}{s-x}+f(x)\frac{s-x+x-t}{s-x} \implies \frac{f(t)-f(x)}{t-x}\leq\frac{f(s)-f(x)}{s-x}.$$

Putting these inequalities together gives (3.2.2), and strict convexity makes the inequalities strict. Thus, by differentiability of f,

$$f'(x)=\lim_{t\downarrow 0}\frac{f(t)-f(x)}{t-x}\leq\lim_{s\downarrow 0}\frac{f(s)-f(y)}{s-y}=f'(y).$$

If we have strict convexity, then we can choose some interval (t',y') where $x<t'<y'<y$ and take $x<t<t'$. Applying (3.2.2) to the pair (y',t') as well, we have

$$f'(x)=\lim_{t\downarrow 0}\frac{f(t)-f(x)}{t-x}<\frac{f(y')-f(t')}{y'-t'}<\lim_{s\downarrow 0}\frac{f(s)-f(y)}{s-y}=f'(y).$$

We have proven that if f is (strictly) convex then f' is (strictly) increasing.
Now suppose that f' is increasing. Let $x,y\in I$ with $x<y$. Let $0<\lambda<1$ and

define $z = (1-\lambda)x + \lambda y$. Note that $f(z) = f(x) + f(z) - f(x)$, so to show $f(z) \le (1-\lambda)f(x) + \lambda f(y)$, we may show $f(z) - f(x) \le \lambda(f(y) - f(x))$.

$$\begin{aligned}
f(z) - f(x) &= \int_x^z f'(x')\,\mathrm{d}x' = \lambda \frac{1}{\lambda}\int_x^z f'(x')\,\mathrm{d}x' = \lambda\left[\int_x^z f'(x')\,\mathrm{d}x' + \left(\frac{1}{\lambda} - 1\right)\int_x^z f'(x')\,\mathrm{d}x'\right] \\
&\le \lambda\left[\int_x^z f'(x')\,\mathrm{d}x' + \left(\frac{1}{\lambda} - 1\right)\int_x^z f'(z)\,\mathrm{d}x'\right] = \lambda\left[\int_x^z f'(x')\,\mathrm{d}x' + \left(\frac{1}{\lambda} - 1\right)\lambda(y-x)f'(z)\right] \\
&= \lambda\left[\int_x^z f'(x')\,\mathrm{d}x' + \int_z^y f'(z)\,\mathrm{d}x'\right] \le \lambda\left[\int_x^z f'(x')\,\mathrm{d}x' + \int_z^y f'(x')\,\mathrm{d}x'\right] = \lambda(f(y) - f(x)).
\end{aligned}$$

If f' is strictly increasing, then we have strict inequality since $\int_z^y f'(z)\,\mathrm{d}x' < \int_z^y f'(x')\,\mathrm{d}x'$. (What's a good justification for this? XD) This completes the proof that f' is (strictly) increasing iff f is (strictly) convex.

Now we show that f'' is nonnegative if and only if f' is increasing. Suppose that f' is increasing. Then, for any $x \in I$, $f''(x) = \lim_{t \downarrow x} \frac{f(t) - f(x)}{t - x} \ge 0$, since $f(t) \ge f(x)$ for $t \ge x$. Furthermore, if $f'' \ge 0$, then $f'(t) = \int_x^t f''(x')\,\mathrm{d}x' + f'(x) \ge f'(x)$, with strict inequality if $f'' > 0$. Therefore, f' is (strictly) increasing if f'' is (strictly) positive.

This suffices to show that f is convex iff f'' is nonnegative, and strictly convex if f'' is strictly positive.

(c) Deduce that if f is twice continuously differentiable on S then f is convex if and only if its Hessian matrix is positive semidefinite everywhere on S, and f is strictly convex if its Hessian matrix is positive definite everywhere on S. (Hint: Apply part (b) to the function g defined by $g(t) = f(x + td)$ for small real t, points $x \in S$, and directions $d \in \mathbb{E}$.)

First let's prove that

$$f''(x + td) := \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(x + td) = d^\top \nabla^2 f(x + td) d. \tag{3.2.5}$$

Here is an approach that is easy, but opaque due to utilizing the multivariable chain rule (i.e., literally following multivariable calculus rules).

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}}{\mathrm{d}t} f(x + td)\right) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}f(x+td)}{\mathrm{d}(x+td)}\frac{\mathrm{d}(x+td)}{\mathrm{d}t}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\nabla f(x+td)^\top d \\
&= d^\top \frac{\mathrm{d}}{\mathrm{d}t}\nabla f(x+td) = d^\top\left(\frac{\mathrm{d}\nabla f(x+td)}{\mathrm{d}(x+td)}\frac{\mathrm{d}(x+td)}{\mathrm{d}t}\right) \\
&= d^\top \nabla^2 f(x+td) d.
\end{aligned}$$

If we want to stick more to single variable calculus to "see under the hood," at least for the calculation of the second derivative (assume $f'(x+td) = \langle \nabla f(x+td), d\rangle$),

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \nabla f(x+td), d\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^n \nabla f(x+td)_i d_i\right) = \sum_{i=1}^n d_i \frac{\mathrm{d}}{\mathrm{d}t}\nabla f(x+td)_i.$$

This roughly says that the change in $\langle \nabla f(x+td), d\rangle$ is the sum of the changes in each coordinate of $\nabla f(x+td)$, weighted by the corresponding coordinate of d_i.

$$\frac{\mathrm{d}}{\mathrm{d}t}\nabla f(x+td)_i = \sum_{j=1}^{n} \frac{\mathrm{d}\nabla f(x+td)_i}{\mathrm{d}(x+td)_j}\frac{\mathrm{d}(x+td)_j}{\mathrm{d}t} = \sum_{j=1}^{n} \nabla^2 f(x+td)_{ij} d_j.$$
$$\implies \frac{\mathrm{d}}{\mathrm{d}t}\langle \nabla f(x+td), d\rangle = \sum_{i=1}^{n} d_i \sum_{j=1}^{n} \nabla^2 f(x+td)_{ij} d_j = d^\top \nabla^2 f(x+td)_{ij} d.$$

Now let us use (3.2.5) to prove the statement. Given $x \in S$ and $d \in \mathbb{E}$, define $g_{x,d} : I \to \mathbb{R}$, $g_{x,d}(t) = f(x+td)$ where I is the nonempty open interval containing 0 such that $x + td \in S$. If f is convex, then $g_{x,d}$ is convex for every $x \in S, d \in \mathbb{E}$, so by part (b),

$$\forall x \in S,\ d \in \mathbb{E}\ :\ g''_{x,d}(0) = d^\top \nabla^2 f(x) d \geq 0.$$

This implies that $\nabla^2 f$ is everywhere PSD. Now if $\nabla^2 f$ is everywhere PSD, then given $x, y \in S$, $g''_{x,y-x} \geq 0$, implying $g''_{x,y-x}$ is convex. Thus, for any $\lambda \in [0,1]$,

$$g(\lambda) = f((1-\lambda)x + \lambda y) \leq (1-\lambda)g(0) + \lambda g(1) = (1-\lambda)f(x) + \lambda f(y),$$

proving that f is convex. If $\nabla^2 f$ is everywhere PD, then we have $g''_{x,y-x}$ is strictly convex for any $x, y \in S$, implying that f is strictly convex.

(d) Find a strictly convex function $f : (-1,1) \to \mathbb{R}$ with $f''(0) = 0$.
The function x^4 is strictly convex, because its derivative is $4x^3$ which is strictly increasing. On the other hand, its second derivative, $12x^2$, at 0 is 0.

(e) Prove that a continuous function $h : \mathrm{cl}\, S \to \mathbb{R}$ is convex if and only if its restriction to S is convex. What about strictly convex functions?
First suppose that h is convex on S. Then take $x, y \in \mathrm{cl}\, S$, so there are sequences $(x^i)_i \to x$ and $(y^i)_i \to y$ in S. Let $\lambda \in [0,1]$. Thus, $(\lambda x^i + (1-\lambda)y^i)_i \to \lambda x + (1-\lambda)y$ is a sequence in S. Thus, using continuity and convexity of h on S,

$$h(\lambda x+(1-\lambda)y) = \lim_{i\to\infty} h(\lambda x^i+(1-\lambda)y^i) \leq \lim_{i\to\infty} \lambda h(x^i)+(1-\lambda)h(y^i) = \lambda h(x)+(1-\lambda)h(y).$$

In other words, h is convex on $\mathrm{cl}\, S$.
Now if h is (strictly) convex on $\mathrm{cl}\, S$, clearly its restriction to S is (strictly) convex.
What if h is strictly convex on S? Does this mean h is strictly convex on $\mathrm{cl}\, S$? Consider two points $x, y \in \mathrm{cl}\, S$ where $x \neq y$. Define the function $g : [0,1] \to \mathbb{R}$, $g(t) = h((1-t)x + ty)$. If $\{(1-t)x + ty : t \in (0,1)\} \subset S$, then given $0 < t < 1$, take any $0 < s < t < u < 1$. By convexity of h on $\mathrm{cl}\, S$ (as proven earlier), $g(s) \leq (1-s)g(0) + sg(1)$ and $g(u) \leq (1-u)g(0) + ug(1)$. By strict convexity of h on S, g is strictly convex on $(0,1)$, so

$$\begin{aligned} h((1-t)x+ty) = g(t) &< \frac{u-t}{u-s}g(s) + \frac{t-s}{u-s}g(u) \\ &\leq \frac{u-t}{u-s}\left((1-s)g(0)+sg(1)\right) + \frac{t-s}{u-s}\left((1-u)g(0)+ug(1)\right) \\ &= (1-t)g(0) + tg(1) = (1-t)h(x) + th(y). \end{aligned}$$

Therefore, if the open line segment $\{(1-t)x+ty : t \in (0,1)\}$ is in S, then h is strictly convex along this line. In fact, by the accessibility lemma (Section 1.1, Exercise 11), if any z in the line segment is in $\operatorname{int} S$, then the entire open line segment is in $\operatorname{int} S$, since the lemma says $\lambda \operatorname{int} S + (1-\lambda)\operatorname{cl} S \subset \operatorname{int} S$ for any $0 < \lambda \leq 1$.

Therefore, to produce an example where h is strictly convex on S but not on $\operatorname{cl} S$, we need a set S such that there is a line segment completely contained in the boundary $\operatorname{bd} S$, and show that h is not strictly convex on this line.

Take the function $f(x,y) = y^4/x$. The Hessian is $\begin{bmatrix} 2y^4/x^3 & -4y^3/x^2 \\ -4y^3/x^2 & 12y^2/x \end{bmatrix}$, with determinant $8y^6/x^4$. This is strictly positive on $(0,1)^2$, indicating strict convexity. However, it is linear (in fact, constant) on the line $y = 0$.

13 (Local convexity). Suppose the function $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable near 0 and $\nabla^2 f(0)$ is positive definite. Prove $f_{|\delta B}$ is convex for some real $\delta > 0$.

Proof. By Exercise 1 from Section 1.2, the set of positive definite matrices $\mathbb{S}^n_{++}$ is the interior of the set of positive semidefinite matrices $\mathbb{S}^n_+$. If $\nabla^2 f(0) \in \operatorname{int} S^n_{++}$, then there exists $\epsilon > 0$ where $\nabla^2 f(0) + \epsilon B \subset \operatorname{int} S^n_{++}$. By continuity of the second derivative at 0, there exists $\delta > 0$ such that $x \in \delta B \implies \nabla^2 f(x) \in \nabla^2 f(0) + \epsilon B \subset \mathbb{S}^n_{++}$. In other words, $\nabla^2 f_{|\delta B}$ is PD, implying that $f_{|\delta B}$ is strictly convex. $\square$

14 (Examples of convex functions). As we shall see in Section 4.2, most natural convex functions occur in pairs. The table in Section 3.3 lists many examples on $\mathbb{R}$. Use Exercise 12 to prove each function f and f^* in the table is convex. Ordered pairs $(f, \operatorname{dom} f)$ describe a function and its domain, and $(f, \operatorname{dom} f) \leftrightarrow (g, \operatorname{dom} g)$ means that $(f, \operatorname{dom} f)$ and $(g, \operatorname{dom} g)$ are conjugate pairs. Also, let D denote the derivative operator.

1. $(0, \mathbb{R}) \leftrightarrow (0, \{0\})$. Obviously, 0 is convex on any convex domain. It's a constant linear affine function (any of these three descriptors implies convexity).

2. $(0, \mathbb{R}_+) \leftrightarrow (0, -\mathbb{R}_+)$. See above.

3. $(0, [-1,1]) \leftrightarrow (|y|, \mathbb{R})$. Being a norm, $|\cdot|$ is convex.

4. $(|x|^p/p, \mathbb{R}) \leftrightarrow (|y|^q/q, \mathbb{R})$ $(p > 1, \frac{1}{p} + \frac{1}{q} = 1)$. $D(|x|^p/p) = x|x|^{p-2} = \operatorname{sgn}(x)|x|^{p-1}$ (see Section 2.3 Exercise 6 (a)). This function is strictly increasing, so $|x|^p/p$ is strictly convex. By the same token, $|y|^q/q$ is strictly convex ($1/p + 1/q = 1$ and $p > 1$ implies $q > 1$).

5. $(|x|^p/p, \mathbb{R}_+) \leftrightarrow (|y^+|^q/q, \mathbb{R})$, $(p > 1, \frac{1}{p} + \frac{1}{q} = 1)$. See above, and note that the derivative of $|y^+|^q/q$ when $y < 0$ becomes 0.

6. $(-x^p/p, \mathbb{R}_+) \leftrightarrow (-(-y)^q/q, -\mathbb{R}_{++})$, $(p \in (0,1), \frac{1}{p} + \frac{1}{q} = 1)$. $D(-x^p/p) = -x^{p-1} = -1/x^{|p-1|}$. $x^{|p-1|}$ is strictly increasing on $\mathbb{R}_{++}$, so $1/x^{|p-1|}$ is strictly decreasing on $\mathbb{R}_{++}$, so $-1/x^{|p-1|}$ is strictly increasing on $\mathbb{R}_{++}$. By Exercise 12, $-x^p/p$ is strictly convex on $\mathbb{R}_{++}$. By the comment at the end of 12 (d), $-x^p/p$ is strictly convex on $\mathbb{R}_+$. $D(-(-y)^q/q) = (-y)^{q-1}$. Since $1/p + 1/q = 1$ and $p \in (0,1)$, $q < 0$. Thus, $q - 1 < 0$. Thus, on $-\mathbb{R}_{++}$, $(-y)^{q-1} = 1/|y|^{|q-1|}$ is strictly increasing. Thus, $-(-y)^q/q$ on $\mathbb{R}_{++}$ is strictly convex.

7. $(\sqrt{1+x^2}, \mathbb{R}) \leftrightarrow (-\sqrt{1-y^2}, [-1,1])$. $D(\sqrt{1+x^2}) = x/\sqrt{1+x^2}$ can be shown to be strictly increasing, but let's make our lives easier by taking another derivative.

$$D^2(\sqrt{1+x^2}) = \frac{\sqrt{1+x^2} - \frac{x^2}{\sqrt{1+x^2}}}{1+x^2} = \frac{1}{(1+x^2)^{3/2}} > 0,$$

implying by Exercise 12 that $\sqrt{1+x^2}$ is strictly convex.
$D(-\sqrt{1-y^2}) = y/\sqrt{1-y^2}$.

$$D^2(-\sqrt{1-y^2}) = \frac{\sqrt{1-y^2} + \frac{y^2}{\sqrt{1-y^2}}}{1-y^2} = \frac{1}{(1-y^2)^{3/2}}.$$

The second derivative is positive on $(-1,1)$, implying strict convexity of $-\sqrt{1-y^2}$ on $(-1,1)$. By the comment at the end of Exercise 12 (d), this implies strict convexity on $[-1,1]$.

8. $(-\log x, \mathbb{R}_{++}) \leftrightarrow (-1-\log(-y), -\mathbb{R}_{++})$. $D(-\log x) = -1/x$, which is strictly increasing on $\mathbb{R}_{++}$, so $-\log x$ is strictly convex.
$D(-1-\log(-y)) = -1/y$, which is strictly increasing on $-\mathbb{R}_{++}$. Thus $-1-\log(-y)$ is strictly convex.

9. $(\cosh x, \mathbb{R}) \leftrightarrow (y\sinh^{-1}(y) - \sqrt{1+y^2}, \mathbb{R})$. $D^2(\cosh x) = \cosh x = (e^x + e^{-x})/2 > 0$. Therefore, $\cosh x$ is strictly convex.
Let us first compute $\sinh^{-1} y$. We have

$$\begin{aligned} y = \frac{e^x - e^{-x}}{2} &\implies 2y = e^x - e^{-x} \implies e^{2x} - 2ye^x - 1 = 0 \\ &\implies e^x \in \left\{y \pm \sqrt{y^2+1}\right\} \implies e^x = y + \sqrt{y^2+1} \\ &\implies \sinh^{-1}(y) = x = \log(y + \sqrt{y^2+1}), \end{aligned}$$

where we used the fact that $e^x > 0$ whereas $y - \sqrt{y^2+1} < 0$.
Now let us show that $y\sinh^{-1}(y) - \sqrt{1+y^2} = y\log(y+\sqrt{y^2+1}) - \sqrt{1+y^2}$ is convex.

$$\begin{aligned} &D(y\log(y+\sqrt{y^2+1}) - \sqrt{1+y^2}) \\ &= \log(y+\sqrt{y^2+1}) + \frac{y\left(1+\frac{y}{\sqrt{y^2+1}}\right)}{y+\sqrt{y^2+1}} - \frac{y}{\sqrt{y^2+1}} \\ &= \log(y+\sqrt{y^2+1}) + \frac{y}{\sqrt{y^2+1}} - \frac{y}{\sqrt{y^2+1}} \\ &= \log(y+\sqrt{y^2+1}). \end{aligned}$$

Clearly, $\log(y+\sqrt{y^2+1})$ is increasing on $\mathbb{R}_+$. Since

$$D\left(\sqrt{y^2+1} - y\right) = \frac{y}{\sqrt{y^2+1}} - 1 < 0,$$

on $-\mathbb{R}_+$, $\log(y+\sqrt{y^2+1}) = \log(\sqrt{z^2+1}-z)$ where $z=-y$. $\log(\sqrt{z^2+1}-z)$ is decreasing in z, therefore it is increasing in y. Thus $y\sinh^{-1}(y)-\sqrt{1+y^2}$ is strictly convex.

10. $(-\log(\cos x),(-\frac{\pi}{2},\frac{\pi}{2})) \leftrightarrow (y\tan^{-1}(y)-\frac{1}{2}\log(1+y^2),\mathbb{R})$. $D(-\log(\cos x)) = \tan x$ is strictly increasing on $(-\pi/2,\pi/2)$. Therefore, $-\log\cos x$ is strictly convex on $(-\pi/2,\pi/2)$. Note $D(\tan^{-1}(y)) = \frac{1}{1+y^2}$. One may use the formulas $D(f^{-1}(y)) = \frac{1}{f'(f^{-1}(y))}$, $D(\tan x) = \sec^2 x$, and $\sec^2 x = \tan^2 x + 1$ to derive this. Thus,

$$D(y\tan^{-1}(y)-\frac{1}{2}\log(1+y^2)) = \tan^{-1}(y)+\frac{y}{1+y^2}-\frac{y}{1+y^2} = \tan^{-1}(y),$$

which is a strictly increasing function. Therefore, $y\tan^{-1}(y)-\frac{1}{2}\log(1+y^2)$ is strictly convex.

11.
$$(e^x,\mathbb{R}) \leftrightarrow \left(\begin{cases} y\log y - y & y>0 \\ 0 & y=0 \end{cases},\mathbb{R}_+\right).$$

e^x is just so very very convex (it's strictly convex). $D(y\log y - y) = \log y$, which is strictly increasing on $\mathbb{R}_{++}$. Furthermore, $\lim_{y\to 0} y\log y = \lim_{y\to 0}\frac{1/y}{-1/y^2} = 0$. Thus, $y\log y - y$ is continuous, which makes it strictly convex on $\mathbb{R}_+$.

12.
$$(\log(1+e^x),\mathbb{R}) \leftrightarrow \left(\begin{cases} y\log y + (1-y)\log(1-y) & y\in(0,1) \\ 0 & y=0,1 \end{cases},[0,1]\right)$$

$D(\log(1+e^x)) = e^x/(1+e^x) = 1/(1+e^{-x})$ is strictly increasing because e^{-x} is strictly decreasing. Thus, $\log(1+e^x)$ is strictly convex.

$$D(y\log y + (1-y)\log(1-y)) = \log y + 1 - \log(1-y) - 1 = \log[y/(1-y)].$$

$y/(1-y)$ is strictly increasing on $(0,1)$, so $y\log y + (1-y)\log(1-y)$ is strictly convex on $(0,1)$. Moreover, $\lim_{y\to 0} y\log y = 0$ and likewise $\lim_{y\to 1}(1-y)\log(1-y)\to 0$. Coupled with $\log 1 = 0$, this implies that the given function is continuous, and thus strictly convex on $[0,1]$.

13.
$$(-\log(1-e^x),-\mathbb{R}_{++}) \leftrightarrow \left(\begin{cases} y\log y - (1+y)\log(1+y) & y>0 \\ 0 & y=0 \end{cases},\mathbb{R}_+\right).$$

$D(-\log(1-e^x)) = e^x/(1-e^x) = 1/(e^{-x}-1)$. This function is strictly increasing on $-\mathbb{R}_{++}$, so $-\log(1-e^x)$ is convex on $-\mathbb{R}_{++}$.

$$D(y\log y + (1+y)\log(1+y)) = \log y + 1 + \log(1+y) + 1.$$

The derivative is strictly increasing on $\mathbb{R}_+$, so $y\log y + (1+y)\log(1+y)$ is strictly convex on $\mathbb{R}+$.

15 (Examples of convex functions). Prove the following functions of $x \in \mathbb{R}$ are convex:

(a) $\log\left(\frac{\sinh ax}{\sinh x}\right)$ for $a \geq 1$.
First, assume that $x > 0$.

$$(\log\sinh)'(x) = \frac{\sinh'(x)}{\sinh x} \implies (\log\sinh)''(x) = \frac{\sinh''(x)\sinh(x) - (\sinh'(x))^2}{\sinh^2(x)}$$
$$\implies (\log\sinh)''(x) = \frac{\sinh^2(x) - \cosh^2(x)}{\sinh^2(x)} = -\frac{1}{\sinh^2(x)} = -\frac{4}{(e^x - e^{-x})^2}.$$

Therefore, $D^2(\log\left(\frac{\sinh ax}{\sinh x}\right)) = \frac{4}{(e^x-e^{-x})^2} - \frac{4a^2}{(e^{ax}-e^{-ax})^2}$. To show this is ≥ 0, we can show that $(e^{ax} - e^{-ax})/a \geq e^x - e^{-x}$. To see this we use power series:

$$\frac{e^{ax} - e^{-ax}}{a} = \frac{1}{a}\left(\sum_{i=0}^{\infty}\frac{(ax)^i}{i!} - \sum_{i=0}^{\infty}\frac{(-ax)^i}{i!}\right) = \frac{2}{a}\sum_{i=0}^{\infty}\frac{(ax)^{2i+1}}{(2i+1)!}$$
$$= 2\sum_{i=0}^{\infty}\frac{a^{2i}x^{2i+1}}{(2i+1)!} \geq 2\sum_{i=0}^{\infty}\frac{x^{2i+1}}{(2i+1)!} = e^x - e^{-x},$$

with strict inequality if $x > 0$ and $a > 1$. Now let's check the derivative at 0.

$$D(\log\left(\frac{\sinh ax}{\sinh x}\right)) = \frac{a\cosh ax}{\sinh ax} - \frac{\cosh x}{\sinh x} = \frac{a\cosh ax\sinh x - \cosh x\sinh ax}{\sinh ax\sinh x}$$
$$= \frac{\frac{a}{2}(\sinh(a+1)x - \sinh(a-1)x) - \frac{1}{2}(\sinh(a+1)x + \sinh(a-1)x)}{\frac{1}{2}(\cosh(a+1)x - \cosh(a-1)x)}$$
$$= \frac{\frac{a-1}{2}\sinh(a+1)x - \frac{a+1}{2}\sinh(a-1)x}{\frac{1}{2}(\cosh(a+1)x - \cosh(a-1)x)}.$$

Applying L'Hopital's rule (differentiate the numerator and denominator twice), we see $\frac{(a-1)(a+1)^2\sinh(a+1)x-(a+1)(a-1)^2\sinh(a-1)x}{(a+1)^2\cosh(a+1)x-(a-1)^2\cosh(a-1)x}$ goes to 0 as $x \to 0$. Therefore, the derivative at 0 is 0. Now $\log\left(\frac{\sinh ax}{\sinh x}\right)$ is an even function. Therefore, its derivative is an odd function. We have shown that its derivative from $x = 0$ to $+\infty$ starts at 0 and increases. Therefore, its derivative from $x = 0$ to $-\infty$ starts at 0 and decreases. In other words, the derivative is increasing, i.e. it is convex.

(b) $\log\left(\frac{e^{ax}-1}{e^x-1}\right)$ for $a \geq 1$. $\log(\frac{\sinh ax}{\sinh x})$ composed with linear $\frac{1}{2}\cdot$ plus a linear function is convex.

$$\log\left(\frac{e^{ax}-1}{e^x-1}\right) = \log\left(\frac{e^{\frac{ax}{2}} - e^{-\frac{ax}{2}}}{e^{\frac{x}{2}} - e^{-\frac{x}{2}}}\right) + \frac{ax}{2} - \frac{x}{2} = \log\left(\frac{\sinh(\frac{ax}{2})}{\sinh(\frac{x}{2})}\right) + \frac{(a-1)x}{2}.$$

16 * (Bregman distances). For a function $\phi : \mathbb{E} \to (-\infty, +\infty]$ that is strictly convex and differentiable on $\operatorname{int}(\operatorname{dom}\phi)$, define the *Bregman distance* $d_\phi : \operatorname{dom}\phi \times \operatorname{int}(\operatorname{dom}\phi) \to \mathbb{R}$ by

$$d_\phi(x, y) = \phi(x) - \phi(y) - \phi'(y)(x - y).$$

(a) Prove $d_\phi(x, y) \geq 0$, with equality if and only if $x = y$.
$\phi(x) - \phi(y) - \phi'(y)(x - y) = \int_y^x (\phi'(x') - \phi'(y))\,\mathrm{d}x'$. If $x > y$, then for some $\epsilon > 0$, $\phi'(z) - \phi'(y) > \epsilon$ for every $z > (x + y)/2$ since ϕ' is strictly increasing. Therefore, $\int_y^x (\phi'(x) - \phi'(y))\,\mathrm{d}x' \geq \epsilon(x - y)/2 > 0$. Likewise, if $y > x$, then for some $\epsilon > 0$, $\phi(y) - \phi(x') > \epsilon$ for all $x' < (x + y)/2$. Thus $\int_y^x (\phi'(x') - \phi'(y))\,\mathrm{d}x' > \epsilon(y - x)/2 > 0$. Clearly if $x = y$ then $d_\phi(x, y) = 0$.

(b) Compute d_ϕ when $\phi(t) = t^2/2$ and when ϕ is the function p defined in Exercise 27.

$$d_{t^2/2}(x, y) = \frac{x^2}{2} - \frac{y^2}{2} - y(x - y) = \frac{x^2}{2} - xy + \frac{y^2}{2} = \frac{1}{2}(x - y)^2.$$

$p(u) = u \log u - u$, and $p(0) = 0$ and $p(u) = +\infty$ for $u < 0$. So $\operatorname{dom} p = \mathbb{R}_+$.

$$d_p(x, y) = x \log x - x - y \log y + y - \log y(x - y) = x \log \frac{x}{y} - x + y.$$

(c) Suppose ϕ is three times differentiable. Prove d_ϕ is convex if and only if $-1/\phi''$ is convex on $\operatorname{int}(\operatorname{dom} \phi)$.

$$\frac{\mathrm{d}}{\mathrm{d}x}(d_\phi(x, y)) = \phi'(x) - \phi'(y), \quad \frac{\mathrm{d}}{\mathrm{d}y}(d_\phi(x, y)) = -\phi'(y) - \phi''(y)(x - y) + \phi'(y) = -\phi''(x - y)$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}(d_\phi(x, y)) = \phi''(x), \quad \frac{\mathrm{d}^2}{\mathrm{d}y\mathrm{d}x}(d_\phi(x, y)) = -\phi''(y), \quad \frac{\mathrm{d}^2}{\mathrm{d}y^2}(d_\phi) = -\phi'''(y)(x - y) + \phi''(y).$$

Therefore, $\det \nabla^2 d_\phi(x, y) = \phi''(x)\phi'''(y)(y - x) + \phi''(x)\phi''(y) - (\phi''(y))^2$.
This proof assumes that $\phi''(x)$ is nonzero for all $x \in \operatorname{int}(\operatorname{dom} \phi)$.
Suppose that $-1/\phi''$ is convex. For $x, y \in \operatorname{int}(\operatorname{dom} \phi)$ with $y > x$. Look at

$$\frac{\det \nabla^2 d_\phi(x, y)}{(y - x)\phi''(x)\phi''(y)} = \frac{\phi'''(y)}{\phi''(y)} - \frac{\phi''(y) - \phi''(x)}{(y - x)\phi''(x)}. \tag{3.2.6}$$

Since $-1/\phi''$ is convex, $(y - x)^{-1}(-1/\phi''(y) + 1/\phi''(x)) = \frac{\phi''(y) - \phi''(x)}{(y - x)\phi''(y)\phi''(x)}$ increases in x. This implies (3.2.6) is ≥ 0, which implies $\det \nabla^2 d_\phi(x, y) \geq 0$ since $y - x, \phi''(x), \phi''(y) > 0$.
Now suppose that d_ϕ is convex. This implies that $\det \nabla^2 d_\phi(x, y) \geq 0$. By moving the right term of (3.2.6) to the other side and dividing by $\phi''(y)$, we get

$$\frac{\phi'''(y)}{(\phi''(y))^2} \geq \frac{\phi''(y) - \phi''(x)}{(y - x)\phi''(x)\phi''(y)}.$$

Now if we swapped x and y in (3.2.6), the inequality would change from $\geq$ to $\leq$, since $x - y$ is negative. Thus,

$$\frac{\phi'''(x)}{(\phi''(x))^2} \leq \frac{\phi''(x) - \phi''(y)}{(x - y)\phi''(x)\phi''(y)}.$$

Together with the previous inequality, we get $\frac{\phi'''(y)}{(\phi''(y))^2} \geq \frac{\phi'''(x)}{(\phi''(x))^2}$, i.e. the derivative of $-1/\phi''$ is increasing.

(d) Extend the results above to the function

$$D_\phi : (\operatorname{dom}\phi)^n \times (\operatorname{int}(\operatorname{dom}\phi))^n \to \mathbb{R}$$

defined by $D_\phi(x, y) = \sum_i d_\phi(x_i, y_i)$.

Clearly, $D_\phi(x, y) = 0$ iff $x = y$ (if any $x_i \neq y_i$, then part (a) implies the sum is positive). Furthermore,

$$D_{\frac{t^2}{2}}(x, y) = \sum_{i=1}^{n} \frac{(x_i - y_i)^2}{2} = \frac{1}{2}\|x - y\|^2$$
$$D_p(x, y) = \sum_{i=1}^{n} x_i \log\left(\frac{x_i}{y_i}\right) - \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i.$$

Clearly, if $D_\phi(x, y)$ is convex, then $d_\phi(x, y)$ is convex, so $-1/\phi''$ is convex. On the other hand, if $-1/\phi''$ is convex, then $d_\phi(x, y)$ is convex, which implies that $D_\phi(x, y)$, being a sum of convex functions, is convex.

17 * (Convex functions on $\mathbb{R}^2$). Prove the following functions of $x \in \mathbb{R}^2$ are convex:

(a)

$$f(x) = \begin{cases} (x_1 - x_2)(\log x_1 - \log x_2) & \text{if } x \in \mathbb{R}^2_{++} \\ 0 & \text{if } x = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Hint: See Exercise 16.

For $x \in \mathbb{R}^2_{++}$, $f(x) = d_p(x_1, x_2) + d_p(x_2, x_1)$ (see Exercise 16 for definitions). $-1/p'' = -u$ is convex, so d_p is convex. Therefore, $f(x)$ is convex on $\mathbb{R}_{++}$. Since it is continuous on $\mathbb{R}_+$

(b)

$$f(x) = \begin{cases} \frac{x_1^2}{x_2} & \text{if } x_2 > 0 \\ 0 & \text{if } x = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

The Hessian is $\begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix}$. The determinant is $\frac{4x_1^2}{x_2^4} - \frac{4x_1^2}{x_2^4} = 0$ and the trace is positive, so the Hessian is PSD on the interior of its domain, which implies convexity on the interior of its domain. We just need to verify convexity involving the point 0. This is rather convenient to show: $\frac{(\lambda x_1)^2}{\lambda x_2} = \lambda \frac{x_1^2}{x_2}$, i.e. the function is positively homogeneous. I find it so weird that $\frac{x_1^2}{x_2}$ is positively homogeneous yet discontinuous at 0.

18. * Prove the function

$$f(x) = \begin{cases} -(x_1 x_2 \dots x_n)^{1/n} & \text{if } x \in \mathbb{R}^n_+ \\ +\infty & \text{otherwise} \end{cases}$$

is convex.

Proof. Denote $f(x) = (x_1 \dots x_n)^{1/n}$. We will show that for $\lambda \in [0,1]$ and $x, y \in \mathbb{R}^n_+$,

$$(\lambda f(x) + (1-\lambda) f(y))^n \le f(\lambda x + (1-\lambda) y)^n,$$

which implies the convexity of $-f(x) = -(x_1 \dots x_n)^{1/n}$ as desired.

$$\begin{aligned}
(\lambda f(x) + (1-\lambda) f(y))^n &= \left(\lambda (x_1 \dots x_n)^{1/n} + (1-\lambda)(y_1 \dots y_n)^{1/n}\right)^{1/n} \\
&= \sum_{k=0}^{n} \binom{n}{k} \lambda^k (1-\lambda)^{n-k} (x_1 \dots x_n)^{\frac{k}{n}} (y_1 \dots y_n)^{\frac{k-n}{n}} \\
f(\lambda x + (1-\lambda) y)^n &= (\lambda x_1 + (1-\lambda) y_1) \dots (\lambda x_n + (1-\lambda) y_n) \\
&= \sum_{k=0}^{n} \lambda^k (1-\lambda)^{n-k} \sum_{\substack{S \subset [n] \\ |S| = k}} \left(\prod_{i \in S} x_i\right) \left(\prod_{i \notin S} y_i\right).
\end{aligned}$$

By the AM-GM inequality (See Section 1.1, Exercise 2(b)) and the fact that $|\{S \subset [n] : |S| = k\}| = \binom{n}{k}$,

$$\begin{aligned}
&\frac{1}{\binom{n}{k}} \sum_{\substack{S \subset [n] \\ |S| = k}} \left(\prod_{i \in S} x_i\right) \left(\prod_{i \notin S} y_i\right) \ge \left[\prod_{\substack{S \subset [n] \\ |S| = k}} \left(\prod_{i \in S} x_i\right) \left(\prod_{i \notin S} y_i\right)\right]^{\frac{1}{\binom{n}{k}}} \\
&= \left[\left(\prod_{i=1}^{n} x_i^{\binom{n-1}{k-1}}\right) \left(\prod_{i=1}^{n} y_i^{\binom{n-1}{k}}\right)\right]^{\frac{1}{\binom{n}{k}}} = \left(\prod_{i=1}^{n} x^{\frac{k}{n}}\right) \left(\prod_{i=1}^{n} y^{\frac{n-k}{n}}\right).
\end{aligned}$$

To obtain the above, note that for any index i, there are $\binom{n-1}{k-1}$ sets of size S containing i, and $\binom{n-1}{k}$ sets not containing i. Thus,

$$\begin{aligned}
f(\lambda x + (1-\lambda) y)^n &= \sum_{k=0}^{n} \lambda^k (1-\lambda)^{n-k} \sum_{\substack{S \subset [n] \\ |S| = k}} \left(\prod_{i \in S} x_i\right) \left(\prod_{i \notin S} y_i\right) \\
&\ge \sum_{k=0}^{n} \binom{n}{k} \lambda^k (1-\lambda)^{n-k} \left(\prod_{i=1}^{n} x^{\frac{k}{n}}\right) \left(\prod_{i=1}^{n} y^{\frac{n-k}{n}}\right) = (\lambda f(x) + (1-\lambda) f(y))^n
\end{aligned}$$

as desired. I wanted to solve this using the Hessian which AFAIK is

$$\frac{\mathrm{d}^2(-f)}{\mathrm{d}x_i^2} = \left(1 - \frac{1}{n}\right) \frac{(x_1 \dots x_n)^{1/n}}{n x_i^2}, \qquad \frac{\mathrm{d}^2(-f)}{\mathrm{d}x_j \mathrm{d}x_i} = -\frac{(x_1 \dots x_n)^{1/n}}{n^2 x_i x_j}.$$

but I don't know how to show it's positive semidefinite. $\square$

19 (Domain of subdifferential). If the function $f : \mathbb{R}^2 \to (-\infty, +\infty]$ is defined by

$$f(x_1, x_2) = \begin{cases} \max\{1 - \sqrt{x_1}, |x_2|\} & \text{if } x_1 \ge 0 \\ +\infty & \text{otherwise,} \end{cases}$$

prove that f is convex but that $\operatorname{dom} \partial f$ is not convex.

Proof. Note $1-\sqrt{x_1}$ is convex since its second derivative is $\frac{1}{4x_1^{3/2}}$, which is positive if $x_1 > 0$, and $\sqrt{x_1}$ is continuous on $\mathbb{R}_+$. Furthermore, $|x_2|$ being a norm (composed with the second coordinate function) is convex. Finally, the max of convex functions is convex. Therefore, f is convex.
$\int \operatorname{dom} f = \{x : x_1 > 0\} \subset \operatorname{dom} \partial f$ by Theorem 3.1.8 (3.6). Now suppose $x_1 = 0$ and $|x_2| > 1$. Since $\operatorname{sgn}(x_2) \in \partial |x_2|$, we have

$$\langle \operatorname{sgn}(x_2)e_2, y-x\rangle \leq |y_2| - |x_2| = |y_2| - f(x) \leq f(y) - f(x),$$

i.e. $\operatorname{sgn}(x_2)e_2 \in \partial f(x)$.
For $|x_2| \leq 1$, $\phi = -\infty e_1$ appears to work.

$$\langle -\infty e_1, y-x\rangle = \mathbb{1}\{y_1 > 0\}(-\infty) \leq f(y) - 1 = f(y) - f(x),$$

since if $y_1 > 0$ the inequality holds automatically because the LHS is $-\infty$, and if $y_1 = 0$ then $f(y) \geq 1 \implies f(y) - 1 \geq 0$. So, it seems that $\phi \in \partial f(x)$. But I suppose that infinite values are not allowed for the subgradient (at least with vectors). There is no finite subgradient, because i.e. at 0, $\langle \phi, \frac{e_1}{2|\phi_1|^2}\rangle \geq -\frac{1}{2|\phi_1|} > -\frac{1}{\sqrt{2}|\phi_1|} = f\left(\frac{e_1}{2|\phi_1|^2}\right) - f(0)$. If $\phi_1 = 0$, then clearly, $\langle \phi, x\rangle = 0 > -\sqrt{x_1}$. □

20 * (Monotonicity of gradients). Suppose that the set $S \subset \mathbb{R}^n$ is open and convex and that the function $f : S \to \mathbb{R}$ is differentiable. Prove f is convex if and only if

$$\langle \nabla f(x) - \nabla f(y), x-y\rangle \geq 0 \text{ for all } x, y \in S,$$

and f is strictly convex if and only if the above inequality holds strictly whenever $x \neq y$. (You may use Exercise 12.)

Proof. Assume f is convex. We show the function $g : \mathbb{R} \to \mathbb{R}$, $g(t) = f(y + t(x-y))$ is convex, and strictly convex if f is strictly convex and $x \neq y$. Take $s, t \in \mathbb{R}$. Given $\lambda \in (0,1)$,

$$\begin{aligned} g(\lambda s + (1-\lambda)t) &= f(y + (\lambda s + (1-\lambda)t)(x-y)) = f(\lambda(y + s(x-y)) + (1-\lambda)(y + t(x-y))) \\ &\leq \lambda f(y + s(x-y)) + (1-\lambda) f(y + t(x-y)) = \lambda g(s) + (1-\lambda) g(t). \end{aligned}$$

We have strict inequality if the things mentioned hold and $t \neq s$, since then $y + s(x-y) \neq y + t(x-y)$. Thus,

$$\langle \nabla f(y), x-y\rangle = \int_0^1 g'(0)\,\mathrm{d}t \leq \int_0^1 g'(t)\,\mathrm{d}t = \int_0^1 \langle \nabla f(y + t(x-y)), x-y\rangle\,\mathrm{d}t = f(x) - f(y),$$

with strict inequality if strict convexity holds. Therefore, $f(x) - f(y) \geq \langle \nabla f(y), x-y\rangle$, and $f(y) - f(x) \geq \langle \nabla f(x), y-x\rangle$ (switching x and y). Adding these, we obtain $0 \geq \langle \nabla f(y), x-y\rangle + \langle \nabla f(x), y-x\rangle$, or $\langle \nabla f(x) - \nabla f(y), x-y\rangle \geq 0$ with strict inequality if f is strictly convex and $x \neq y$.
If we assume $\langle \nabla f(x) - \nabla f(y), x-y\rangle \geq 0$ for any $x, y \in S$, then for any $t > s$,

$$\begin{aligned} &\langle \nabla f(y + t(x-y)) - \nabla f(y + s(x-y)), (t-s)(x-y)\rangle \geq 0 \\ \implies &\langle \nabla f(y + t(x-y)), x-y\rangle \geq \langle \nabla f(y + s(x-y)), x-y\rangle. \end{aligned}$$

In other words, the derivative of $g(t) = f(y + t(x - y))$ is increasing (strictly if the above inequalities are strict). Thus, g is (strictly) convex. $g(\lambda) \le (<) \lambda g(0) + (1-\lambda) g(1)$ then gives (strict) convexity of f. $\square$

21 ** (The log barrier). Use Exercise 20 (Monotonicity of gradients), Exercise 10 in Section 2.1 and Exercise 8 in Section 1.2 to prove that the function $f : \mathbb{S}^n_{++} \to \mathbb{R}$ defined by $f(X) = -\log\det X$ is strictly convex. Deduce the uniqueness of the minimum volume ellipsoid in Section 2.3, Exercise 8 and the matrix completion in Section 2.1, Exercise 12.

Proof. Using the fact that $\nabla - \log\det X = -X^{-1}$,

$$\langle \nabla f(X) - \nabla f(Y), X - Y \rangle = \langle Y^{-1} - X^{-1}, X - Y \rangle = \langle Y^{-1} X \rangle + \langle X^{-1} Y \rangle - 2n \ge 0,$$

since $(Y^{-1}X)^{-1} = X^{-1}Y$ and using Exercise 8 from Section 1.2. Furthermore, we have equality iff $Y^{-1}X = I$, i.e. $X = Y$, so if $X \neq Y$ we have strict inequality. By Exercise 20, f is strictly convex.
Exercise 12 of section 2.1 minimizes $\langle C, X \rangle - \log\det X$ over a convex set (containing a minimizer, deduced from nonemptiness). Since $\langle C, X \rangle$ is linear, $\langle C, X \rangle - \log\det X$ is strictly convex and has a unique minimum.
Similarly, Exercise 13 minimizes $-\log\det X$ over a convex set over which it has a minimizer (the set $\{X \in \mathbb{S}^n_{++} : \|Xy^i\| \le 1, \ \forall i \in [m]\}$, where $\{y^i\}_{i \in [m]}$ spans $\mathbb{R}^n$), which by strict convexity is unique. $\square$

22. Prove the function (2.2.5) is convex on $\mathbb{R}^n$ by calculating its Hessian.

Proof. The function (2.2.5) is

$$f(x) = \log\left(\sum_{i=0}^m \exp\langle a^i, x \rangle \right), \qquad \{a^0, a^1, \ldots, a^m\} \subset \mathbb{R}^n.$$

The derivative w.r.t the jth coordinate is

$$\frac{\partial f(x)}{\partial x_j} = \sum_{i=0}^m \frac{a^i_j \exp\langle a^i, x \rangle}{\sum_{i'=0}^m \exp\langle a^{i'}, x \rangle}.$$

The second order derivatives are (denote $E(x) = \sum_{i=0}^m \exp\langle a^i, x \rangle$):

$$\begin{aligned}
\frac{\partial^2 f(x)}{\partial x_j^2} &= \sum_{i=0}^m \frac{(a^i_j)^2 \exp\langle a^i, x \rangle \sum_{i'=0}^m \exp\langle a^{i'}, x \rangle - a^i_j \exp\langle a^i, x \rangle \sum_{i'=0}^m a^{i'}_j \exp\langle a^{i'}, x \rangle}{\left(\sum_{i'=0}^m \exp\langle a^{i'}, x \rangle\right)^2} \\
&= E(x)^{-2} \sum_{i,i'=0}^m (a^i_j a^i_j - a^i_j a^{i'}_j) \exp\langle a^i + a^{i'}, x \rangle. \\
\frac{\partial^2 f(x)}{\partial x_k x_j} &= \sum_{i=0}^m \frac{a^i_k a^i_j \exp\langle a^i, x \rangle \sum_{i'=0}^m \exp\langle a^{i'}, x \rangle - a^i_j \exp\langle a^i, x \rangle \sum_{i'=0}^m a^{i'}_k \exp\langle a^{i'}, x \rangle}{\left(\sum_{i'=0}^m \exp\langle a^{i'}, x \rangle\right)^2} \\
&= E(x)^{-2} \sum_{i,i'=0}^m (a^i_j a^i_k - a^i_j a^{i'}_k) \exp\langle a^i + a^{i'}, x \rangle.
\end{aligned}$$

In fact, the diagonal and nondiagonal elements of the Hessian end up being basically the same (in form). Now let us show $\nabla^2 f(x)$ is PSD. Let $y \in \mathbb{R}^n$.

$$\begin{aligned}
E(x)^2 y^\top \nabla^2 f(x) y &= \sum_{j,k=1}^{n} y_j y_k \sum_{i,i'=0}^{m} (a_j^i a_k^i - a_j^i a_k^{i'}) \exp\langle a^i + a^{i'}, x\rangle \\
&= \sum_{j,k=1}^{n} y_j y_k \sum_{i<i'}^{m} (a_j^i a_k^i + a_j^{i'} a_k^{i'} - a_j^i a_k^{i'} - a_j^{i'} a_k^i) \exp\langle a^i + a^{i'}, x\rangle \\
&= \sum_{j,k=1}^{n} y_j y_k \sum_{i<i'}^{m} (a_j^i - a_j^{i'})(a_k^i - a_k^{i'}) \exp\langle a^i + a^{i'}, x\rangle \\
&= \sum_{i<i'} \exp\langle a^i + a^{i'}, x\rangle \sum_{j,k=1}^{n} y_j y_k (a_j^i - a_j^{i'})(a_k^i - a_k^{i'}) \\
&= \sum_{i<i'} \exp\langle a^i + a^{i'}, x\rangle \left(\sum_{j=1}^{n} y_j (a_j^i - a_j^{i'}) \right)^2 \geq 0.
\end{aligned}$$

□

23. * If the function $f : \mathbb{E} \to (-\infty, +\infty]$ is essentially strictly convex, prove all distinct points x and y in $\mathbb{E}$ satisfy $\partial f(x) \cap \partial f(y) = \emptyset$. Deduce that f has at most one minimizer.

Proof. Let $\bar{x}_1$ and $\bar{x}_2$ be distinct points in $\operatorname{dom} \partial f$. Suppose that $\phi \in \partial f(\bar{x}_1) \cap \partial f(\bar{x}_2)$. Then, for any $x \in \mathbb{E}$ and $\lambda \in (0,1)$,

$$\begin{aligned}
&\langle \phi, \lambda(x - \bar{x}_1)\rangle \leq \lambda(f(x) - f(\bar{x}_1)) \\
&\langle \phi, (1-\lambda)(x - \bar{x}_2)\rangle \leq (1-\lambda)(f(x) - f(\bar{x}_2)) \\
\implies &\langle \phi, x - (\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2\rangle \leq f(x) - (\lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_2)) \leq f(x) - f(\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2).
\end{aligned}$$

The last line follows by convexity of f. (This problem seems to require that f is convex; the definition of essentially convex given doesn't seem to imply convexity.) Since x was arbitrary, in fact $\phi \in \partial f(\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2)$. Therefore, $\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2 \in \operatorname{dom} f$ for every $\lambda \in [0,1]$. In other words, f is strictly convex along the line segment from $\bar{x}_1$ to $\bar{x}_2$. Now we may revisit the above inequality for any $\lambda \in (0,1)$:

$$\langle \phi, x - (\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2\rangle \leq f(x) - (\lambda f(\bar{x}_1) + (1-\lambda) f(\bar{x}_2)) < f(x) - f(\lambda\bar{x}_1 + (1-\lambda)\bar{x}_2).$$

by strict convexity. (It's interesting how we redo this inequality after using it to get strict convexity.) However, this presents a contradiction if we plug in $x = \lambda\bar{x}_1 + (1-\lambda)\bar{x}_2$: this results in $0 < 0$. Therefore, the initial assumption $\phi \in \partial f(\bar{x}_1) \cap \partial f(\bar{x}_2)$ was faulty, i.e. any distinct points $x, y \in \mathbb{E}$ do not share elements of their subgradients.
This implies a unique minimizer because by Proposition 3.1.5 (3.3), $\bar{x}$ is a minimizer of a proper f iff $0 \in \partial f(\bar{x})$. Having two minimizers means two elements share 0 in their subgradients; which we proved cannot happen for essentially strictly convex f. □

24 (Minimizers of essentially smooth functions). Prove that any minimizer of an essentially smooth function f must lie in $\operatorname{core}(\operatorname{dom} f)$.

Proof. We will show that $\partial\,\mathrm{dom}\, f \subset \mathrm{core}(\mathrm{dom}\, f)$, which implies that any minimizer lies in $\mathrm{core}(\mathrm{dom}\, f)$ since 0 is in the subgradient at a minimizer.
Suppose $x \in \partial\,\mathrm{dom}\, f$ such that there exists d where $f(x+td) = +\infty$ for any $t > 0$. Then, $f'(x,d) = +\infty$. Since f is differentiable at x, $\langle \nabla f(x), d\rangle = +\infty$, which is a contradiction. Thus, $x \in \partial\,\mathrm{dom}\, f$ implies for any $d \in \mathbb{E}$, there exists $t > 0$ such that $x + td \in \mathrm{dom}\, f$, i.e. $x \in \mathrm{core}(\mathrm{dom}\, f)$. □

25 ** (Convex matrix functions). Consider a matrix C in $\mathbb{S}^n_+$.

(a) For matrices $X \in \mathbb{S}^n_{++}$ and $D \in \mathbb{S}^n$, use a power series expansion to prove

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathrm{Tr}(C(X+tD)^{-1})\bigg|_{t=0} \geq 0.$$

If the function $f(t)$ has the power series expansion

$$f(t) = c_0 + c_1 t + c_2 t^2 + \ldots = f(0) + f'(0)t + \frac{f''(0)t^2}{2} + \ldots = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)t^i}{i!},$$

then we can identify $f''(0)$ with $2c_2$. Noting that for small enough x, $(1+x)^{-1} = \sum_{i=0}^{\infty}(-x)^i$, we have

$$\begin{aligned}\mathrm{Tr}(C(X+tD)^{-1}) &= \mathrm{Tr}(CX^{-1/2}(I + tX^{-1/2}DX^{-1/2})^{-1}X^{-1/2}) \\ &= \mathrm{Tr}\left[CX^{-1/2}\left(I - tX^{-1/2}DX^{-1/2} + t^2X^{-1/2}DX^{-1}DX^{-1/2} + \ldots\right)X^{-1/2}\right]\end{aligned}$$

Thus, $\frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathrm{Tr}\left[C(X+tD)^{-1}\right]\Big|_{t=0} = 2\langle C, X^{-1}DX^{-1}DX^{-1}\rangle \geq 0.$

(b) Deduce $X \in \mathbb{S}^n_{++} \mapsto \mathrm{Tr}(CX^{-1})$ is convex.

In Exercise 12 (c), we proved that $\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(X+tD)\Big|_{t=0} = \langle D, \nabla^2 f(X)D\rangle$. Therefore, we have shown that the Hessian of $\mathrm{Tr}\, CX^{-1}$ is PSD, i.e. the function is convex.

(c) Prove similarly the function $X \in \mathbb{S}^n \mapsto \mathrm{Tr}(CX^2)$ and the function $X \in \mathbb{S}^n_+ \mapsto -\mathrm{Tr}(CX^{1/2})$ are convex.

$$\mathrm{Tr}(C(X+tD)^2) = \mathrm{Tr}(CX^2 + tXD + tDX + t^2D^2).$$

Therefore, $\frac{\mathrm{d}^2}{\mathrm{d}t^2} \mathrm{Tr}(C(X+tD)^2)\Big|_{t=0} = 2\langle C, D^2\rangle \geq 0.$

Using the power series $\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \ldots,$

$$\begin{aligned}\mathrm{Tr}(C(X+tD)^{1/2}) &= \mathrm{Tr}(CX^{1/4}(I + tX^{-1/2}DX^{-1/2})^{1/2}X^{1/4}) \\ &= \mathrm{Tr}\left[CX^{1/4}\left(I + \frac{tX^{-1/2}DX^{-1/2}}{2} - \frac{t^2X^{-1/2}DX^{-1}DX^{-1/2}}{8} + \ldots\right)X^{1/4}\right].\end{aligned}$$

Thus, $\frac{\mathrm{d}^2}{\mathrm{d}t^2} - \mathrm{Tr}(C(X+tD)^{1/2})\Big|_{t=0} = \langle C, X^{-1/4}DX^{-1}DX^{-1/4}\rangle/4 \geq 0.$

26 ** (Log-convexity). Given a convex set $C \subset \mathbb{E}$, we say that a function $f : C \to \mathbb{R}_{++}$ is *log-convex* if $\log f(\cdot)$ is convex.

(a) Prove any log-convex function is convex, using Section 1.1, Exercise 9 (Composing convex functions).
exp is a convex, isotonic (in this case, monotonically increasing since its domain is 1d) function. Therefore, $\log f$ convex implies $\exp \log f = f$ convex.

(b) If a polynomial $p : \mathbb{R} \to \mathbb{R}$ has all real roots, prove $1/p$ is log-convex on any interval on which p is strictly positive.
If p has all real roots, it can be factored into a product of degree one polynomials $p = C \prod_{i=1}^{k} (x - r_i)$. Assume that $r_1 \geq r_2 \geq \ldots \geq r_k$ WLOG. If p is strictly positive on the interval I, then for some $1 \leq j \leq k-1$, $r_j > I > r_{j+1}$ (if $r_i \in I$ for some i, then p is zero at some point in I). Therefore, for every $x \in I$,

$$\log 1/p = -\log C - \sum_{i=1}^{j} \log(r_i - x) - \sum_{i=j+1}^{k} \log(x - r_i).$$

We have seen that $-\log x$ is convex, and convex functions composed with linear ones are convex. Plus, the sum of convex functions is convex. Therefore, the $\log 1/p$ is convex on I.

(c) One version of *Hölder's inequality* states, for real $p, q > 1$ satisfying $p^{-1} + q^{-1} = 1$ and functions $u, v : \mathbb{R}_+ \to \mathbb{R}$,

$$\int uv \leq \left(\int |u|^p \right)^{1/p} \left(\int |v|^q \right)^{1/q}$$

when the right hand side is well-defined. Use this to prove the *gamma function* $\Gamma : \mathbb{R} \to \mathbb{R}$ given by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, \mathrm{d}t$$

is log-convex.
Let $\lambda \in (0, 1)$ and $x, y \in \mathbb{R}$.

$$\begin{aligned}
\Gamma(\lambda x + (1-\lambda) y) &= \int_0^\infty t^{\lambda x + (1-\lambda) y - 1} e^{-t} \, \mathrm{d}t = \int_0^\infty (t^{x-1} e^{-t})^\lambda (t^{y-1} e^{-t})^{1-\lambda} \, \mathrm{d}t \\
&\leq \left(\int t^{x-1} e^{-t} \, \mathrm{d}t \right)^\lambda \left(\int t^{y-1} e^{-t} \, \mathrm{d}t \right)^{1-\lambda} \quad p = \frac{1}{\lambda}, q = \frac{1}{1-\lambda}.
\end{aligned}$$

Taking logs on both sides gives $\log \Gamma(\lambda x + (1-\lambda) y) \leq \lambda \log \Gamma(x) + (1-\lambda) \log \Gamma(y)$.

27 ** (Maximum entropy [36]). Define a convex function $p : \mathbb{R} \to (-\infty, +\infty]$ by

$$p(u) = \begin{cases} u \log u - u & \text{if } u > 0 \\ 0 & \text{if } u = 0 \\ +\infty & \text{if } u < 0 \end{cases}$$

and a convex function $f : \mathbb{R}^n \to (-\infty, +\infty]$ by

$$f(x) = \sum_{i=1}^{n} p(x_i).$$

Suppose $\hat{x}$ lies in the interior of $\mathbb{R}^n_+$.

(a) Prove f is strictly convex on $\mathbb{R}^n_+$ with compact level sets.
The derivative of p is $\log u$, which is strictly increasing. Therefore, p is strictly convex on $\mathbb{R}_{++}$. However, since p is continuous ($\lim_{u\to 0} u \log u = \log u/(1/u) = (1/u)/(-1/u^2) = u \to 0$), p is also strictly convex on $\mathbb{R}_+$. If $x \neq y \in \mathbb{R}^n_+$ and $\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) = \sum_{i=1}^{n} p(\lambda x_i + (1-\lambda)y_i) < \sum_{i=1}^{n} \lambda p(x_i) + (1-\lambda)p(y_i),$$

because each i satisfies $p(\lambda x_i + (1-\lambda)y_i) \leq \lambda p(x_i) + (1-\lambda)p(y_i)$, and the inequality is strict for $x_i \neq y_i$ (which must occur). Note for any $u \in \mathbb{R}^n_+$, there is an i with $u_i \geq \|u\|/\sqrt{n}$. Furthermore, by CS $\sum_{i=1}^{n} u_i \leq \sqrt{n}\|u\|$. Thus,

$$\frac{f(x)}{\|x\|} \geq \frac{1}{\sqrt{n}} \log \frac{\|u\|}{\sqrt{n}} - \sqrt{n} \xrightarrow{\|u\|\to\infty} \infty.$$

This means f satisfies the growth condition 1.1.4 in Section 1.1, which implies its level sets are bounded and thus compact since $\mathbb{R}^n_+$ is closed.

(b) Prove $f'(x; \hat{x} - x) = -\infty$ for any point x on the boundary of $\mathbb{R}^n_+$.

$$\begin{aligned} f'(x; \hat{x} - x) &= \lim_{t\downarrow 0} \frac{f(x + t(\hat{x} - x)) - f(x)}{t} \lim_{t\downarrow 0} \frac{\sum_{i=1}^{n} p(x_i + t(\hat{x}_i - x_i)) - p(x_i)}{t} \\ &= \sum_{i=1}^{n} \lim_{t\downarrow 0} \frac{p(x_i + t(\hat{x}_i - x_i)) - p(x_i)}{t} = -\infty. \end{aligned}$$

Exchanging limits and sums is uncontroversial for finite sums, at least one $i \in [n]$ satisfies $x_i = 0, \hat{x}_i > 0$ (recall x is on the boundary while $\hat{x}$ is in the interior). The directional derivative is then $\lim_{t\downarrow 0} \frac{tx \log tx - tx}{t} = x \log x + x \log t - x = -\infty$.

(c) Suppose the map $G : \mathbb{R}^n \to \mathbb{R}^m$ is linear with $G\hat{x} = b$. Prove for any vector $c \in \mathbb{R}^n$ that the problem

$$\inf\{f(x) + \langle c, x\rangle \mid Gx = b, x \in \mathbb{R}^n\}$$

has a unique optimal solution $\bar{x}$, lying in $\mathbb{R}^n_{++}$.
$f(x) + \langle c, x\rangle$ is strictly convex because f is and $\langle c, x\rangle$ is convex. Furthermore, $f(x) + \langle c, x\rangle$ satisfies the growth condition since f does, and $|\langle c, x\rangle|/\|x\| \leq \|c\|$. Finally, the set $\{x \in \mathbb{R}^n : Gx = b\}$ is nonempty and closed. Therefore the level sets of $f(x) + \langle c, x\rangle$ intersected with this set are compact (and convex). Thus $f(x) + \langle c, x\rangle$ has a minimizer, which is unique by strict convexity.

The minimizer cannot lie on the boundary, because by Proposition 2.1.1 (first order necessary condition) the minimizer $\bar{x}$ of any function g must satisfy $g(\bar{x}; x - \bar{x}) \geq 0$ for all $x \in \operatorname{dom} g$. However we proved above that $f(\bar{x}; \hat{x} - \bar{x}) = -\infty$ if $\bar{x}$ is on the boundary. Thus, $\bar{x}$ must lie in the interior.

(d) Use Corollary 2.1.3 (First order conditions for linear constraints) to prove that some vector $\lambda \in \mathbb{R}^m$ satisfies $\nabla f(\bar{x}) = G^*\lambda - c$, and deduce $\bar{x}_i = \exp(G^*\lambda - c)_i$.
By Corollary 2.1.3, the gradient of $f(x) + \langle c, x \rangle$ at the minimum must lie in the range of G^*. Writing this out, $\nabla f(\bar{x}) + c = G^*\lambda$ for some $\lambda \in \mathbb{R}^m$, which gives $\nabla f(\bar{x}) = G^*\lambda - c$. Since $\nabla f(\bar{x}) = \log(\bar{x})$, this gives $\bar{x} = \exp(G^*\lambda - c)$.

28 ** (DAD problems [36]). Consider the following example of Exercise 27 (Maximum entropy). Suppose the $k \times k$ matrix A has each entry a_{ij} nonnegative. We say A *has doubly stochastic pattern* if there is a doubly stochastic matrix with exactly the same zero entries as A. Define a set $Z = \{(i,j) : a_{ij} > 0\}$ and let $\mathbb{R}^Z$ denote the set of vectors with components indexed by Z and $\mathbb{R}^Z_+$ denote those vectors in $\mathbb{R}^Z$ with all nonnegative components. Consider the problem

$$\begin{array}{ll} \inf & \displaystyle\sum_{(i,j)\in Z} (p(x_{ij}) - x_{ij} \log a_{ij}) \\ \text{subject to} & \displaystyle\sum_{i:(i,j)\in Z} x_{ij} = 1 \text{ for } j = 1, 2, \ldots, k \\ & \displaystyle\sum_{j:(i,j)\in Z} x_{ij} = 1 \text{ for } i = 1, 2, \ldots, k \\ & x \in \mathbb{R}^Z. \end{array}$$

(a) Suppose A has doubly stochastic pattern. Prove there is a point $\hat{x}$ in the interior of $\mathbb{R}^Z_+$ which is feasible for the problem above. Deduce that the problem has a unique optimal solution $\bar{x}$, and, for some vectors λ and μ in $\mathbb{R}^k$, $\bar{x}$ satisfies

$$\bar{x}_{ij} = a_{ij} \exp(\lambda_i + \mu_j) \text{ for } (i,j) \in Z.$$

If A has doubly stochastic pattern, then by definition there exists a doubly stochastic matrix $X \in \mathbb{R}^{k\times k}$ such that $X_{ij} > 0$ for every $(i,j) \in Z$ and $X_{ij} = 0$ otherwise. Thus, for every $i \in [k]$, $\sum_{j=1}^k X_{ij} = 1$ and for every $j \in [k]$, $\sum_{i=1}^k X_{ij} = 1$. The nonzero entries of X thus constitute a feasible point in the interior of $\mathbb{R}^Z_+$.
By Exercise 26, since the problem is of the form $\inf p(x) + \langle c, x \rangle$ subject to $Gx = b$ for some G, b for which there is a solution $\hat{x} \in \mathbb{R}^Z_{++}$, a unique optimal solution $\bar{x}$ exists.
Also by Exercise 26, $\bar{x}$ satisfies $\bar{x}_{ij} = \exp((G^*\lambda)_{ij} + c_{ij})$ for each $(i,j) \in Z$. Note that G has $2k$ rows because there are $2k$ constraints. Moreover, the ijth column of G has a 1 in the jth and $k+i$th entries, since the variable x_{ij} is involved in the $k+i$th constraint $\sum_{j':(i,j')\in Z} x_{ij'} = 1$ and the j constraint $\sum_{i':(i',j)\in Z} x_{i'j}$, both with a coefficient of 1. Denoting the first k entries of $\lambda \in \mathbb{R}^{2k}$ as $\mu \in \mathbb{R}^k$ and the second k entries as $\lambda' \in \mathbb{R}^k$, we have

$$\bar{x}_{ij} = \exp(\mu_j + \lambda'_i + \log a_{ij}) = a_{ij} \exp(\lambda'_i + \mu_j).$$

(b) Deduce that A has doubly stochastic pattern if and only if there are diagonal matrices D_1 and D_2 with strictly positive diagonal entries and $D_1 A D_2$ doubly stochastic.
If $D_1 A D_2$ is doubly stochastic, then since the zero set of $D_1 A D_2$ is exactly the same as that of A (the ijth entry of A gets multiplied by $(D_1)_i (D_2)_j > 0$, which preserves zero and nonzero entries), A has doubly stochastic pattern.
Now suppose A has doubly stochastic pattern. By part (a), there exists a doubly stochastic matrix $\bar{X} \in \mathbb{R}^{k \times k}$ such that for every $(i,j) \in [k] \times [k]$, $\bar{X}_{ij} = A_{ij} \exp(\lambda_i + \mu_j)$ for some $\lambda \in \mathbb{R}^k$ and $\mu \in \mathbb{R}^k$. In other words, $D_1 A D_2$ is doubly stochastic, where $D_1 = \operatorname{diag} \exp(\lambda)$ and $D_2 = \operatorname{diag} \exp(\mu)$ are diagonal with strictly positive diagonal entries.

29 ** (Relativizing the Max formula). If $f : \mathbb{E} \to (-\infty, +\infty]$ is a convex function then for points $\bar{x} \in \operatorname{ri}(\operatorname{dom} f)$ and directions $d \in \mathbb{E}$, prove the subdifferential $\partial f(\bar{x})$ is nonempty and

$$f'(\bar{x}; d) = \sup\{\langle \phi, d \rangle \mid \phi \in \partial f(\bar{x})\},$$

with attainment when finite.
Define $S := \operatorname{aff}(\operatorname{dom} f) - \bar{x}$, the linear subspace defined by the translate of the affine span of $\operatorname{dom} f$ by $\bar{x}$ (it doesn't matter which element of $\operatorname{dom} f$ we translate by).
First assume that $d \in S$. Then by "restricting f to S" by a function $g : S \to (-\infty, +\infty]$, $x \mapsto f(\bar{x} + x)$, we have $0 \in \operatorname{core}(\operatorname{dom} g)$ because $\bar{x} \in \operatorname{ri}(\operatorname{dom} f)$, which means we can apply Theorem 3.1.8 (3.6) to conclude that $f'(\bar{x}; d) = g'(0; d) = \langle \phi, d \rangle$ for some $\phi \in \partial g(0) \subset \partial f(\bar{x})$ ($\langle \phi, d \rangle \le g(d) - g(0) = f(\bar{x} + d) - f(\bar{x})$ for every $d \in S$ implies $\langle \phi, x - \bar{x} \rangle \le f(x) - f(\bar{x})$ for any $x \in \operatorname{aff} \operatorname{dom} f$, and if $x \notin \operatorname{aff} \operatorname{dom} f$, the inequality is true because the RHS is infinity).
If $d \notin S$, then for any $t > 0$, $td \notin S \implies \bar{x} + td \notin \operatorname{aff} \operatorname{dom} f \implies \bar{x} + td \notin \operatorname{dom} f$. Therefore, $f'(\bar{x}; d) = +\infty$. We can now conclude that we have shown $\sup\{\langle \phi, d \rangle : \phi \in \partial f(\bar{x})\}$ is attained and equal to $f'(\bar{x}; d)$ whenever $f'(\bar{x}; d)$ is finite.
Now we show that if $v \in S^\perp$ and $\phi \in \partial f(\bar{x})$, then $\phi + v \in \partial f(\bar{x})$. For any $d \in S$, $\langle \phi + v, d \rangle = \langle \phi, d \rangle \le f(\bar{x} + d) - f(\bar{x})$. If $d \notin S$, then the RHS is infinite.
Thus, if $d \notin S$ and has orthogonal projection $v \neq 0$ onto $S^\perp$,

$$\sup\{\langle \phi, d \rangle : \phi \in \partial f(\bar{x})\} = \sup\{\langle cv, d \rangle : c \in \mathbb{R}\} = +\infty = f'(\bar{x}; d).$$

3.3 3.2 The Value Function

Here is another approach to KKT conditions in the convex case. Consider the *inequality constrained convex program*

$$\inf\{f(x) \mid g_i(x) \le 0 \text{ for } i = 1, 2, \dots, m,\ x \in \mathbb{E}\}, \tag{3.3.1}$$

where $f, g_1, g_2, \dots, g_m : \mathbb{E} \to (-\infty, +\infty]$ are convex and satisfy $\emptyset \ne \operatorname{dom} f \subset \cap_i \operatorname{dom} g_i$. Denoting the vector with components $g_i(x)$ by $g(x)$, the function $L : \mathbb{E} \times \mathbb{R}^m_+ \to (-\infty, +\infty]$ defined by

$$L(x; \lambda) = f(x) + \lambda^\top g(x), \tag{3.3.2}$$

is called the *Lagrangian.* A *feasible solution* is a point $x \in \operatorname{dom} f$ satisfying the constraints. Presently we say a vector $\bar{\lambda} \in \mathbb{R}^m_+$ is a *Lagrange multiplier vector* for a feasible solution $\bar{x}$ if $\bar{x}$ minimizes the function $L(\cdot; \bar{\lambda})$ over $\mathbb{E}$ and $\bar{\lambda}$ satisfies complementary slackness: $\bar{\lambda}_i = 0$ if $g_i(\bar{x}) < 0$.

Proposition 3.9 (Lagrangian sufficient conditions (3.2.3))**.** *If the point $\bar{x}$ is feasible for the convex program* (3.3.1) *and there is a Lagrange multiplier vector, then $\bar{x}$ is optimal.*

The KKT conditions are a converse to the above when f, g_i are convex. We *perturb* the problem (3.3.1) and analyze the resulting *(optimal) value function* $v : \mathbb{R}^m \to [-\infty, +\infty]$, defined by the equation

$$v(b) = \inf\{f(x) \mid g(x) \le b\}. \tag{3.3.3}$$

To generalize the definition of convex functions, we introduce the *epigraph* of h:

$$\operatorname{epi}(h) = \{(y, r) \in \mathbb{E} \times \mathbb{R} \mid h(y) \le r\}, \tag{3.3.4}$$

and we say h is a *convex function* if $\operatorname{epi}(h)$ is a convex set. An exercise shows in this case that the domain

$$\operatorname{dom}(h) = \{y \mid h(y) < \infty\}$$

is convex, and that the value function v (3.3.3) is convex. We say h is *proper* if $\operatorname{dom} h$ is nonempty and h never takes the value $-\infty$.

Lemma 3.10 (3.2.6)**.** *If the function $h : \mathbb{E} \to [-\infty, +\infty]$ is convex and some point $\hat{y}$ in* $\operatorname{core}(\operatorname{dom} h)$ *satisfies $h(\hat{y}) > -\infty$, then h never takes the value $-\infty$.*

Here is a regularity condition called the *Slater constraint qualification*, for (3.3.1):

$$\text{There exists } \hat{x} \text{ in } \operatorname{dom} f \text{ with } g_i(\hat{x}) < 0 \text{ for } i = 1, \dots, m. \tag{3.3.5}$$

Theorem 3.11 (Lagrangian necessary conditions (3.2.8))**.** *Suppose that the point $\bar{x}$ in* $\operatorname{dom} f$ *is optimal for the convex program* (3.3.1) *and that the Slater condition* (3.3.5) *holds. Then there is a Lagrange multiplier vector for $\bar{x}$.*

3.4 Exercises for 3.2

1. Prove the Lagrangian sufficient conditions (3.9).

Proof. Suppose that $\bar{x}$ is feasible for (3.3.1) and there is a Lagrange multiplier vector $\bar{\lambda} \in \mathbb{R}^m_+$. Pick any feasible $x \in \mathbb{E}$.

$$f(\bar{x}) = f(\bar{x}) + \bar{\lambda}^\top g(\bar{x}) \leq f(x) + \bar{\lambda} g(x) \leq f(x).$$

The first equality is by complementary slackness ($\bar{\lambda}^\top g(\bar{x}) = 0$); the first inequality is because $\bar{x}$ minimizes $L(\cdot; \bar{\lambda})$ over $\mathbb{E}$, and the last inequality is because x is feasible ($g(x) \leq 0$). □

2. Use the Lagrangian sufficient conditions (3.9) to solve the following problems.

(a)
$$\begin{array}{ll} \inf & x_1^2 + x_2^2 - 6x_1 - 2x_2 + 10 \\ \text{subject to} & 2x_1 + x_2 - 2 \leq 0 \\ & x_2 - 1 \leq 0 \\ & x \in \mathbb{R}^2. \end{array}$$

The Lagrangian $L(x; \lambda)$ is

$$L(x; \lambda) = x_1^2 + x_2^2 - 6x_1 - 2x_2 + 10 + \lambda_1(2x_1 + x_2 - 2) + \lambda_2(x_2 - 1).$$

For any $\lambda \in \mathbb{R}$, $L(x; \lambda)$ is convex in x, being a sum of a strictly convex function $(x_1^2 + x_2^2)$ and a linear function. Therefore, by Proposition 2.2 (first order sufficient condition for convex functions), the global minimizer in x (for any fixed λ) may be computed by setting the gradient to 0:

$$\begin{bmatrix} 2x_1 - 6 + 2\lambda_1 \\ 2x_2 - 2 + \lambda_1 + \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{aligned} x_1 &= 3 - \lambda_1 \\ x_2 &= 1 - \frac{\lambda_1 + \lambda_2}{2}. \end{aligned} \tag{3.4.1}$$

The above enables to find a $(\bar{x}, \bar{\lambda})$ satisfying the first requirement of a Lagrange multiplier vector. We also need to make sure $\bar{x}$ is feasible, and that $(\bar{x}, \bar{\lambda})$ satisfy complementary slackness: $\lambda_1(2x_1 + x_2 - 2) = 0$ and $\lambda_2(x_2 - 1) = 0$. If we try setting $x_2 = 1$, then the second equation in (3.4.1) implies $\lambda_1 + \lambda_2 = 0$, which since we need $\lambda \geq 0$ implies $\lambda = 0$. Then (3.4.1) gives $x_1 = 3$. But then $2x_1 + x_2 - 2 = 6 + 1 - 2 > 0$, i.e. x is not feasible.
Therefore, to satisfy $\lambda_2(x_2 - 1) = 0$, $\lambda_2 = 0$ necessarily. If $\lambda_1 = 0$, then $x_2 = 1$, which we just ruled out. Therefore, $\lambda_1 > 0$ and to satisfy $\lambda_1(2x_1 + x_2 - 2) = 0$, $2x_1 + x_2 - 2 = 0$. Plugging in (3.4.1) with $\lambda_2 = 0$,

$$6 - 2\lambda_1 + 1 - \frac{\lambda_1}{2} - 2 = 5 - \frac{5\lambda_1}{2} = 0 \implies \lambda_1 = 2.$$

Therefore, our optimal $(\bar{x}, \bar{\lambda})$ pair is

$$\bar{x} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \bar{\lambda} = \begin{bmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

It can be checked that the optimal objective value is $L(\bar{x}; \bar{\lambda}) = 5$. Furthermore, $\bar{x}$ is feasible as $2\bar{x}_1 + \bar{x}_2 - 2 = 0$ and $\bar{x}_2 - 1 = -1 < 0$.

(b)
$$\begin{aligned} \inf \quad & -2x_1 + x_2 \\ \text{subject to} \quad & x_1^2 - x_2 \le 0 \\ & x_2 - 4 \le 0 \\ & x \in \mathbb{R}^2. \end{aligned}$$

The Lagrangian for this problem is

$$L(x;\lambda) = -2x_1 + x_2 + \lambda_1(x_1^2 - x_2) + \lambda_2(x_2 - 4).$$

Noting that $L(x;\lambda)$ is convex in x for any $\lambda \ge 0$,

$$L(x;\lambda) = \inf_{x' \in \mathbb{R}^2} L(x';\lambda) \iff \begin{bmatrix} -2 + 2\lambda_1 x_1 \\ 1 - \lambda_1 + \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{matrix} \lambda_1 x_1 = 1 \\ \lambda_1 - \lambda_2 = 1 \end{matrix}. \tag{3.4.2}$$

Complementary slackness reads

$$\lambda_1(x_1^2 - x_2) = 0, \qquad \lambda_2(x_2 - 4) = 0. \tag{3.4.3}$$

Since $\lambda_1 x_1 = 1$, we have $\lambda_1 > 0$. Then by (3.4.3), $x_1^2 = x_2$.
Now suppose $\lambda_2 = 0$. (3.4.2) implies $\lambda_1 = 1 \implies x_1 = 1 \implies x_2 = 1$. Thus,

$$\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \bar{\lambda} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

satisfy (3.4.2), (3.4.3). $\bar{x}$ is feasible and $\bar{\lambda} \ge 0$, thus $\bar{x}$ is optimal. The value is $-2\bar{x}_1 + x_2 = -2$.

(c)
$$\begin{aligned} \inf \quad & x_1 + \frac{2}{x_2} \\ \text{subject to} \quad & -x_2 + \frac{1}{2} \le 0 \\ & -x_1 + x_2^2 \le 0 \\ & x \in \{(x_1, x_2) \mid x_2 > 0\}. \end{aligned}$$

The Lagrangian is

$$L(x;\lambda) = x_1 + \frac{2}{x_2} + \lambda_1\left(-x_2 + \frac{1}{2}\right) + \lambda_2\left(-x_1 + x_2^2\right).$$

Notice that the constraint $x_2 \ge \frac{1}{2}$ already implies $x \in \{(x_1, x_2) \mid x_2 > 0\}$. Thus, $L(x;\lambda)$ is convex for any feasible x and $\lambda \ge 0$. So for feasible x,

$$L(x;\lambda) = \inf_{x' \text{ feasible}} L(x';\lambda) \impliedby \begin{bmatrix} 1 - \lambda_2 \\ -\frac{2}{x_2^2} - \lambda_1 + 2\lambda_2 x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{matrix} \lambda_2 = 1 \\ 2x_2^3 - \lambda_1 x_2^2 - 2 = 0 \end{matrix}. \tag{3.4.4}$$

Complementary slackness is

$$\lambda_1\left(-x_2 + \frac{1}{2}\right) = 0, \qquad \lambda_2\left(-x_1 + x_2^2\right) = 0. \tag{3.4.5}$$

Since $\lambda_2 = 1$ by (3.4.4), $x_1 = x_2^2$ by (3.4.5).
If $\lambda_1 = 0$, then by (3.4.4), $2x_2^3 - 2 = 0 \implies x_2 = 1$. Then $x_1 = x_2^2 = 1$. Thus,

$$\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \bar{\lambda} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

satisfy (3.4.4),(3.4.5). Further note that $\bar{x}$ is feasible and $\bar{\lambda} \geq 0$. Therefore, $\bar{x}$ is optimal, with value 3.

3. Given strictly positive reals $a_1, a_2, \dots, a_n, c_1, c_2, \dots, c_n$ and b, use the Lagrangian sufficient conditions to solve the problem

$$\inf \left\{ \sum_{i=1}^n \frac{c_i}{x_i} \;\middle|\; \sum_{i=1}^n a_i x_i \leq b, x \in \mathbb{R}^n_{++} \right\}.$$

Proof. The Lagrangian for this problem is

$$L(x;\lambda) = \sum_{i=1}^n \frac{c_i}{x_i} + \lambda \left(\sum_{i=1}^n a_i x_i - b \right).$$

By convexity of $L(x;\lambda)$ for $x \in \mathbb{R}^n_{++}$,

$$L(x;\lambda) = \inf_{\text{feasible } x'} L(x';\lambda) \impliedby -\frac{c_i}{x_i^2} + \lambda a_i = 0 \iff \lambda x_i^2 = \frac{c_i}{a_i} \quad \forall i \in [n]. \tag{3.4.6}$$

Complementary slackness says (after using $\lambda > 0$, easy from (3.4.6))

$$\lambda \left(\sum_{i=1}^n a_i x_i - b \right) = 0 \implies \sum_{i=1}^n a_i x_i = b. \tag{3.4.7}$$

Utilizing $\lambda > 0$, we get $x_i = \sqrt{\frac{c_i}{\lambda a_i}}$. From (3.4.7), we get

$$\frac{1}{\sqrt{\lambda}} \sum_{i=1}^n \sqrt{a_i c_i} = b \implies \sqrt{\lambda} = \frac{\sum_{i=1}^n \sqrt{a_i c_i}}{b}.$$

Thus, $x_i = \frac{b\sqrt{c_i/a_i}}{\sum_{i'=1}^n \sqrt{a_{i'} c_{i'}}}$, and the optimal objective value is

$$\sum_{i=1}^n \frac{c_i}{x_i} = \sqrt{\lambda} \sum_{i=1}^n \sqrt{c_i a_i} = \frac{\left(\sum_{i=1}^n \sqrt{c_i a_i} \right)^2}{b}.$$

The fact that increasing a, c make the value go up and increasing b makes it go down make sense. If $\sqrt{a}$ and $\sqrt{c}$ (elementwise) are approximately in the same direction, it becomes harder to minimize the objective, and if they are in nearly different directions, it becomes easy to minimize the objective. ☐

4. For a matrix $A \in \mathbb{S}^n_{++}$ and a real $b > 0$, use the Lagrangian sufficient conditions to solve the problem

$$\inf\{-\log\det X \mid \operatorname{Tr} AX \leq b,\ X \in \mathbb{S}^n_{++}\}.$$

You may use the fact that the objective function is convex with derivative $-X^{-1}$ (see Section 3.1, Exercise 21 (The log barrier)).

Proof. The Lagrangian for this problem is

$$L(x;\lambda) = -\log\det X + \lambda(\operatorname{Tr} AX - b)$$

which is convex in x for $x \in \mathbb{S}^n_{++}$. Thus,

$$L(x;\lambda) = \inf_{x' \in \mathbb{S}^n_{++}} L(x';\lambda) \iff -X^{-1} + \lambda A = 0 \iff X = \frac{1}{\lambda}A^{-1}. \tag{3.4.8}$$

By complementary slackness, since clearly $\lambda \neq 0$, $\operatorname{Tr} AX = b$. Thus, right multiplying (3.4.8) by X and taking the trace, $\operatorname{Tr}(-I) + \lambda \operatorname{Tr} AX = 0 \implies \lambda = n/b$. Therefore, $X = \frac{b}{n}A^{-1}$. This yields objective value

$$\begin{aligned} -\log\det \frac{b}{n}A^{-1} &= -\sum_{i=1}^n \log \lambda(A^{-1})_i - \log(b/n) = -\sum_{i=1}^n \log \lambda(A)_i^{-1} - \log(b/n) \\ &= \sum_{i=1}^n \log \lambda(A)_i - \log(b/n) = \log\det A - \log(b) + \log(n). \end{aligned}$$

□

5 * (Mixed constraints). Consider the convex program (3.3.1) with some additional linear constraints $\langle a^j, x\rangle = d_j$ for vectors a^j in $\mathbb{E}$ and reals d_j. By rewriting each equality as two inequalities (or otherwise), prove a version of the Lagrangian sufficient conditions for this problem.

Proof. Suppose there are k linear constraints $\langle a^j, x\rangle = d_j$ for $j = 1, \ldots, k$, which we can express as $h_j(x) = 0$ with $h_j(x) = \langle a^j, x\rangle - d$. We wish to find a version of the Lagrangian sufficient conditions for this problem.
We begin by casting this problem into the form we already know by replacing the constraints $h_j(x) = 0$ with the constraints $h_j(x) \leq 0$, $-h_j(x) \leq 0$. As h is affine these constraints are all convex, though this does not affect the sufficient condition. The Lagrangian may be written as $L : \mathbb{E} \times \mathbb{R}^m_+ \times \mathbb{R}^k_+ \times \mathbb{R}^k_+ \to \mathbb{R}$,

$$\begin{aligned} L(x;\lambda,\mu^+,\mu^-) &= f(x) + \lambda^\top g(x) + (\mu^+)^\top h(x) - (\mu^+)^\top h(x) \\ &= f(x) + \lambda^\top g(x) + (\mu^+ - \mu^-)^\top h(x). \end{aligned}$$

The Lagrangian sufficient condition says, if $\bar{x}$ is feasible, $\bar{\lambda}, \bar{\mu}^+, \bar{\mu}^- \geq 0$, $\bar{x}$ minimizes $L(x;\bar{\lambda},\bar{\mu}^+,\bar{\mu}^-)$ over the "base set" (e.g. $\mathbb{R}^n$, or $\mathbb{S}^n_{++}$), and complementary slackness holds: $\lambda^\top g(\bar{x}) = 0$, $(\mu^+)^\top h(\bar{x}) = 0$, $(\mu^-)^\top h(\bar{x}) = 0$, then $\bar{x}$ is an optimal solution to the optimization program. Our alternate conditions will be a streamlined way of stating the above. If $\bar{x}$ is feasible, then

$h(\bar{x}) = 0$ so there is no need for the complementary slackness constraints $(\mu^{+,-})^\top h(x) = 0$. Furthermore, notice that any $\mu \in \mathbb{R}^k$ can be written as $\mu = \mu^+ - \mu^-$ for $\mu^+, \mu^- \in \mathbb{R}^k_+$. Therefore, we can instead look for $(\bar{x}, \bar{\lambda}, \bar{\mu})$ where $\bar{\mu} \in \mathbb{R}^k$ such that $\bar{x}$ minimizes

$$\mathcal{L}(x; \bar{\lambda}, \bar{\mu}) = f(x) + \bar{\lambda}^\top g(x) + \bar{\mu}^\top h(x)$$

over the base set. If $\bar{x}$ is also feasible and $\bar{\lambda}^\top g(x) = 0$, then the previous sufficient conditions hold (replacing μ with $\mu^+ - \mu^-$) and thus $\bar{x}$ is optimal.
Therefore, our new Lagrangian sufficient conditions can be stated as finding feasible $\bar{x}$ and $\bar{\lambda} \in \mathbb{R}^m_+$, $\bar{\mu} \in \mathbb{R}^k$ such that $\mathcal{L}(x; \bar{\lambda}, \bar{\mu})$ is minimized over x in the base set by $\bar{x}$, and complementary slackness w.r.t the inequality constraints holds. $\square$

6 (Extended convex functions).

(a) Give an example of a convex function that takes the values 0 and $-\infty$.
Consider the function $f : \mathbb{R} \to [-\infty, +\infty]$

$$f(x) = \begin{cases} -\infty & x < 0 \\ 0 & x = 0 \\ +\infty & x > 0 \end{cases}, \implies \operatorname{epi}(f) = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 < 0\} \cup \{(0, x_2) \in \mathbb{R}^2 : x_2 \geq 0\}.$$

Clearly if $x, y \in \operatorname{epi} f$ with $x_2 < 0$ and $y_2 < 0$, any convex combination of x and y has second coordinate < 0 and is thus in $\operatorname{epi} f$. Now suppose $\lambda \in (0,1)$ and $x, y \in \operatorname{epi} f$ are such that $x_2 < 0$ and $y_2 = 0$. Then $(\lambda x + (1-\lambda)y)_2 = \lambda x_2 + (1-\lambda)y_x < 0$, so $\lambda x + (1-\lambda)y \in \operatorname{epi} f$.
If x, y are such that $x_2 = y_2 = 0$, then $x_1 \geq 0$ and $y_1 \geq 0$, and thus $\lambda x + (1-\lambda)y$ has second coordinate 0 and first coordinate ≥ 0, making it an element of $\operatorname{epi} f$.
This proves that $\operatorname{epi} f$ is convex, i.e. that f is convex.

(b) Prove the value function v defined by equation (3.3.3) is convex.
Let $\lambda \in [0,1]$ and (b_1, r_1) and (b_2, r_2) be in

$$\operatorname{epi} v = \{(b, r) \in \mathbb{R}^{m+1} \mid v(b) \leq r\} = \{(b, r) \in \mathbb{R}^{m+1} \mid \inf\{f(x) \mid g(x) \leq b\} \leq r\}.$$

Consider the element $z = (\lambda b_1 + (1-\lambda)b_2, \lambda r_1 + (1-\lambda)r_2)$. Since $(b_1, r_1) \in \operatorname{epi} v$, given arbitrary $\epsilon > 0$ there exists $x_1 \in \mathbb{E}$ with $g(x_1) \leq b_1$ and $f(x_1) \leq r_1 + \epsilon$. Similarly, $(b_2, r_2) \in \operatorname{epi} v$ gives x_2 with $g(x_2) \leq b_2$ and $f(x_2) \leq r_2 + \epsilon$. Applying the original definition of convexity,

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda b_1 + (1-\lambda)b_2 \quad \text{(component-wise)},$$
$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda r_1 + (1-\lambda)r_2 + \epsilon.$$

In other words, $\lambda x_1 + (1-\lambda)x_2$ is an ϵ-certificate that $v(\lambda b_1 + (1-\lambda)b_2) \leq \lambda r_1 + (1-\lambda)r_2 + \epsilon$. Since ϵ is arbitrary, $v(\lambda b_1 + (1-\lambda)b_2) \leq \lambda r_1 + (1-\lambda)r_2$, i.e. $z \in \operatorname{epi} v$. This proves that v is convex.

(c) Prove that a function $h : \mathbb{E} \to [-\infty, +\infty]$ is convex if and only if it satisfies the inequality

$$h(\lambda x + (1-\lambda)y) \le \lambda h(x) + (1-\lambda)h(y)$$

for any points x and y in $\operatorname{dom} h$ (or $\mathbb{E}$ if h is proper) and any real λ in $(0,1)$.

Suppose f is convex and $x, y \in \operatorname{dom} f$. Let $\lambda \in (0,1)$. Let r_1, r_2 be any values in $\mathbb{R}$ where $f(x) \le r_1$ and $f(y) \le r_2$, in other words $(x, r_1), (y, r_2) \in \operatorname{epi} f$. By convexity, their convex combination is in $\operatorname{epi} f$, i.e.

$$f(\lambda x + (1-\lambda)y) \le \lambda r_1 + (1-\lambda) r_2.$$

WLOG suppose $f(x) = -\infty$. Then, we can take $r_1 \to -\infty$ to obtain $f(\lambda x + (1-\lambda)y) = -\infty = \lambda f(x) + (1-\lambda)f(y)$. Otherwise, $f(x), f(y) \in \mathbb{R}$, in which case we can just take $r_1 = f(x)$, $r_2 = f(y)$ to get $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$.

Now suppose that f satisfies the above inequality for all points $x, y \in \operatorname{dom} f$. Let (x_1, r_1) and (x_2, r_2) be in $\operatorname{epi} f$. Then $x_1, x_2 \in \operatorname{dom} f$ (since $f(x_i) \le r_i < +\infty$), so

$$\lambda f(x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) \le \lambda r_1 + (1-\lambda) r_2,$$

implying $\operatorname{epi} f$ is convex, i.e. f is convex.

If f is proper, then the proof of the second direction does not change (if the inequality holds for all $x, y \in \mathbb{E}$, clearly it holds for $x, y \in \operatorname{dom} f$). For the first direction, if $f(x) = +\infty$, then $f(\lambda x + (1-\lambda)y) \le +\infty = \lambda f(x) + (1-\lambda)f(y)$ since $f(y) > -\infty$ (as f is proper). Therefore, the convexity inequality holds for all $x, y \in \mathbb{E}$.

(d) Prove that if the function $h : \mathbb{E} \to [-\infty, +\infty]$ is convex then $\operatorname{dom} h$ is convex.

By the previous part, if $x, y \in \operatorname{dom} h$, then

$$h(\lambda x + (1-\lambda)y) \le \lambda h(x) + (1-\lambda)h(y) < +\infty$$

for any $\lambda \in [0,1]$, since both $h(x), h(y) < +\infty$. Therefore, $\lambda x + (1-\lambda)y \in \operatorname{dom} h$.

7 (Nonexistence of multiplier). For the function $f : \mathbb{R} \to (-\infty, +\infty]$ defined by $f(x) = -\sqrt{x}$ for $x \in \mathbb{R}_+$ and $+\infty$ otherwise, show there is no Lagrange multiplier at the optimal solution of $\inf\{f(x) \mid x \le 0\}$.

Proof. The optimal solution is clearly $x = 0$ with value 0. The Lagrangian is $L(x;\lambda) = f(x;\lambda) = f(x) + \lambda x$. Given any $\lambda \ge 0$, 0 cannot be a minimizer of $L(\cdot;\lambda)$: if $\lambda = 0$, then any $x > 0$ gives

$$L(x;\lambda) = -\sqrt{x} < 0 = L(0;\lambda).$$

If $\lambda > 0$, then $x = 1/4\lambda^2$ gives

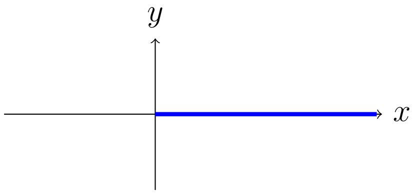
$$L(x;\lambda) = -\frac{1}{2\lambda} + \frac{1}{4\lambda} = -\frac{1}{4\lambda} < L(0;\lambda).$$

□

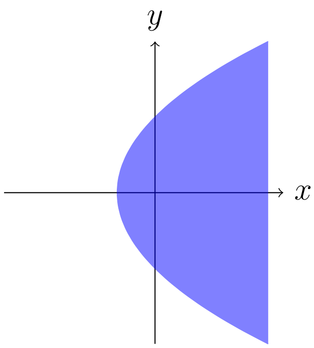
8 (Duffin's duality gap). Consider the following problem (for real b):

$$\inf\{e^{x_2} \mid \|x\| - x_1 \le b,\ x \in \mathbb{R}^2\}. \tag{3.4.9}$$

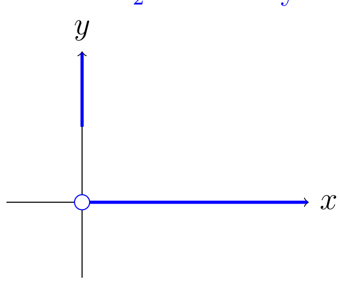
(a) Sketch the feasible region for $b > 0$ and for $b = 0$.
For $b = 0$, the feasible region is $\{x \in \mathbb{R}^2 | \|x\| \leq x_1\}$, i.e. the nonnegative x axis:



For $b > 0$, the feasible region is $\{x | \|x\| \leq x_1 + b\}$. $\|x\| \leq x_1 + b \iff x_1^2 + x_2^2 \leq x_1^2 + 2x_1 b + b^2 \iff x_2^2 \leq 2x_1 b + b^2$. The first statement is an $\iff$ because $x_1 + b$ must be nonnegative if it is $\geq \|x\|$. Thus for $b = 1$, we get the set of $(x, y) \in \mathbb{R}^2$ with $y^2/2 - 1/2 \leq x$, or



(b) Plot the value function v.
Considering (3.4.9) as a function $v(b)$ of b, the above part makes it clear that $v(0) = 1$ (as $x_2 = 0 \implies e^{x_2} = 1$) while for any $b > 0$, $v(b) = 0$ (since the feasible region has a point x with x_2 arbitrarily small). For $b < 0$, $v(b) = +\infty$ as the feasible region is



empty.

(c) Show that when $b = 0$ there is no Lagrange multiplier for any feasible solution. Explain why the Lagrangian necessary conditions (3.11) do not apply.
For $x \neq 0$, $\nabla \|x\| = \frac{x}{\|x\|}$ (In Section 3.1, Exercise 8 we found that $\partial \|x\| = \{x/|x\|\}$). Therefore, the directional derivative of the Lagrangian in the direction of the negative second coordinate is

$$L'(x; \lambda; -\mathbf{e}_2) = -L'(x; \lambda; \mathbf{e}_2) = -e^{x_2} - \lambda \frac{x_2}{\|x\|}.$$

If x is feasible, $x_2 = 0$, which means that $L'(x; \lambda; -\mathbf{e}_2) = -e^{x_2} < 0$. By Proposition 2.1.1 (first order necessary condition), x does not minimize $L(\cdot; \lambda)$.

If $x = 0$, then for $d \in \mathbb{R}^2$,

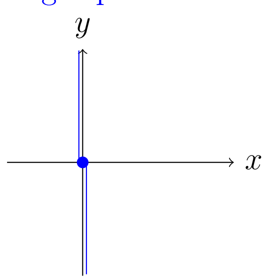
$$L'(0; \lambda; d) = \lim_{t \downarrow 0} \frac{e^{td_2} + \lambda(\|td\| - td_1) - e^0}{t} = d_2 + \lambda(\|d\| - d_1). \tag{3.4.10}$$

If $\lambda = 0$ then we can take $d_2 < 0$ to make $L'(0; \lambda; d) < 0$. If $\lambda > 0$, then for $d_1 > 0$,

$$\begin{aligned} L'(0; \lambda; d) = d_2 + \lambda(\sqrt{d_1^2 + d_2^2} - d_1) &= d_2 + \lambda(d_1\sqrt{1 + \frac{d_2^2}{d_1^2}} - d_1) \\ &\leq d_2 + \lambda(d_1(1 + \frac{d_2^2}{2d_1^2}) - d_1) = d_2 + \lambda\frac{d_2^2}{2d_1}. \end{aligned}$$

We can take $d_1 = 1$ and $d_2 = -\frac{1}{\lambda}$ to get $L'(0; \lambda; d) = -\frac{1}{2\lambda} < 0$. This proves that 0 does not minimize $L(\cdot; \lambda)$ either. Thus, there is no Lagrange multiplier.
Theorem 3.2.8 (3.11) does not apply because the Slater condition does not hold when $b = 0$: $|x| - x_1 \leq 0$ must always be satisfied with equality.

(d) Repeat the above exercises with the objective function e^{x_2} replaced by x_2.
The feasible region stays the same so there is no need to repeat part (a). To calculate the value function, note that for $b < 0$ still $v(b) = +\infty$ as the feasible set is empty and $v(0) = 0$ since $x_2 = 0$. For $b > 0$, since x_2 can get arbitrarily small $v(b) = -\infty$. One might plot v as



The Lagrangian when $b = 0$ is

$$L(x; \lambda) = x_2 + \lambda(\|x\| - x_1).$$

For any feasible x, $L(x; \lambda) = 0$. In fact, $L(d; \lambda)$ is equal to (3.4.10) from part (c), for which we showed there is a $d \in \mathbb{R}^2$ returning a negative value. In this case, it means there is a (non-feasible) $x \in \mathbb{R}^2$ such that $L(x; \lambda) < 0 = L(x'; \lambda)$ for any feasible x', meaning λ cannot be a Lagrange multiplier.

9 ** (Karush-Kuhn-Tucker vectors [167]). Consider the convex program (3.3.1). Suppose the value function v given by equation (3.3.3) is finite at 0. We say the vector $\bar{\lambda} \in \mathbb{R}^m_+$ is a *Karush-Kuhn-Tucker vector* if it satisfies $v(0) = \inf\{L(x; \bar{\lambda}) \mid x \in \mathbb{E}\}$.

(a) Prove that the set of Karush-Kuhn-Tucker vectors is $-\partial v(0)$.
Let $\bar{\lambda} \in \mathbb{R}^m_+$ be a KKT vector and $b \in \mathbb{R}^m$. Then,

$$\begin{aligned} v(0) = \inf\{f(x) + \bar{\lambda}^\top g(x) : x \in \mathbb{E}\} &\leq \inf\{f(x) + \bar{\lambda}^\top g(x) : g(x) \leq b\} \\ &\leq \inf\{f(x) : g(x) \leq b\} + \bar{\lambda}^\top b = v(b) + \bar{\lambda}^\top b. \end{aligned}$$

The second inequality follows because when $g(x) \le b$, it follows that $\bar\lambda^\top g(x) \le \bar\lambda^\top b$ as $\bar\lambda \ge 0$. Rearranging, we obtain that $-\bar\lambda \in \partial v(0)$.
Now if $-\lambda \in \partial v(0)$, since $v(b) \le v(0)$ for any $b \ge 0$ we obtain $v(0) \le v(b) + \bar\lambda^\top b \le v(0) + \bar\lambda^\top b$. This implies that $\bar\lambda \ge 0$. By definition of v, $f(x) \ge v(g(x))$. Thus,

$$f(x) \ge v(g(x)) \ge v(0) - \bar\lambda^\top g(x).$$

I.e., $v(0) \le f(x) + \bar\lambda^\top g(x)$ for every $x \in \mathbb{E}$. Thus, $v(0) \le \inf\{L(x;\bar\lambda)\}$. On the other hand, since any $\bar x$ with $g(\bar x) \le 0$ satisfies $f(\bar x) + \bar\lambda^\top g(\bar x) \le f(\bar x)$, we have $\inf\{L(x;\bar\lambda)\} \le v(0)$. This proves that $v(0) = \inf\{L(x;\bar\lambda)\}$ and that $\bar\lambda$ is a KKT vector.

(b) Suppose the point $\bar x$ is an optimal solution of problem (3.3.1). Prove that the set of Karush-Kuhn-Tucker vectors coincides with the set of Lagrange multiplier vectors for $\bar x$.
Let $\bar\lambda \in \mathbb{R}^m_+$ be a Lagrange multiplier vector. Then,

$$v(0) = f(\bar x) = f(\bar x) + \bar\lambda^\top g(x) = \inf\{L(x;\bar\lambda) : x \in \mathbb{E}\}.$$

Therefore, $\bar\lambda$ is a KKT vector.
Now suppose $\bar\lambda \in \mathbb{R}^m_+$ is a KKT vector. Then

$$f(\bar x) + \bar\lambda^\top g(\bar x) \ge \inf\{f(x) + \bar\lambda^\top g(x)\} = v(0) = f(\bar x) \ge f(\bar x) + \bar\lambda^\top g(\bar x).$$

The first equality is since $\bar\lambda$ is a KKT vector, and the last inequality is because $g(\bar x) \le 0$ while $\bar\lambda \ge 0$. This proves both that $\bar\lambda^\top g(\bar x) = 0$, i.e. complementary slackness, and that $f(\bar x) + \bar\lambda^\top g(\bar x) = \inf\{L(x;\bar\lambda) : x \in \mathbb{E}\}$, i.e. $\bar\lambda$ is a Lagrange multiplier vector.

(c) Prove the Slater condition ensures the existence of a Karush-Kuhn-Tucker vector.
The Slater condition ensures that $0 \in \operatorname{core}(\operatorname{dom} v)$. By Lemma 3.2.6 (3.10), v never takes the value $-\infty$. Then, by Theorem 3.1.8, $\partial v(0)$ is nonempty. I.e., by part (a), there exists a KKT vector. (proof uses convexity of (3.3.1)).

(d) Suppose $\bar\lambda$ is a Karush-Kuhn-Tucker vector. Prove a feasible point $\bar x$ is optimal for problem (3.3.1) if and only if $\bar\lambda$ is a Lagrange multiplier vector for $\bar x$.
If $\bar x$ is optimal, then by part (b), $\bar\lambda$ is a Lagrange multiplier vector for $\bar x$.
If $\bar\lambda$ is a Lagrange multiplier vector for $\bar x$, then by Proposition 3.2.3 (3.9), $\bar x$ is optimal.

10. Prove the equivalence of the Slater and Mangasarian-Fromovitz conditions asserted at the end of the section.

Proof. Suppose $g_1, \ldots, g_m$ are differentiable.
Suppose the Slater condition holds, that is there exists $\hat x \in \operatorname{dom} f$ such that $g(\hat x) < 0$. Now let feasible $x \in \mathbb{E}$ and let $I(x)$ be the set of tight constraints for x, $I(x) = \{i \in [m] : g_i(x) = 0\}$. For each $i \in I$, $g_i(x + t(\hat x - x))$ is convex. Thus, $g_i(x + t(\hat x - x))/t$ is increasing in t, and

$$\langle \nabla g_i(x), \hat x - x\rangle = \lim_{t \downarrow 0} \frac{g_i(x + t(\hat x - x) - g(x)}{t} \le g_i(\hat x) - g_i(x) = g_i(\hat x) < 0.$$

Therefore, $\hat{x} - x$ is a certificate for the Mangasarian-Fromovitz condition.
Now suppose that the Mangasarian-Fromovitz condition holds, i.e. there exists $d \in \mathbb{E}$ such that $\langle \nabla g_i(x), d\rangle < 0$ for all $i \in I(x)$. Then for all $t > 0$ small enough, $g_i(x + td) < g_i(x) = 0$ for every $i \in I(x)$. If we pick $t > 0$ even smaller enough, then for every $i \notin I(x)$ we still have $g_i(x + td) < 0$ by continuity of g (which follows from differentiability). $\square$

11 (Normals to epigraphs). For a function $f : \mathbb{E} \to (-\infty, +\infty]$ and a point $\bar{x}$ in $\operatorname{core}(\operatorname{dom} f)$, calculate the normal cone $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$.

Proof.

$$N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) = \{(d, c) \in \mathbb{E} \times \mathbb{R} : \forall (x, r) \in \operatorname{epi} f,\ d^\top (x - \bar{x}) + c(r - f(\bar{x})) \le 0\}.$$

The condition can be written as

$$\forall x \in \operatorname{dom} f,\ r \ge f(x),\ \langle d, x - \bar{x}\rangle \le -c(r - f(\bar{x})),$$

since $(x, r) \in \operatorname{epi} f \implies x \in \operatorname{dom} f$. If $c < 0$, then the above holds iff $-d/c$ is a subgradient. If $c = 0$, then since $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$, we need $d = 0$. If $c > 0$, then for any $x \in \mathbb{E}$, we can take $r \to +\infty$ to reach a contradiction. Thus, $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) = \{(-c\phi, c) \in \mathbb{E} \times \mathbb{R}_- : \phi \in \partial f(\bar{x})\}$. $\square$

12 (Normals to level sets). Suppose the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex. If the point $\bar{x}$ lies in $\operatorname{core}(\operatorname{dom} f)$ and is not a minimizer for f, prove that the normal cone at $\bar{x}$ to the level set

$$C = \{x \in \mathbb{E} \mid f(x) \le f(\bar{x})\}$$

is given by $N_C(\bar{x}) = \mathbb{R}_+ \partial f(\bar{x})$. Is the assumption $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$ and $f(\bar{x}) > \inf f$ necessary?

Proof. The real way to do it: given $d \in N_C(\bar{x})$, consider the problem $\inf_{x \in C} -\langle d, x - \bar{x}\rangle$. Thanks Sinho! Suppose that $\phi \in \partial f(\bar{x})$. Then, for any $x \in C$, $\langle \phi, x - \bar{x}\rangle \le f(x) - f(\bar{x}) \le 0$, i.e. $\phi \in N_C(\bar{x})$. Since $N_C(\bar{x})$ is a cone, this implies $\mathbb{R}_+ \partial f(\bar{x}) \subset N_C(\bar{x})$.
Now suppose that $d \in N_C(\bar{x})$. We begin by noting that

$$f'(\bar{x}; d') < 0 \implies \langle d, d'\rangle < 0. \tag{3.4.11}$$

To the contrary, suppose that $f'(\bar{x}; d') < 0$ while $\langle d, d'\rangle = 0$. Consider for some small $s > 0$ the directional derivative in the direction $d' + sd$. By sublinearity of the directional derivative at $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$ (Proposition 3.1.2),

$$f'(\bar{x}; d' + sd) \le f'(\bar{x}; d') + s f'(\bar{x}; d) < 0$$

if we pick s small enough. This implies for some $t > 0$ that

$$f(\bar{x} + t(d' + sd)) - f(\bar{x}) < 0$$

yet $\langle d, t(d' + sd)\rangle = ts\langle d, d\rangle > 0$, which contradicts the fact that $d \in N_C(\bar{x})$.
Furthermore, we'll show

$$f'(\bar{x}; d') = 0 \implies \langle d, d'\rangle \le 0. \tag{3.4.12}$$

Since $\bar{x}$ is not a minimizer, there is a point $x \neq \bar{x}$ with $f(x) < f(\bar{x})$, so that if $d'' := x - \bar{x}$ then $\langle d, d'' \rangle < 0$. By sublinearity, we have $f'(\bar{x}; d' + td'') \leq f'(\bar{x}; d') + tf'(\bar{x}; d'') < 0$ for any $t > 0$ since $f'(\bar{x}; d') = 0$ and $f'(\bar{x}, d'') < 0$. Thus, by (3.4.11), $\langle d, d' + td'' \rangle < 0$, which since $t > 0$ is arbitrary implies $\langle d, d' \rangle \leq 0$.
We'll proceed by showing two properties of any sublinear g in relation to a linear function $\langle d, \cdot \rangle$.
First we'll show that if (3.4.11) and (3.4.12) hold with $g(d')$ in place of $f'(\bar{x}; d')$, and g is finite everywhere, then there exists a $C > 0$ such that $g(d_1) \geq C\langle d, d_1 \rangle$ for any d_1 where $g(d_1) \leq 0$. Consider the set $D_1 := \{d_1 \in \mathbb{E} : g(d_1) = -1\}$. By (3.4.11), for all $d_1 \in D_1$, $\langle d, d_1 \rangle < 0$. Therefore, the set

$$S := \{-\langle d, d_1 \rangle \mid d_1 \in D_1\}$$

is a set of positive reals. Suppose elements of S get arbitrarily close to 0, i.e. there exists a sequence $(d_1^i)_{i=1}^\infty$ such that $-\langle d, d_1^i \rangle \downarrow 0$. Then $\langle d, \frac{d_1^i}{\langle d, d_1^i \rangle} \rangle = 1$ while $g(-\frac{d_1^i}{\langle d, d_1^i \rangle}) = -\frac{1}{\langle d, d_1^i \rangle} g(d_1^i) = \frac{1}{\langle d, d_1^i \rangle} \to -\infty$. Consider the set $H := \{\hat{d}_1 \in \mathbb{E} : \langle d, \hat{d}_1 \rangle = -1\}$. The sequence $(-\frac{d_1^i}{\langle d, d_1^i \rangle})_{i=1}^\infty$ is contained in H. Take some $d_0 \in H$. By Section 1.1 Exercise 12, $H - d_0$ is a linear subspace containing $(s_i)_{i=1}^\infty = (-\frac{d_1^i}{\langle d, d_1^i \rangle} - d_0)_{i=1}^\infty$.
For any $M > 0$, there exists N such that $i > N$ implies $g(s_i + d_0) \leq -M$. Since $g(s_i + d_0) \geq g(s_i) - g(-d_0)$, we have $g(s_i) \leq -M + g(-d_0)$.
Then, $g(-s_i - d_0) \geq g(-d_0) - g(s_i) \geq M$, and $g(s_i - d_0) \leq g(s_i) + g(-d_0) \leq -M + 2g(d_0)$. If M is large enough, then $g(-s_i - d_0) > 0$ while $g(s_i - d_0) < 0$. Thus, by continuity of $g(\cdot s_i - d_0)$ (HEHEHE unproved XD) there exists a $t \in \mathbb{R}$ such that $g(ts_i - d_0) = 0$. On the other hand, $\langle d, ts_i - d_0 \rangle = -\langle d, -ts_i + d_0 \rangle = 1$, since $ts_i \in H - d_0 \implies -ts_i \in H - d_0$ by linearity. But, $g(ts_i - d_0) = 0$ while $\langle d, ts_i - d_0 \rangle = 1$ contradicts (3.4.12). Therefore, $\inf S = C > 0$. Thus, for any $d_1 \in D_1, g(d_1) = -1 \geq \langle \frac{d}{C}, d_1 \rangle$. Then for any d_1 where $g(d_1) < 0$, we have $g(d_1) \geq \langle \frac{d}{C}, d_1 \rangle$, by scaling. If $g(d_1) = 0$ by (3.4.12) $\langle d, d_1 \rangle \leq 0$. Furthermore, there exists a sequence d_1^i such that $g(d_1^i) = -1$ while $\langle \frac{d}{C}, d_1^i \rangle \uparrow -1$.
Now we will show for any sublinear g, if $\langle d, d_1 \rangle \leq g(d_1)$ for any $d_1 \in \{d_1 \in \mathbb{E} : g(d_1) \leq 0\}$ and there exists a sequence d_1^i such that $g(d_1^i) = -1$ while $\langle d, d_1^i \rangle \uparrow -1$, then $g(\cdot) \geq \langle d, \cdot \rangle$. For any $d_2 \in \mathbb{E}$ such that $g(d_2) > 0$,

$$\begin{aligned}
&g(d_2 + g(d_2)d_1^i) \leq g(d_2) + g(d_2)g(d_1^i) = g(d_2) - g(d_2) = 0, \\
&\implies g(d_2) \geq g(d_2 + g(d_2)d_1^i) - g(d_2)g(d_1^i) \\
&\qquad \geq \langle d, d_2 + g(d_2)d_1^i \rangle - g(d_2)(\langle d, d_1^i \rangle + \epsilon) \\
&\qquad = \langle d, d_2 \rangle - g(d_2)\epsilon.
\end{aligned}$$

By taking $i \to \infty$ we can let $\epsilon > 0$ be arbitrary, i.e. $g(d_2) \geq \langle d, d_2 \rangle$. Since $g(d_1) \geq \langle d, d_1 \rangle$ already for all d_1 with $g(d_1) \leq 0$, we have $g(\cdot) \geq \langle d, \cdot \rangle$.
Since we proved $f'(\bar{x}; \cdot)$ satisfies (3.4.11) and (3.4.12) and is finite, with $g = f'(\bar{x}; \cdot)$ we have $g(d_1) \geq C\langle d, d_1 \rangle$ for all d_1 with $g(d_1) \leq 0$, and a sequence d_1^i such that $g(d_1^i) = -1$ while $\langle Cd, d_1 \rangle \uparrow -1$. Then, we have by the next result $g(\cdot) \geq \langle Cd, \cdot \rangle$, which implies that $Cd \in \partial f(\bar{x})$. Therefore, $N_C(\bar{x}) \subset \mathbb{R}_+ \partial f(\bar{x})$. □

13 * (Subdifferential of max-function). Consider convex functions

$$g_1, g_2, \ldots, g_m : \mathbb{E} \to (-\infty, +\infty],$$

and define a function $g(x) = \max_i g_i(x)$ for all points $x \in \mathbb{E}$. For a fixed point $\bar{x}$ in $\mathbb{E}$, define the index set $I = \{i \mid g_i(\bar{x}) = g(\bar{x})\}$ and let

$$C = \bigcup \left\{ \partial \left(\sum_{i \in I} \lambda_i g_i \right) (\bar{x}) \,\middle|\, \lambda \in \mathbb{R}^I_+, \sum_{i \in I} \lambda_i = 1 \right\}.$$

(a) Prove $C \subset \partial g(\bar{x})$.

Suppose $\phi \in C$. Then for some $\lambda \in \mathbb{R}^I_+$ where $\sum_{i \in I} \lambda_i = 1$, we have for any $x \in \mathbb{E}$,

$$\langle \phi, x - \bar{x} \rangle \leq \sum_{i \in I} \lambda_i g_i(x) - g(\bar{x}) \leq g(x) - g(\bar{x}),$$

so $\phi \in \partial g(\bar{x})$.

(b) Suppose $0 \in \partial g(\bar{x})$. By considering the convex program

$$\inf_{t \in \mathbb{R},\, x \in \mathbb{E}} \{t \mid g_i(x) - t \leq 0 \text{ for } i = 1, 2, \ldots, m\},$$

prove $0 \in C$.

For any t, x with $t \geq g(x)$, since $\bar{x}$ is a minimizer of g we have $t \geq g(\bar{x})$. Moreover, by setting $t = g(\bar{x})$ and $x = \bar{x}$ we achieve the infimum. By taking t strictly larger than some $g(x)$, we have a strictly feasible point with finite objective, so Slater's condition holds. Thus, by Theorem 3.2.8 (3.11), there exists a Lagrange multiplier vector for the optimal solution $(g(\bar{x}), \bar{x})$. This implies that there exists a $\lambda \in \mathbb{R}^I_+$ such that for any $t \in \mathbb{R}, x \in \mathbb{E}$,

$$g(\bar{x}) \leq t + \sum_{i \in I} \lambda_i (g_i(x) - t) = (1 - \sum_{i \in I} \lambda_i) t + \sum_{i \in I} \lambda_i g_i(x).$$

If $1 \neq \sum_{i \in I} \lambda_i$, then taking $t \to -\infty$ would yield a contradiction. Therefore, $g(\bar{x}) = \sum_{i \in I} \lambda_i g_i(\bar{x})$ is indeed a minimum of $\sum_{i \in I} \lambda_i g_i(x)$. Thus, $0 \in C$.

(c) Deduce $\partial g(\bar{x}) = C$.

Let $\phi \in \partial g(\bar{x})$ and consider the function $g(x) - \langle \phi, x - \bar{x} \rangle$. Notice that since $\phi \in \partial g(\bar{x})$, we have $g(x) - \langle \phi, x - \bar{x} \rangle \geq g(\bar{x})$ with equality at $x = \bar{x}$. In other words, $\bar{x}$ is a minimizer for $g(\cdot) - \langle \phi, \cdot - \bar{x} \rangle$, which is equivalent to $\max_i \{g_i(x) - \langle \phi, x - \bar{x} \rangle\}$. By part (c), (notice the set I is the same), $\bar{x}$ is thus a minimizer for $\sum_{i \in I} \lambda_i (g_i(x) - \langle \phi, x - \bar{x} \rangle)$ for some $\lambda \in \mathbb{R}^I_+$, $\sum_{i \in I} \lambda_i = 1$, so for any $x \in \mathbb{E}$,

$$g(\bar{x}) = \sum_{i \in I} \lambda_i (g(\bar{x}) - \langle \phi, \bar{x} - \bar{x} \rangle) \leq \sum_{i \in I} \lambda_i (g_i(x) - \langle \phi, x - \bar{x} \rangle).$$

Rearranging, we have $\langle \phi, x - \bar{x} \rangle \leq \sum_{i \in I} \lambda_i g(x) - g(\bar{x})$, which implies that $\phi \in \partial \left(\sum_{i \in I} \lambda_i g_i(\bar{x}) \right)$.

14 ** (Minimum volume ellipsoid). Denote the standard basis of $\mathbb{R}^n$ by $\{e^1, e^2, \ldots, e^n\}$ and consider the minimum volume ellipsoid problem (see Section 2.3, Exercise 8)

$$\begin{array}{ll} \inf & -\log\det X \\ \text{subject to} & \|Xe^i\|^2 - 1 \leq 0 \text{ for } i = 1, 2, \ldots, n \\ & X \in \mathbb{S}^n_{++}. \end{array}$$

Use the Lagrangian sufficient conditions (3.9) to prove $X = I$ is the unique optimal solution. (Hint: Use Section 3.1, Exercise 21 (The log barrier).) Deduce the following special case of *Hadamard's inequality:* Any matrix $(x_1\, x_2\, \ldots x^n)$ in $\mathbb{S}^n_{++}$ satisfies

$$\det(x^1\, x^2\, \ldots\, x^n) \leq \|x^1\|\|x^2\| \ldots \|x^n\|.$$

By Section 3.1, Exercise 21, if a minimizer exists it is unique. The Lagrangian is

$$L(X; \lambda) = -\log\det X + \sum_{i=1}^{n} \lambda_i(\|Xe_i\|^2 - 1).$$

This is a convex objective. The gradient is

$$-X^{-1} + \sum_{i=1}^{n} \lambda_i (Xe_i e_i^\top + e_i e_i^\top X).$$

If we set $X = I$ and $\lambda_i = \frac{1}{2}$ for each $i \in [n]$, the gradient equals 0. This means if $\bar{\lambda} = \frac{1}{2}\mathbf{1}$ then $L(I; \bar{\lambda}) = \inf_X L(X; \bar{\lambda})$ by convexity of the Lagrangian in X. Furthermore, I satisfies the constraints of the problem with equality. Therefore, the Lagrangian sufficient conditions imply that I is the unique optimal.
Now applying the above result, we have for any matrix $X = (x^1\, \ldots, x^n) \in \mathbb{S}^n_{++}$ with unit norm columns that $\det X \leq \det I = 1$, by monotonicity of log. Thus,

$$\det(\frac{x^1}{\|x^1\|}\, \frac{x^2}{\|x^2\|}\, \ldots\, \frac{x^n}{\|x^n\|}) \leq 1,$$

which implies the result by the homogeneity of det in each column.

3.5 3.3 The Fenchel Conjugate

The *Fenchel conjugate* of a function $h : \mathbb{E} \to [-\infty, +\infty]$ is the function $h^* : \mathbb{E} \to [-\infty, +\infty]$ defined by

$$h^*(\phi) = \sup_{x \in \mathbb{E}} \{\langle \phi, x \rangle - h(x)\}.$$

The function h^* is convex and if the domain of h is nonempty then h^* never takes the value $-\infty$. The conjugacy operation is *order-reversing*: $f \geq g \implies f^* \leq g^*$.
A subtle example is the function $g : \mathbb{E} \to (-\infty, +\infty]$ defined, for points $a^0, a^1, \ldots, a^m \in \mathbb{E}$, by

$$g(z) = \inf_{x \in \mathbb{R}^{m+1}} \left\{ \sum_i \exp^*(x_i) \,\middle|\, \sum_i x_i = 1,\ \sum_i x_i a^i = z \right\}. \tag{3.5.1}$$

The conjugate is the function in Section 2.2:

$$g^*(y) = 1 + \log\left(\sum_i \exp\langle a^i, y \rangle\right). \tag{3.5.2}$$

Many important convex functions h equal their *biconjugates* h^{**}.

Proposition 3.12 (Log barriers (3.3.3))**.** *The functions* $\mathrm{lb} : \mathbb{R}^n \to (-\infty, +\infty]$ *and* $\mathrm{ld} : \mathbb{S}^n \to (-\infty, +\infty]$ *defined by*

$$\mathrm{lb}(x) = \begin{cases} -\sum_{i=1}^n \log x_i & \text{if } x \in \mathbb{R}^n_{++} \\ +\infty & \text{otherwise} \end{cases}$$

and

$$\mathrm{ld}(X) = \begin{cases} -\log\det X & \text{if } X \in \mathbb{S}^n_{++} \\ +\infty & \text{otherwise} \end{cases}$$

are essentially smooth, and strictly convex on their domains. They satisfy the conjugacy relations

$$\mathrm{lb}^*(x) = \mathrm{lb}(-x) - n \text{ for all } x \in \mathbb{R}^n, \text{ and}$$
$$\mathrm{ld}^*(X) = \mathrm{ld}(-X) - n \text{ for all } X \in \mathbb{S}^n.$$

The perturbed functions $\mathrm{lb} + \langle c, \cdot \rangle$ *and* $\mathrm{ld} + \langle C, \cdot \rangle$ *have compact level sets for any vector* $c \in \mathbb{R}^n_{++}$ *and matrix* $C \in \mathbb{S}^n_{++}$*, respectively.*

Note the simple relationships $\mathrm{lb} = \mathrm{ld} \circ \mathrm{diag}$ and $\mathrm{ld} = \mathrm{lb} \circ \lambda$.

Proposition 3.13 (Fenchel-Young inequality (3.3.4))**.** *Any points* $\phi \in \mathbb{E}$ *and* x *in the domain of a function* $h : \mathbb{E} \to (-\infty, +\infty]$ *satisfy the inequality*

$$h(x) + h^*(\phi) \geq \langle \phi, x \rangle.$$

Equality holds if and only if $\phi \in \partial h(x)$*.*

Analogue of perturbation method gives convex analogue of the chain rule for differentiable functions:

$$\nabla(f + g \circ A)(x) = \nabla f(x) + A^* \nabla g(Ax)$$

for a linear map A. When $A = I$ we obtain a *sum rule.*
In this section we fix a Euclidean space $\mathbb{Y}$ and denote the set of points where a function $g : \mathbb{Y} \to [-\infty, +\infty]$ is finite and continuous by $\operatorname{cont} g$.

Theorem 3.14 (Fenchel duality and convex calculus (3.3.5))**.** *For given functions $f : \mathbb{E} \to (-\infty, +\infty]$ and $g : \mathbb{Y} \to (-\infty, +\infty]$ and a linear map $A : \mathbb{E} \to \mathbb{Y}$, let $p, d \in [-\infty, +\infty]$ be primal and dual values defined, respectively, by the* ***Fenchel problems***

$$p = \inf_{x \in \mathbb{E}} \{f(x) + g(Ax)\} \tag{3.5.3}$$

$$d = \sup_{\phi \in \mathbb{Y}} \{-f^*(A^*\phi) - g^*(-\phi)\}. \tag{3.5.4}$$

These values satisfy the ***weak duality*** *inequality $p \geq d$. If, furthermore f and g are convex and satisfy the condition*

$$0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f) \tag{3.5.5}$$

or the stronger condition

$$A \operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset \tag{3.5.6}$$

then the values are equal ($p = d$) and the supremum in the dual problem (3.5.4) *is attained if finite.*
At any point $x \in \mathbb{E}$, the calculus rule

$$\partial(f + g \circ A)(x) \supset \partial f(x) + A^* \partial g(Ax) \tag{3.5.7}$$

holds, with equality if f and g are convex and either condition (3.5.5) *or* (3.5.6) *holds.*

Corollary 3.15 (Fenchel duality for linear constraints (3.3.11))**.** *Given any function $f : \mathbb{E} \to (-\infty, +\infty]$, and linear map $A : \mathbb{E} \to \mathbb{Y}$, and any element b of $\mathbb{Y}$, the weak duality inequality*

$$\inf_{x \in \mathbb{E}} \{f(x) \mid Ax = b\} \geq \sup_{\phi \in \mathbb{Y}} \{\langle b, \phi \rangle - f^*(A^*\phi)\}$$

holds. If f is convex and b belongs to $\operatorname{core}(A \operatorname{dom} f)$ *then equality holds, and the supremum is attained when finite.*

The *(negative) polar cone* of the set $K \subset \mathbb{E}$ is the convex cone

$$K^- = \{\phi \in \mathbb{E} \mid \langle \phi, x \rangle \leq 0 \text{ for all } x \in K\}$$

and the cone K^{--} is called the *bipolar*. Note $N_C(x) = (C - x)^-$.

Proposition 3.16 (Self-dual cones (3.3.12))**.**

$$(\mathbb{R}^n_+)^- = -\mathbb{R}^n_+ \text{ and } (\mathbb{S}^n_+)^- = -\mathbb{S}^n_+.$$

Corollary 3.17 (Krein-Rutman polar cone calculus (3.3.13))**.** *Any cones $H \subset \mathbb{Y}$ and $K \subset \mathbb{E}$ and linear map $A : \mathbb{E} \to \mathbb{Y}$ satisfy*

$$(K \cap A^{-1}H)^- \supset A^*H^- + K^-.$$

Equality holds if H and K are convex and satisfy $H - AK = Y$ (or in particular $AK \cap \operatorname{int} H \neq \emptyset$).

Theorem 3.18 (Bipolar cone (3.3.14))**.** *The bipolar cone of any nonempty set $K \subset \mathbb{E}$ is given by $K^{--} = \operatorname{cl}(\operatorname{conv}(R_+K))$.*

For example, we deduce immediately that the normal cone $N_C(x)$ to a convex set C at a point $x \in C$ and the *(convex) tangent cone* to C at x defined by $T_C(x) = \operatorname{cl}\mathbb{R}_+(C - x)$ are polars of each other.

Theorem 3.19 (Pointed cones (3.3.15))**.** *If $K \subset \mathbb{E}$ is a closed convex cone, then K is pointed if and only if there is an element y of $\mathbb{E}$ for which the set $C = \{x \in K \mid \langle x, y\rangle = 1\}$ is compact and generates K (that is, $K = \mathbb{R}_+C$).*

3.6 Exercises for 3.3

1. For each of the functions f in Table 3.1, check the calculation of f^* and check $f = f^{**}$. Fenchel conjugate pairs along with domains will be denoted $(f, \operatorname{dom} f) \leftrightarrow (g, \operatorname{dom} g) = (f^*, \operatorname{dom} f^*)$.

1. $(0, \mathbb{R}) \leftrightarrow (0, \{0\})$. $f^*(y) = \sup_{x\in\mathbb{R}} yx - 0x$. If $y \neq 0$, then we can take $x \to +\infty$ or $x \to -\infty$ to get $g(y) = +\infty$. If $y = 0$, clearly $g(y) = 0$. This proves $g = f^*$.
   Now consider $f^{**}(x) = \sup_{y\in\mathbb{R}} xy - g(y) = x0 - g(0) = 0$, since any $y \neq 0$ results in $-\infty$. Thus $f^{**} = f$.

2. $(0, \mathbb{R}_+) \leftrightarrow (0, -\mathbb{R}_+)$. We may as well do $\mathbb{R}^n_+$ and $-\mathbb{R}^n_+$. $f^*(y) = \sup_{x\in\mathbb{E}}\langle y, x\rangle - \delta_{\mathbb{R}^n_+}(x) = \sup_{x\in\mathbb{R}^n_+}\langle y, x\rangle$, since any x not in the set defined by the indicator returns $-\infty$. If y has a positive coordinate, then the sup returns $+\infty$ by taking x to $+\infty$ in that coordinate. Otherwise, $y \in -\mathbb{R}^n_+$, and $\langle y, x\rangle \leq 0$ with equality at $x = 0$. Thus $f^* = g$.
   Now $f^{**}(x) = \inf_{y\in\mathbb{E}}\langle x, y\rangle - \delta_{-\mathbb{R}^n_+}(y) = \inf_{y\in-\mathbb{R}^n_+}\langle x, y\rangle = \delta_{\mathbb{R}^n_+}(x) = f(x)$ by the same reasoning.

3. $(0, [-1, 1]) \leftrightarrow (|y|, \mathbb{R})$. Let us remember from now on that $\delta^*_C(y) = \sup_{x\in C}\langle y, x\rangle$. Thus, $f^*(y) = \sup_{x\in[-1,1]} yx = |y| = g(y)$.
   $f^{**}(x) = \sup_{y\in\mathbb{R}} xy - |y|$. If $|x| > 1$, then taking $y \to \operatorname{sgn}(x)\infty$ returns $+\infty$. For $x \in [-1, 1]$, we have $xy - |y| \leq |x||y| - |y| \leq |y| - |y| = 0$, with equality when $y = 0$. Therefore, $f^{**}(x) = \delta_{[-1,1]}(x) = f(x)$.

4. $(0, [0, 1]) \leftrightarrow (y^+, \mathbb{R})$. $f^*(y) = \sup_{x\in[0,1]} yx = y^+ = g(y)$. $f^{**}(x) = \sup_{y\in\mathbb{R}} xy - y^+$. If $x < 0$, then $y < 0$ implies $xy - y^+ = |xy| \to +\infty$ as $y \to -\infty$. If $x > 1$, then for $y > 0$, $xy - y^+ = (x - 1)y \to +\infty$ as $y \to \infty$. If $x \in [0, 1]$, $xy - y^+ \leq y^+ - y^+ = 0$, with equality at $x = 0$. Thus $f^{**}(x) = \delta_{[0,1]}(x) = f(x)$.

5. $(\frac{|x|^p}{p}, \mathbb{R}) \leftrightarrow (\frac{|y|^q}{q}, \mathbb{R})$, $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. $f^*(y) = \sup_{x\in\mathbb{R}} yx - \frac{|x|^p}{p}$. The derivative of the objective wrt x is $y - |x|^{p-2}x$, which after setting to 0 yields $x = y|y|^{q-2}$ (see Section 2.3, Exercise 6 for Hölder's inequality). By strict concavity, this yields the unique maximum. Plugging it back in, we get $f^*(y) = |y|^q - \frac{|y|^{p(q-1)}}{p} = |y|^q\left(1 - \frac{1}{p}\right) = \frac{|y|^q}{q}$. $f^{**} = f$ by symmetry of conjugates p, q.

6. $(\frac{|x|^p}{p}, \mathbb{R}_+) \leftrightarrow (\frac{|y^+|^q}{q}, \mathbb{R})$, $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Applying similar reasoning as before, we can restrict the supremum to $\mathbf{dom}\, f$. Our optimality condition for $\sup_{x\in\mathbb{R}_+} yx - \frac{|x|^p}{p}$ is now $x = y|y|^{q-2}$ for $x > 0$ or $y - x|x|^{p-2} \le 0$ for $x = 0$, i.e. $y \le 0$. By inspection, this implies that the optimal $x = |y^+|^{q-1}$. Plugging this back in, when $y < 0$ we get $f^*(y) = 0$. Otherwise, $f^*(y) = \frac{y^q}{q}$ by the same steps as the previous part, so $f^*(y) = \frac{|y^+|^q}{q}$.
$f^{**}(x) = \sup_{y\in\mathbb{R}} xy - \frac{|y^+|^q}{q}$. If $x < 0$, then taking $y \to -\infty$ returns $+\infty$. For $x > 0$, we can take the derivative and find the unique maximum at $y = x^{p-1}$ giving $\frac{x^p}{p}$. Thus $f^{**} = f$.

7. $(-\frac{x^p}{p}, \mathbb{R}_+) \leftrightarrow (-\frac{(-y)^q}{q}, -\mathbb{R}_{++})$, $0 < p < 1$, $\frac{1}{p} + \frac{1}{q} = 1$. $f^*(y) = \sup_{x\in\mathbb{R}_+} yx + \frac{x^p}{p}$. If $y \ge 0$, then we can take $x \to +\infty$ to obtain $f^*(y) = +\infty$. If $y < 0$, let us compute the derivative: $y + x^{p-1} = 0 \implies x^{p-1} = -y = |y| \implies x = (-y)^{q-1}$. Plugging this back in, $f^*(y) = -(-y)^q + \frac{(-y)^q}{q} = -\left(1 - \frac{1}{p}\right)(-(-y)^q) = -\frac{(-y)^q}{q}$.
$f^{**}(x) = \sup_{y\in-\mathbb{R}_{++}} xy + \frac{|y|^q}{q}$. If $x < 0$, then we can take $y \to -\infty$ to get $f^{**} = +\infty$. Otherwise, the derivative is $x + y|y|^{q-2} \implies x = |y|^{q-1} \implies y = -x^{p-1}$. Plugging this back in, we get $f^{**} = -x^p + \frac{x^{q(p-1)}}{q} = -\frac{x^p}{p}$.

8. $(\sqrt{1+x^2}, \mathbb{R}) \leftrightarrow (-\sqrt{1-y^2}, [-1, 1])$. $f^*(y) = \sup_{x\in\mathbb{R}} yx - \sqrt{1+x^2}$. If $|y| > 1$, then by taking x with the same sign as y and the inequality $\sqrt{1+r} \le 1 + \frac{r}{2}$ (well $\sqrt{1+r} \le 1 + \sqrt{r}$ would also work, for $r \ge 1$)

$$yx - \sqrt{1+x^2} = |y||x| - |x|\sqrt{1 + \frac{1}{x^2}} \ge |y||x| - |x| - \frac{1}{2|x|} = (|y|-1)|x| - \frac{1}{2|x|} \xrightarrow{|x|\to\infty} +\infty$$

Therefore, $\mathbf{dom}\, f^* \subset [-1, 1]$. Taking the derivative and setting it equal to zero, $y = \frac{x}{\sqrt{1+x^2}} \implies y^2 = \frac{x^2}{1+x^2} \implies x^2 = \frac{y^2}{1-y^2} \implies x = \frac{y}{\sqrt{1-y^2}}$, which is valid for $y \in (-1, 1)$. Plugging this back in, $f^*(y) = \frac{y^2}{\sqrt{1-y^2}} - \sqrt{1 + y^2/(1-y^2)} = \frac{y^2-1}{\sqrt{1-y^2}} = -\sqrt{1-y^2}$. When $|y| = 1$, $yx - \sqrt{1+x^2}$ is always negative, but since it can equal $|x| - \sqrt{1+x^2} \ge -\frac{1}{2|x|}$, we can take $x \to \operatorname{sgn}(y)\infty$ to make the sup 0, i.e. $-\sqrt{1-y^2}$ with $y = 1$.
$f^{**}(x) = \sup_{y\in[-1,1]} xy + \sqrt{1-y^2}$. The derivative is $x - \frac{y}{\sqrt{1-y^2}} = 0 \implies y = \frac{x}{\sqrt{1+x^2}}$. This returns $f^{**}(x) = \frac{x^2}{\sqrt{1+x^2}} + \frac{1}{\sqrt{1+x^2}} = \sqrt{1+x^2}$.

9. $(-\log x, \mathbb{R}_{++}) \leftrightarrow (-1 - \log(-y), -\mathbb{R}_{++})$. $f^*(y) = \sup_{x\in\mathbb{R}_{++}} yx + \log x$. When $y \ge 0, x \to +\infty \implies f^*(y) = +\infty$. Otherwise, the derivative is $y + \frac{1}{x} = 0 \implies x = -\frac{1}{y}$, giving $f^*(y) = -1 + \log(-1/y) = -\log(-y) - 1$.

$f^{**}(x) = \sup_{y \in -\mathbb{R}_{++}} xy + 1 + \log(-y)$. When $x \le 0$, $y \to -\infty \implies f^{**}(x) = +\infty$. Otherwise, the derivative is $x + \frac{1}{y} = 0 \implies y = -\frac{1}{x}$. This yields $f^{**}(x) = -1 + 1 + \log(1/x) = -\log x$.

10. $(\cosh x, \mathbb{R}) \leftrightarrow (y \sinh^{-1}(y) - \sqrt{1+y^2}, \mathbb{R})$. In Section 3.1 Exercise 14, we found that $\sinh^{-1}(y) = \log(y + \sqrt{y^2+1})$. $f^*(y) = \sup_{x \in \mathbb{R}} yx - \cosh x$. Taking the derivative, $y - \sinh x = 0 \implies x = \sinh^{-1} y$. Plugging back in, $f^*(y) = y \sinh^{-1} y - \cosh(\sinh^{-1} y)$. By the relation $\cosh = \sqrt{1 + \sinh^2}$, we obtain $f^*(y) = y \sinh^{-1}(y) - \sqrt{1+y^2}$.
$f^{**}(x) = \sup_{y \in \mathbb{R}} xy - y\sinh^{-1}(y) + \sqrt{1+y^2}$. In Section 3.1 Exercise 14, we computed the derivative $g'(y) = \log(y + \sqrt{y^2+1}) = \sinh^{-1}(y)$. Thus, we get $x - \sinh^{-1}(y) = 0 \implies y = \sinh x$. This gives $f^{**}(x) = x \sinh x - (\sinh x)x + \sqrt{1 + \sinh^2 x} = \cosh x$.

11. $(-\log(\cos x), (-\frac{\pi}{2}, \frac{\pi}{2})) \leftrightarrow (y \tan^{-1}(y) - \frac{1}{2}\log(1+y^2), \mathbb{R})$. $f^*(y) = \sup_{x \in (-\pi/2, \pi/2)} yx + \log \cos x$. The derivative is $y + \tan x = 0 \implies x = \tan^{-1} y$. Plugging this back in, $f^*(y) = y \tan^{-1} y + \log \cos \tan^{-1} y$. Since $\tan^2 + 1 = \sec^2$ and $x \in (-\pi/2, \pi/2)$ implies $\cos x \ge 0$, we have $\cos x = \frac{1}{\sqrt{\tan^2 x + 1}}$. Therefore, $f^*(y) = y\tan^{-1} y + \log(1/(\sqrt{1+y^2})) = y \tan^{-1} y - \frac{1}{2}\log(1+y^2)$.
$f^{**}(x) = \sup_{y \in \mathbb{R}} xy - y \tan^{-1}(y) + \frac{1}{2}\log(1+y^2)$. The derivative is $x - \tan^{-1}(y) - \frac{y}{1+y^2} + \frac{y}{1+y^2} = 0 \implies y = \tan x$. This yields $f^{**}(x) = x \tan x - (\tan x)x + \frac{1}{2}\log(1 + \tan^2 x) = \frac{1}{2}\log \sec^2(x) = -\log\cos x$ (for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$). If $x \notin (-\frac{\pi}{2}, \frac{\pi}{2})$, then $(x - \tan^{-1}(y))y$ is always nonnegative for one of when $y \to +\infty$ or $y \to -\infty$, which cause $\log(1+y^2) \to +\infty$, returning $f^{**}(x) = +\infty$.

12. $$(e^x, \mathbb{R}) \leftrightarrow \left(\begin{cases} y \log y - y & (y > 0) \\ 0 & (y = 0) \end{cases}, \mathbb{R}_+ \right).$$

$f^*(y) = \sup_{x \in \mathbb{R}} yx - e^x$. When $y < 0$, we can take $x \to -\infty$ to get $f^*(y) = +\infty$. Otherwise, taking the derivative gives $y - e^x = 0 \implies x = \log y$ (if $y = 0$, then taking $x \to -\infty$ gives $f^*(y) = 0$), which gives $f^*(y) = y \log y - y$.
$f^{**}(x) = \sup_{y \in \mathbb{R}_+} xy - y\log y + y$. Taking the derivative gives $x - \log y - 1 + 1 = x - \log y = 0 \implies y = e^x$. Thus, $f^{**}(x) = xe^x - xe^x + e^x = e^x$.

13. $$(\log(1+e^x), \mathbb{R}) \leftrightarrow \left(\begin{cases} y \log y + (1-y)\log(1-y) & (y \in (0,1)) \\ 0 & (y = 0, 1) \end{cases}, [0,1] \right).$$

$f^*(y) = \sup_{x \in \mathbb{R}} yx - \log(1+e^x)$. If $y < 0$, then $x \to -\infty \implies f^*(y) = +\infty$. If $y = 0$, $x \to -\infty \implies f^*(0) = 0$. If $y = 1$, then $x - \log(1+e^x) > 0$ but goes to 0 when $x \to \infty$ since it equals $-\log(1+e^{-x})$. For $y > 1$, taking the derivative, $y - \frac{e^x}{1+e^x} = 0 \implies x = \log \frac{y}{1-y}$. Thus, $f^*(y) = y(\log y - \log(1-y)) + \log(1-y) = y \log y + (1-y)\log(1-y)$.
$f^{**}(x) = \sup_{y \in [0,1]} xy - y \log y - (1-y)\log(1-y)$. Taking the derivative, $x - \log y - 1 + \log(1-y) + 1 = x - \log\frac{y}{1-y} = 0 \implies y = \frac{e^x}{1+e^x}$. This gives $f^{**}(x) = \frac{xe^x}{1+e^x} + \frac{e^x}{1+e^x}(\log(1+e^x) - x) + \frac{1}{1+e^x}\log(1+e^x) = \log(1+e^x)$.

14. $$(-\log(1-e^x), -\mathbb{R}_{++}) \leftrightarrow \left(\begin{cases} y \log y - (1+y)\log(1+y) & (y > 0) \\ 0 & (y = 0) \end{cases}, \mathbb{R}_+ \right).$$

$f^*(y) = \sup_{x<0} yx + \log(1-e^x)$. If $y < 0$, $x \to -\infty \implies f^*(y) = +\infty$. If $y = 0$, then $\log(1-e^x) < 0$ but approaches it as $x \to -\infty$ so $f^*(0) = 0$. If $y > 0$, then taking the derivative $y - \frac{e^x}{1-e^x} = 0 \implies x = \log \frac{y}{1+y}$. Plugging this back in, $f^*(y) = y(\log y - \log(1+y)) - \log(1+y) = y \log y - (1+y)\log(1+y)$.
$f^{**}(x) = \sup_{y\in\mathbb{R}_+} xy - y\log y + (1+y)\log(1+y)$. If $x \geq 0$, then since $y\log(1+y) \geq y \log y$, we can take $y \to \infty$ to get $f^{**}(x) = +\infty$. Otherwise, we can notice the pattern that the relationship between x and y stays the same when computing the biconjugate, i.e. $y = \frac{e^x}{1-e^x}$. This gives $f^{**}(x) = \frac{xe^x}{1-e^x} - \frac{e^x}{1-e^x}(x - \log(1-e^x)) - \frac{1}{1-e^x}\log(1-e^x) = -\log(1-e^x)$.

2 (Quadratics). For all matrices $A \in \mathbb{S}^n_{++}$, prove the function $x \in \mathbb{R}^n \mapsto x^\top Ax/2$ is convex and calculate its conjugate. Use the order reversing property of the conjugacy operation to prove

$$A \succeq B \iff B^{-1} \succeq A^{-1} \text{ for } A \text{ and } B \text{ in } \mathbb{S}^n_{++}.$$

Proof. The Hessian is A which is PD, making it strictly convex. The conjugate is $f^*(y) = \sup_{x\in\mathbb{R}^n} \langle y, x\rangle - \frac{1}{2}x^\top Ax$. Taking gradient, $y = Ax \implies x = A^{-1}y \implies f^*(y) = \frac{1}{2}y^\top A^{-1}y$. If $A \succeq B$, then $\frac{1}{2}x^\top Ax \geq \frac{1}{2}x^\top Bx$ for all $x \in \mathbb{E}$, which implies their conjugates have the reverse relationship, i.e. $\frac{1}{2}x^\top A^{-1}x \leq \frac{1}{2}x^\top B^{-1}x$. Thus, $A, B \in \mathbb{S}^n_{++}$, $A \succeq B \implies B^{-1} \succeq A^{-1}$. We can apply this again to get $B^{-1} \succeq A^{-1}$ implies $(A^{-1})^{-1} = A \succeq (B^{-1})^{-1} = B$. ☐

3. Verify the conjugates of the log barriers lb and ld claimed in Proposition 3.3.3 (3.12).

Proof. $\mathrm{lb}^*(y) = \sup_{x\in\mathbb{R}^n_{++}} \langle y, x\rangle + \sum_{i=1}^n \log x_i$. If $y_i \geq 0$ for some $i \in [n]$, then taking $x_i \to +\infty \implies \mathrm{lb}^*(y) = +\infty$. If $y \in -\mathbb{R}^n_{++}$, taking gradient, $y = -\frac{1}{x}$ where the division is pointwise, implying $x = -\frac{1}{y}$. Then $\mathrm{lb}^*(y) = -n + \sum_{i=1}^n \log(-1/y_i) = -\sum_{i=1}^n \log(-y_i) - n = \mathrm{lb}(-y) - n$, as expected.
$\mathrm{ld}^*(Y) = \sup_{X\in\mathbb{S}^n_{++}} \langle Y, X\rangle + \log\det X$. If $x^\top Yx \geq 0$ for some nonzero x, then we can take $X = I + Mxx^\top$ for $M \to +\infty$ to show $\mathrm{ld}^*(Y) = +\infty$. Otherwise, $Y \in -\mathbb{S}^n_{++}$. Taking the gradient, $Y = -X^{-1} \implies X = -Y^{-1} \implies f^*(Y) = -\langle Y, Y^{-1}\rangle + \log\det(-Y^{-1}) = -\log\det(-Y) - n = \mathrm{ld}(-Y) - n$, as expected. ☐

4 * (Self-conjugacy). Consider functions $f : \mathbb{E} \to (-\infty, +\infty]$.

(a) Prove $f = f^*$ if and only if $f(x) = \|x\|^2/2$ for all points $x \in \mathbb{E}$.
We have for any $y \in \mathbb{E}$, $f(y) = \sup_{x\in\mathbb{E}} \langle y, x\rangle - f(x)$. This implies that $f(y) \geq \langle y, y\rangle - f(y) \implies f(y) \geq \|y\|^2/2$. Then,

$$f(y) = \sup_{x\in\mathbb{E}} \langle y, x\rangle - f(x) \leq \sup_{x\in\mathbb{E}} \langle y, x\rangle - \frac{\|x\|^2}{2} \leq \frac{\|y\|^2}{2},$$

because $\langle y, x\rangle - \frac{\|x\|^2}{2}$ is maximized by $x = y$. The other direction is implied by Exercise 2, taking $A = I$.

(b) Find two distinct functions f satisfying $f(-x) = f^*(x)$ for all points $x \in \mathbb{E}$.
Consider $\delta_{\mathbb{R}^n_+}$, the indicator function of the set of points with nonnegative coordinates. Given a cone K, we have $\delta_K^* = \delta_{K^-}$, the indicator of the bipolar cone. To see this, $\delta_K^*(y) = \sup_{x \in K}\langle y, x\rangle$. If $y \notin K^-$, then there exists $x \in K$ where $\langle y, x\rangle > 0$, and we can take $rx \in K$ with $r \to \infty$ to show $\delta_K^*(y) = +\infty$. Otherwise if $y \in K^-$, every $x \in K$ satisfies $\langle y, x\rangle \leq 0$, with attainment at $x = 0 \in K$. Thus, $\delta_K^* = \delta_{K^-}$. We happen to have $(\mathbb{R}^n_+)^- = -\mathbb{R}^n_+$. Therefore, $\delta^*_{\mathbb{R}^n_+}(x) = \delta_{-\mathbb{R}^n_+}(x) = \delta_{\mathbb{R}^n_+}(-x)$.
Is there a more satisfying answer, i.e. a function that is differentiable, or even just finite, everywhere? XD

5 * (Support functions). The conjugate of the indicator function of a nonempty set $C \subset \mathbb{E}$, namely $\delta_C^* : \mathbb{E} \to (-\infty, +\infty]$ is called the *support function* of C. Calculate it for the following sets: (For each of these sets, we will preemptively consider $\delta_C^*(y) = \sup_{x \in C}\langle y, x\rangle$.)

(a) the halfspace $\{x \mid \langle a, x\rangle \leq b\}$ for $0 \neq a \in \mathbb{E}$ and $b \in \mathbb{R}$
Decompose $y = ca + v$ where $\langle v, a\rangle = 0$. The halfspace contains some feasible point, because $-ra$ for r large enough is feasible. Then $\langle a, ra + v\rangle = \langle a, ra\rangle$, so that adding any v perpendicular to a returns a feasible point. If $v \neq 0$, then note $\langle a, ra + sv\rangle = \langle a, ra\rangle \leq b$, i.e. $ra + sv$ is feasible. But $\langle y, ra + sv\rangle = rc + s\langle v, v\rangle \to +\infty$ as $s \to +\infty$. Otherwise, $y = ca$. Then $\delta_C^*(y) = \sup_{x:\langle a,x\rangle \leq b}\langle ca, x\rangle$. If $c < 0$, then we can take $x = ra$, $r \to -\infty \implies \delta_C^*(y) = +\infty$. If $c = 0$, then $\delta_C^*(y) = \delta_C^*(0) = 0$. If $c > 0$, then $\langle ca, x\rangle = c\langle a, x\rangle \leq cb$ with equality at appropriate choice of x (since a is nonzero). Thus $\delta_C^*(ca) = cb$ when $c \geq 0$.
Summarizing, $\delta_C^*(y) = \delta_{\mathbb{R}_+ a}(y) + \frac{\langle y, a\rangle}{\|a\|^2} b$.

(b) the unit ball B
$\delta_C^*(y) = \sup_{x \in B}\langle y, x\rangle = \|y\|$ by Cauchy-Schwartz (or Section 2.3, Exercise 5).

(c) $\{x \in \mathbb{R}^n_+ \mid \|x\| \leq 1\}$
For $y \in \mathbb{R}^n_+$, since $\arg\max_{x \in B}\langle y, x\rangle = y/\|y\|$, we still have $\delta_C^*(y) = \|y\|$. Notice that for $x \in \mathbb{R}^n_+$, $\langle y, x\rangle \leq \langle y^+, x\rangle$. Thus, $\delta_C^*(y) = \sup_{x \in \mathbb{R}^n_+, \|x\| \leq 1}\langle y^+, x\rangle = \|y^+\|$. So actually, $\delta_C^*(y) = \|y^+\|$.

(d) the *polytope* $\operatorname{conv}\{a^1, a^2, \ldots, a^m\}$ for given points $a^1, a^2, \ldots, a^m$ in $\mathbb{E}$.
Given $\lambda_1, \ldots, \lambda_m \geq 0$ summing to 1,

$$\langle y, \sum_{i=1}^m \lambda_i a_i\rangle = \sum_{i=1}^m \lambda_i\langle y, a_i\rangle \leq \max_{i \in [m]}\langle y, a_i\rangle.$$

We get equality by setting $\lambda_{i^*} = 1$ for some i^* achieving the maximum. Therefore, $\delta_C^*(y) = \max_{i \in [m]}\langle y, a_i\rangle$.

(e) A cone K
In 4(b), I proved that $\delta_K^* = \delta_{K^-}$.

(f) the epigraph of a convex function $f : \mathbb{E} \to (-\infty, +\infty]$
$\delta_C^*(y, b) = \sup_{(x,r):f(x) \leq r}\langle y, x\rangle + rb = \sup_{(x,r):f(x)\leq r,\, x \in \operatorname{dom} f}\langle y, x\rangle + rb$. We can restrict to

$\operatorname{dom} f$ because $\{(x,r) : f(x) \le r,\ f(x) = +\infty\} = \emptyset$. If $b > 0$, then we can take $r \to +\infty$ to get $\delta_C^*(y,b) = +\infty$. Otherwise, $b \le 0$, and we can always take $r = f(x)$ to maximize (otherwise r would be larger, which since $b \le 0$ could only decrease the objective). Thus we obtain $\delta_C^*(y,b) = \sup_{x \in \operatorname{dom} f} \langle y, x\rangle - |b| f(x)$. Therefore, $\delta_C^*(y,b) = (-bf)^*(y)$ if $b \le 0$ and $+\infty$ otherwise.

Perhaps it's possible to show that $f^* = +\infty$ if f is concave, which would let us just have $\delta_C^*(y,b) = (-bf)^*(y)$, but I'm too lazy to show it right now.

(g) the subdifferential $\partial f(\bar{x})$, where the function $f : \mathbb{E} \to (-\infty, +\infty]$ is convex and the point $\bar{x}$ lies in $\operatorname{core}(\operatorname{dom} f)$

Since f does not take on $-\infty$ and $\bar{x} \in \operatorname{core}(\operatorname{dom} f)$, by Theorem 3.1.8 (max formula)

$$f'(\bar{x}; d) = \max_{\phi \in \partial f(\bar{x})} \langle \phi, d\rangle = \delta^*_{\partial f(\bar{x})}(d).$$

(h) $\{Y \in \mathbb{S}^n_+ \mid \operatorname{Tr} Y = 1\}$

$$\sup_{Y \in \mathbb{S}^n_+ | \operatorname{Tr} Y = 1} \langle X, Y\rangle \le \sum_{i=1}^n \lambda_i(X)\lambda_i(Y) \le \lambda_1(X),$$

with equality when $\lambda_1(Y) = 1$ and Y has a simultaneous decomposition with X, i.e. for an orthogonal $U \in \mathbb{R}^{n\times n}$, $X = U \operatorname{diag}\lambda(X) U^\top$, $Y = U \operatorname{diag}\lambda(Y) U^\top$. Thus, $\delta_C^*(X) = \lambda_1(X)$.

Note that in Section 3.1 Exercise 9, we proved that $C = \partial\lambda_1(0)$. Thus, using the previous part of this question, we could compute $\delta_C^*(X)$ by computing $\lambda_1'(0; X)$. Interestingly, this works out to $\lambda_1(X)$, i.e. $\lambda_1'(0; \cdot) = \lambda_1(\cdot)$.

6. Calculate the conjugate and biconjugate of the function

$$f(x_1, x_2) = \begin{cases} \frac{x_1^2}{2x_2} + x_2 \log x_2 - x_2 & \text{if } x_2 > 0 \\ 0 & \text{if } x_1 = x_2 = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

$$f^*(y_1, y_2) = \sup_{x \in \mathbb{R}^2 : x = 0,\, x_2 > 0} x_1 y_1 + x_2 y_2 - \frac{x_1^2}{2x_2} - x_2 \log x_2 + x_2.$$

Taking the derivative of the objective and setting it equal to zero,

$$\begin{bmatrix} y_1 - \frac{x_1}{x_2} \\ y_2 + \frac{x_1^2}{2x_2^2} - \log x_2 \end{bmatrix} = 0, \implies x_1 = y_1 x_2 \implies x = \begin{bmatrix} y_1 e^{y_2 + \frac{y_1^2}{2}} \\ e^{y_2 + \frac{y_1^2}{2}} \end{bmatrix}$$

Plugging this back in,

$$\begin{aligned} f^*(y_1, y_2) &= y_1^2 e^{y_2 + \frac{y_1^2}{2}} + y_2 e^{y_2 + \frac{y_1^2}{2}} - \frac{y_1^2}{2} e^{y_2 + \frac{y_1^2}{2}} - (y_2 + \frac{y_1^2}{2}) e^{y_2 + \frac{y_1^2}{2}} + e^{y_2 + \frac{y_1^2}{2}} \\ &= e^{y_2 + \frac{y_1^2}{2}}. \end{aligned}$$

7 ** (Maximum entropy example).

(a) Prove the function g defined by (3.5.1) is convex.
The function g was

$$g(z) = \inf_{x\in\mathbb{R}^{m+1}} \left\{ \sum_{i=0}^m \exp^*(x_i) \,\middle|\, \sum_{i=0}^m x_i = 1,\ \sum_{i=0}^m x_i a^i = z \right\}.$$

$g(z) = v(z, -z, 1, -1)$ with the value function $v(b) = \inf_x\{f(x) \mid g(x) \le b\}$ with $f(x) = \sum_i \exp^*(x)$ and $g_j(x) = \sum_{i=0}^m a_j^i x_i,\ g_{-j}(x) = -\sum_{i=0}^m a_j^i x_i$ for $j \in [n]$. $g_{n+1}(x) = \sum_{i=0}^m x_i,\ g_{-n-1}(x) = \sum_i -x_i$. Since f, g are convex, the value function v is convex; we proved this in Section 3.2, Exercise 6. Thus for all $z_1, z_2 \in \operatorname{dom} g$ and $\lambda \in (0,1)$, we have

$$\begin{aligned} g(\lambda z_1 + (1-\lambda)z_2) &= v(\lambda(z_1, -z_1, 1, -1) + (1-\lambda)(z_2, -z_2, 1, -1)) \\ &\le \lambda v(z_1, -z_1, 1, -1) + (1-\lambda)v(z_2, -z_2, 1, -1) = \lambda g(z_1) + (1-\lambda)g(z_2). \end{aligned}$$

Therefore, g is convex (see Section 3.2, Exercise 6).

(b) For any point $y \in \mathbb{R}^{m+1}$, prove

$$g^*(y) = \sup_{x\in\mathbb{R}^{m+1}} \left\{ \sum_i (x_i\langle a^i, y\rangle - \exp^*(x_i)) \,\middle|\, \sum_i x_i = 1 \right\}.$$

$$\begin{aligned} g^*(y) &= \sup_{z\in\mathbb{R}^n} \left[\langle y, z\rangle - \inf_{x\in\mathbb{R}^{m+1}} \left\{ \sum_{i=0}^m \exp^*(x_i) \,\middle|\, \sum_{i=0}^m x_i = 1,\ \sum_{i=0}^m x_i a^i = z \right\} \right] \\ &= \sup_{z\in\mathbb{R}^n,\, x\in\mathbb{R}^{m+1}} \left\{ \langle y, z\rangle - \sum_{i=0}^m \exp^*(x_i) \,\middle|\, \sum_{i=0}^m x_i = 1, \sum_{i=0}^m x_i a^i = z \right\} \\ &= \sup_{z\in\mathbb{R}^n,\, x\in\mathbb{R}^{m+1}} \left\{ \sum_{i=0}^m (x_i\langle y, a^i\rangle - \exp^*(x_i)) \,\middle|\, \sum_{i=0}^m x_i = 1 \right\} \end{aligned}$$

and we can drop the z in the sup since it is independent of the quantity in the set notation.

(c) Apply Exercise 27 in Section 3.1 to deduce the conjugacy formula (3.5.2).
Notice that we can express g^* as

$$g^*(y) = -\inf_{x\in\mathbb{R}^{m+1}} \left\{ \sum_{i=0}^m \exp^*(x_i) + \langle c, x\rangle \,\middle|\, \mathbf{1}^\top x = 1 \right\},$$

where $c_i = -\langle a^i, y\rangle$. By Exercise 27 of Section 3.1, the optimal solution exists and is unique (the point $\hat{x} = \mathbf{1}/(m+1)$ satisfies $\mathbf{1}^\top\hat{x} = 1$ and is in the interior of $\mathbb{R}_+^{m+1}$), and moreover the optimal $\bar{x} \in \mathbb{R}_{++}^{m+1}$ satisfies $\bar{x} = \exp(\lambda\mathbf{1} - c)$ for some $\lambda \in \mathbb{R}$. By feasibility of $\bar{x}$,

$$\sum_{i=0}^m \exp(\lambda + \langle a^i, y\rangle) = 1 \implies \lambda = -\log\left(\sum_{i=0}^m \exp\langle a^i, y\rangle \right).$$

Now we can simply evaluate $g^*(y)$:

$$\begin{aligned}
-g^*(y) &= \sum_{i=0}^{m} \bar{x}_i \log \bar{x}_i - \bar{x}_i - \langle a^i, y\rangle x_i \\
&= \sum_{i=0}^{m} \exp(\lambda + \langle a^i, y\rangle)(\lambda + \langle a^i, y\rangle) - \exp(\lambda + \langle a^i, y\rangle) - \langle a^i, y\rangle \exp(\lambda + \langle a^i, y\rangle) \\
&= (\lambda - 1) \sum_{i=0}^{m} \exp(\lambda + \langle a^i, y\rangle) = \lambda - 1 \\
&= -\log\left(\sum_{i=0}^{m} \exp\langle a^i, y\rangle\right) - 1.
\end{aligned}$$

Negating, we obtain the conjugacy formula (3.5.2).

(d) Compute the conjugate of the function of $x \in \mathbb{R}^{m+1}$,

$$\begin{cases} \sum_i \exp^*(x_i) & \text{if } \sum_i x_i = 1 \\ +\infty & \text{otherwise.} \end{cases}$$

This function can be expressed as $g(x)$ where $[a^0,\, a^1,\, \ldots\, a^m] = I \in \mathbb{R}^{(m+1)\times(m+1)}$. Therefore, the conjugate is

$$1 + \log\left(\sum_{i=0}^{m} \exp y_i\right).$$

8. Prove the Fenchel-Young inequality.

Proof. $f^*(y) = \sup_x \langle y, x\rangle - f(x)$ implies that for any x,

$$f^*(y) \geq \langle y, x\rangle - f(x) \implies f(x) + f^*(y) \geq \langle y, x\rangle.$$

$\square$

9 * (Fenchel duality and convex calculus). Fill in the details for the proof of Theorem 3.3.5 as follows.

(a) Prove the weak duality inequality.
The weak duality inequality is

$$p = \inf_{x\in\mathbb{E}} f(x) + g(Ax) \geq \sup_{\phi\in\mathbb{Y}} -f^*(A^*\phi) - g^*(-\phi) = d.$$

By Fenchel duality, $f(x)+f^*(A^*\phi) \geq \langle x, A^*\phi\rangle$ and $g(Ax)+g^*(-\phi) \geq \langle Ax, -\phi\rangle$. Adding these together gives

$$f(x) + f^*(A^*\phi) + g(Ax) + g^*(-\phi) \geq 0.$$

In other words, for any $x \in \mathbb{E}$ and $\phi \in \mathbb{Y}$, $f(x) + g(Ax) \geq -f^*(A^*\phi) - g(-\phi)$, i.e. $p \geq d$.

(b) Prove the inclusion (3.5.7).
Let $\phi_f \in \partial f(x)$ and $\phi_g \in \partial g(Ax)$. Then, for any $y \in \mathbb{E}$,

$$\langle \phi_f + A^*\phi_g, y - x\rangle = \langle \phi_f, y - x\rangle + \langle \phi_g, Ay - Ax\rangle \leq f(y) - f(x) + g(Ay) - g(Ax).$$

This implies that $\phi_f + A^*\phi_g \in \partial(f + g\circ A)(x)$. Thus, $\partial f(x) + A^*\partial g(Ax) \subset \partial(f + g\circ A)(x)$.

Now assume f and g are convex.

(c) Prove the function h defined in the proof is convex with domain $\operatorname{dom} g - A \operatorname{dom} f$. For reference, h was defined as $h(u) = \inf_{x\in\mathbb{E}} f(x) + g(Ax + u)$. Let $(u_1, r_1), (u_2, r_2)$ be in $\operatorname{epi} h$. Let $\lambda \in [0, 1]$. Then there exist x_1, x_2 such that $f(x_1) + g(Ax_1 + u_1) \leq r_1$ and $f(x_2) + g(Ax_2 + u_2) \leq r_2$. Thus

$$\begin{aligned} &f(\lambda x_1 + (1-\lambda)x_2) + g(A(\lambda x_1 + (1-\lambda)x_2) + \lambda u_1 + (1-\lambda)u_2) \\ &\quad \leq \lambda f(x_1) + (1-\lambda)f(x_2) + \lambda g(Ax_1 + u_1) + (1-\lambda)g(Ax_2 + u_2) \leq \lambda r_1 + (1-\lambda)r_2. \end{aligned}$$

Therefore, $\lambda(u_1, r_1) + (1-\lambda)(u_2, r_2) \in \operatorname{epi} h$. Therefore, h is convex.
Note $u \in \operatorname{dom} h \iff \exists x \in \mathbb{E},\ x \in \operatorname{dom} f \wedge Ax + u \in \operatorname{dom} g$. I.e., $u \in \operatorname{dom} g \iff \exists x \in \operatorname{dom} f,\ u \in \operatorname{dom} g - Ax \iff u \in \operatorname{dom} g - A \operatorname{dom} f$.

(d) Prove the implication (3.5.6) $\implies$ (3.5.5).
We wish to prove $A \operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset \implies 0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$. Suppose $y \in A \operatorname{dom} f \cap \operatorname{cont} g$. Let $d \in \mathbb{E}$. Since g is continuous and finite at y, given any fixed $\epsilon > 0$, for all $t \geq 0$ smaller than some positive number, $|g(y + td) - g(y)| \leq \epsilon$, implying $y + td \in \operatorname{dom} g$. Thus, since $y \in A \operatorname{dom} f$, we have $td \in \operatorname{dom} g - A \operatorname{dom} f$. herefore, $0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$.

Finally, assume in addition that condition (3.5.5) holds.

(e) Suppose $\phi \in \partial(f + g\circ A)(\bar{x})$. Use the first part of the theorem and the fact that $\bar{x}$ is an optimal solution of the problem

$$\inf_{x\in\mathbb{E}} \{(f(x) - \langle \phi, x\rangle) + g(Ax)\}$$

to deduce equality in part (b).
We have for any $x \in \mathbb{E}$,

$$\langle \phi, x - \bar{x}\rangle \leq f(x) + g(Ax) - f(\bar{x}) - g(A\bar{x}), \quad \text{i.e.} \quad f(\bar{x}) + g(A\bar{x}) - \langle \phi, \bar{x}\rangle \leq f(x) + g(Ax) - \langle \phi, x\rangle,$$

which shows that $\bar{x}$ is an optimal solution to $\inf_{x\in\mathbb{E}} f(x) - \langle \phi, x\rangle + g(Ax)$. It's worth noting how this observation relates to Fenchel conjugates; if $\phi \in \partial f(\bar{x})$, then $f^*(\phi)$ is obtained by plugging $\bar{x}$ into $\langle \phi, x\rangle - f(x)$.
Since $\bar{x}$ is a solution to the above primal problem, by the first part of the theorem, there is a dual variable which makes the dual objective equal the primal objective with $\bar{x}$

plugged in. First, let us compute $(f(\cdot)+\langle\phi,\cdot\rangle)^*(y) = \sup_x \langle y+\phi, x\rangle - f(x) = f^*(y+\phi)$. Thus, there exists $\phi' \in \mathbb{Y}$ such that

$$f(\bar{x}) - \langle \phi, \bar{x}\rangle + g(Ax) = -f^*(A\phi' + \phi) - g(-\phi').$$

Rearranging and adding and subtracting $\langle A^*\phi, \bar{x}\rangle = \langle \phi, Ax\rangle$,

$$f^*(A^*\phi' + \phi) + g^*(-\phi') = \langle A^*\phi' + \phi, \bar{x}\rangle - f(\bar{x}) + \langle -\phi', Ax\rangle - g(Ax).$$

By definition, $f^*(A^*\phi' + \phi) \geq \langle A^*\phi' + \phi, \bar{x}\rangle - f(\bar{x})$ and $g^*(-\phi') \geq \langle -\phi', A\bar{x}\rangle - g(A\bar{x})$, so we must have equality in both cases. Therefore, $\langle A^*\phi' + \phi, \bar{x}\rangle - f(\bar{x}) \geq \langle A^*\phi' + \phi, x\rangle - f(x)$ for every $x \in \mathbb{E}$, i.e. $A^*\phi' + \phi \in \partial f(\bar{x})$. Similarly, $-\phi' \in \partial g(A\bar{x})$. Thus, $\phi = A^*\phi' + \phi - A^*\phi' \in \partial f(\bar{x}) + A^*\partial g(Ax)$, showing the other inclusion. Therefore $\partial(f + g\circ A)(\bar{x}) = \partial f(\bar{x}) + A^*\partial g(A\bar{x})$, as long as $0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$.

(f) Prove points $\bar{x} \in \mathbb{E}$ and $\bar{\phi} \in \mathbb{Y}$ are optimal for problems (3.5.3) and (3.5.4), respectively, if and only if they satisfy the conditions $A^*\bar{\phi} \in \partial f(\bar{x})$ and $-\bar{\phi} \in \partial g(A\bar{x})$.
Since $\phi \in \partial f(x)$ iff $f^*(\phi) = \langle \phi, x\rangle - f(x)$, if $A^*\bar{\phi} \in \partial f(\bar{x})$ and $-\bar{\phi} \in \partial g(A\bar{x})$, then $f^*(A^*\bar{\phi}) = \langle A^*\phi, \bar{x}\rangle - f(\bar{x})$ and $g^*(-\bar{\phi}) = \langle -\bar{\phi}, Ax\rangle - g(A\bar{x})$. Adding these together and rearranging, we obtain

$$p \leq f(\bar{x}) + g(A\bar{x}) = -f^*(A^*\bar{\phi}) - g^*(-\bar{\phi}) \leq d \leq p.$$

This implies $\bar{x}$ achieves the primal optimum value and $\bar{\phi}$ achieves the dual optimum value.
Now suppose $\bar{x}$ and $\bar{\phi}$ are a pair of optima for the primal and dual. Then,

$$\begin{aligned} & f(\bar{x}) + g(A\bar{x}) = -f^*(A^*\bar{\phi}) - g^*(-\bar{\phi}) \\ \implies\ & f^*(A^*\bar{\phi}) + g^*(-\bar{\phi}) = \langle A^*\bar{\phi}, \bar{x}\rangle - f(\bar{x}) + \langle -\bar{\phi}, Ax\rangle - g(A\bar{x}), \end{aligned}$$

which by similar reasoning to the previous part implies $f^*(A^*\bar{\phi}) = \langle A^*\bar{\phi}, \bar{x}\rangle - f(\bar{x})$, and $g^*(-\bar{\phi}) = \langle -\bar{\phi}, Ax\rangle - g(A\bar{x})$, and subsequently $A^*\bar{\phi} \in \partial f(\bar{x})$ and $-\bar{\phi} \in \partial g(A\bar{x})$.

10 (Normals to an intersection). If the point x lies in two convex subsets C and D of $\mathbb{E}$ satisfying $0 \in \operatorname{core}(C - D)$ (or in particular $C \cap \operatorname{int} D \neq \emptyset$), use Section 3.1, Exercise 4 (Subgradients and normal cones) to prove

$$N_{C\cap D}(x) = N_C(x) + N_D(x).$$

If $0 \in \operatorname{core}(C - D)$, then $0 \in \operatorname{core}(\operatorname{dom}\delta_C - \operatorname{dom}\delta_D)$. Furthermore, the functions δ_C and δ_D are convex. Therefore, by Theorem 3.3.5 (3.14),

$$\partial(\delta_C + \delta_D)(\cdot) = \partial\delta_C(\cdot) + \partial\delta_D(\cdot).$$

Note that $\delta_C + \delta_D = \delta_{C\cap D}$. By Section 3.1, Exercise 4, for any set S, $\partial\delta_S(\bar{x}) = N_S(\bar{x})$. Thus,

$$N_{C\cap D}(\cdot) = N_C(\cdot) + N_D(\cdot).$$

11 * (Failure of convex calculus).

(a) Find convex functions $f, g : \mathbb{R} \to (-\infty, +\infty]$ with

$$\partial f(0) + \partial g(0) \neq \partial(f+g)(0).$$

(Hint: Section 3.1, Exercise 5.)
Consider the function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x = 0 \\ +\infty & \text{otherwise.} \end{cases}.$$

$\partial f(0) = \emptyset$, since for any $\phi \in \mathbb{R}$, $\phi x > -1 = f(x) - f(0)$ if we choose $x < 0$ with small enough magnitude. However, $\delta_{\{0\}} + f = \delta_{\{0\}} + 1$. The RHS has subgradient $\mathbb{R}$ at 0. This gives a contradiction.

(b) Find a convex function $g : \mathbb{R}^2 \to (-\infty, +\infty]$ and a linear map $A : \mathbb{R} \to \mathbb{R}^2$ with $A^*\partial g(0) \neq \partial(g \circ A)(0)$.
Consider a 2d version of the function in the previous part: $g(x, y) = f(x)$. Take $\phi \in \mathbb{E}$. By taking (x, y) with x negative and $\|(x, y)\|^2$ small enough,

$$\langle \phi, (x, y) \rangle > -1 = g(x, y) - g(0).$$

Therefore, $\partial g(0) = \emptyset$. So if $A = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then $A^*\partial g(0) = \emptyset$. On the other hand, $g \circ A = g(0) = 1$, a constant. Therefore, $\partial(g \circ A)(0) = \{0\}$, which does not equal the empty set.

12 * (Infimal convolution). If the functions $f, g : \mathbb{E} \to (-\infty, +\infty]$ are conex, we define the *infimal convolution* $f \odot g : \mathbb{E} \to [-\infty, +\infty]$ by

$$(f \odot g)(y) = \inf_x \{f(x) + g(y - x)\}.$$

(a) Prove $f \odot g$ is convex. (On the other hand, if g is concave prove so is $f \odot g$.)
Let $y_1, y_2 \in \mathbb{E}$ and $\lambda \in [0, 1]$. For any $x_1, x_2 \in \mathbb{E}$,

$$\begin{aligned} &f(\lambda x_1 + (1-\lambda)x_2) + g(\lambda y_1 + (1-\lambda)y_2 - \lambda x_1 + (1-\lambda)x_2) \\ &\qquad \leq \lambda[f(x_1) + g(y_1 - x_1)] + (1-\lambda)[f(x_2) + g(y_2 - x_2)]. \end{aligned}$$

Taking the infimum over x_1 and x_2 on the RHS, we obtain $(f \odot g)(\lambda y_1 + (1-\lambda)y_2) \leq \lambda(f \odot g)(y_1) + (1-\lambda)(f \odot g)(y_2)$.
Now suppose g is concave. For any $x \in \mathbb{E}$, we have

$$\begin{aligned} f(x) + g(\lambda y_1 + (1-\lambda)y_2 - x) &\geq \lambda(f(x) + g(y_1 - x)) + (1-\lambda)(f(x) + g(y_2 - x)) \\ &\geq \lambda(f \odot g)(y_1) + (1-\lambda)(f \odot g)(y_2), \end{aligned}$$

which after taking an infimum on the LHS implies that $(f \odot g)(\lambda y_1 + (1-\lambda)y_2) \geq \lambda(f \odot g)(y_1) + (1-\lambda)(f \odot g)(y_2)$. Notice that we didn't need f to be convex or concave.

(b) Prove $(f \odot g)^* = f^* + g^*$.

$$
\begin{aligned}
(f \odot g)^*(z) &= \sup_{y\in\mathbb{E}} \left\{ \langle z, y\rangle - \inf_{x\in\mathbb{E}}\{f(x) + g(y-x)\} \right\} \\
&= \sup_{y,x\in\mathbb{E}} \{\langle z, y\rangle - f(x) - g(y-x)\} \\
&= \sup_{y,x\in\mathbb{E}} \{\langle z, y-x\rangle - g(y-x) + \langle z, x\rangle - f(x)\} \\
&= g^*(z) + f^*(z),
\end{aligned}
$$

The last step follows by taking the supremum over x and then the supremum over y.

(c) If $\operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset$, prove $(f+g)^* = f^* \odot g^*$.
Is there a reason why the condition $0 \in \operatorname{core}(\operatorname{dom} f - \operatorname{dom} g)$ is not used here? Theorem 3.3.5 (3.14) can be rephrased as saying that if $0 \in \operatorname{core}(\operatorname{dom} g - \operatorname{dom} f)$ or in particular $\operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset$, then

$$
\sup_{x\in\mathbb{E}}\{-f(x) - g(x)\} = \inf_{\phi\in\mathbb{E}}\{f^*(\phi) + g^*(\phi)\}.
$$

Noting that $\operatorname{dom}(f - \langle y, \cdot\rangle) = \operatorname{dom}(f)$ for any $y \in \mathbb{E}$, and that $(f - \langle y, \cdot\rangle)^*(\phi) = f^*(\phi + y)$, we have

$$
\sup_{x\in\mathbb{E}}\{\langle y, x\rangle - f(x) - g(x)\} = \inf_{\phi\in\mathbb{E}}\{f^*(\phi + y) + g^*(-\phi)\}.
$$

But the LHS is $(f+g)^*(y)$ and the RHS is $(f^* \odot g^*)(y)$.

(d) Given a nonempty set $C \subset \mathbb{E}$, define the *distance function* by

$$
d_C(x) = \inf_{y\in C} \|x - y\|.
$$

(i) Prove d_C^2 is a difference of convex functions, by observing

$$
(d_C(x))^2 = \frac{\|x\|^2}{2} - \left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x).
$$

$(\inf_{y\in C}\|x-y\|)^2 = \inf_{y\in C}\|x-y\|^2$, because $\|a\| \geq \|b\| \iff \|a\|^2 \geq \|b\|^2$. Therefore, $d_C(x)^2 = \inf_{y\in C}\|x-y\|^2 = \inf_{y\in C}\{\|x\|^2 - 2\langle x, y\rangle + \|y\|^2\}$. Continuing,

$$
\begin{aligned}
d_C(x)^2 &= \|x\|^2 + \inf_{y\in C}\{-2\langle x, y\rangle + \|y\|^2\} \\
&= \|x\|^2 - 2\sup_{y\in C}\{\langle x, y\rangle - \frac{\|y\|^2}{2}\} \\
&= \|x\|^2 - 2\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x).
\end{aligned}
$$

I think the LHS in the question is missing a factor of $\frac{1}{2}$.

Now suppose C is convex.

(ii) Prove d_C is convex and $d_C^* = \delta_B + \delta_C^*$.
For any $y_1, y_2 \in C$ and $\lambda \in [0,1]$, the point $\lambda y_1 + (1-\lambda)y_2$ is also in C, by convexity. Also,

$$\|\lambda x_1 + (1-\lambda)x_2 - \lambda y_1 - (1-\lambda)y_2\| \leq \lambda\|x_1 - y_1\| + (1-\lambda)\|x_2 - y_2\|.$$

This implies that $d_C(\lambda x_1 + (1-\lambda)x_2) \leq \lambda d_C(x_1) + (1-\lambda)d_C(x_2)$, i.e. d_C is convex. Now for the conjugate,

$$d_C^*(z) = \sup_{x\in\mathbb{E}}\left\{\langle z, x\rangle - \inf_{y\in C}\|x-y\|\right\} = \sup_{x\in\mathbb{E}, y\in C}\langle z, x\rangle - \|x-y\|.$$

To handle $\|z\| > 1$, notice $\langle z, x\rangle - \|x-y\| \geq \langle z, x\rangle - \|x\| - \|y\|$. If we pick x in the same direction as z, this is equal to $(\|z\|-1)\|x\| - \|y\| \to +\infty$ as $\|x\| \to +\infty$. This explains the presence of the δ_B. Otherwise, $\|z\| \leq 1$. Then, $\langle z, x-y\rangle - \|x-y\| \leq \|z\|\|x-y\| - \|x-y\| \leq 0$, which implies $\langle z, x\rangle - \|x-y\| \leq \langle z, y\rangle$. Thus, we can restrict the sup to $x = y$, giving

$$d_C^*(z) = \sup_{y\in C}\langle z, y\rangle.$$

This proves $d_C^*(z) = \delta_B + \delta_C^*$.

(iii) For $x \in C$ prove $\partial d_C(x) = B \cap N_C(x)$.
For the inclusion $\partial d_C(x) \subset B \cap N_C(x)$, suppose $\phi \in \partial d_C(x)$. Then for any $z \in \mathbb{E}$ and $y \in \mathbb{E}$,

$$\langle \phi, z - x\rangle \leq \inf_{y\in C}\|z-y\| - \inf_{y\in C}\|x-y\| \leq \|z-y\| - \inf_{y\in C}\|x-y\|.$$

Since C is nonempty ($x \in C$), for any $\epsilon > 0$ there exists $y \in \mathbb{E}$ such that $\|x-y\| \leq \inf_{y\in C}\|x-y\| + \epsilon$. Therefore,

$$\langle \phi, z-x\rangle \leq \|z-y\| - \|x-y\| + \epsilon \leq \|z-x\| + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, $\langle \phi, z-x\rangle \leq \|z-x\|$. By taking $z = x + \phi$, this implies $\phi \in B$. Finally, by the subgradient property, for $y \in C$,

$$\langle \phi, y - x\rangle \leq -\inf_{y\in C}\|x-y\| \leq 0.$$

Wait, we were supposed to assume $x \in C$ which implies $d_C(x) = 0$, but I think this proves $\partial d_C(x) \subset B \cap N_C(x)$ without assuming $x \in C$. Now we show $B \cap N_C(x) \subset \partial d_C(x)$. Suppose $\phi \in B \cap N_C(x)$. For any other $y \in C$,

$$\langle \phi, y-x\rangle \leq 0 = d_C(y) - d_C(x),$$

since $d_C(y) = d_C(x) = 0$. If $z \in \mathbb{E}$ in general, then since $\langle \phi, y-x\rangle \leq 0$,

$$\langle \phi, z-x\rangle \leq \langle \phi, z-y\rangle \leq \|z-y\|.$$

By taking the inf over y, we get $\langle z - x\rangle \leq d_C(z)$, implying $\phi \in \partial d_C(x)$.

(iv) If C is closed and $x \notin C$, prove

$$\nabla d_C(x) = d_C(x)^{-1}(x - P_C(x)),$$

where $P_C(x)$ is the nearest point to x in C.
First we show that $\frac{x-P_C(x)}{d_C(x)}$ is a subgradient. Let $z \in \mathbb{E}$. By the nearest point characterization, $\langle x - P_C(x), P_C(x) - P_C(z)\rangle \geq 0$. Therefore,

$$\begin{aligned}\left\langle \frac{x - P_C(x)}{d_C(x)}, z - x \right\rangle &\leq \left\langle \frac{x - P_C(x)}{d_C(x)}, z - P_C(z) + P_C(x) - x \right\rangle \\ &\leq \|z - P_C(z)\| - \|x - P_C(x)\| = d_C(z) - d_C(z).\end{aligned}$$

The second inequality follows because $\|\frac{x-P_C(x)}{d_C(x)}\| = 1$ and $\langle x - P_C(x), x - P_C(x)\rangle / d_C(x) = d_C(x)$. Now let us show this must be the only subgradient.
Suppose $\phi \in \partial d_C(x)$. First we show that $\phi \in B$.

$$\langle \phi, z - x\rangle = \|z - P_C(z)\| - \|x - P_C(x)\| \leq \|z - P_C(x)\| - \|x - P_C(x)\| \leq \|z - x\|.$$

Since z was arbitrary, $\|\phi\| \leq 1$. Now we conclude ϕ must equal $x - P_C(x)$, scaled to unit norm. Applying the subgradient inequality to $P_C(x)$, we obtain

$$\langle \phi, P_C(x) - x\rangle \leq -\|x - P_C(x)\|.$$

By Cauchy-Schwarz, the only way this is possible while $\|\phi\| \leq 1$ is if $\phi = \frac{x-P_C(x)}{d_C(x)}$.

(v) If C is closed, prove

$$\nabla \frac{d_C^2}{2}(x) = x - P_C(x)$$

for all points x.
Recall the formula $\frac{d_C^2}{2}(x) = \frac{\|x\|^2}{2} - \left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x)$. If we can show that $\frac{\|x\|^2}{2}$ and $-\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)$ are differentiable, then we can simply compute their gradients and add them. By calculus, $\nabla \frac{\|x\|^2}{2} = x$.
To compute the other gradient, we make two general observations about subgradients and conjugates. First of all,

$$f^*(\phi) = \langle \phi, x\rangle - f(x) \iff \phi \in \partial f(x). \tag{3.6.1}$$

This is because $\forall x' \in \mathbb{E},\ \langle \phi, x\rangle - f(x) \geq \langle \phi, x'\rangle - f(x') \iff \forall x' \in \mathbb{E},\ \langle \phi, x' - x\rangle \leq f(x') - f(x)$. In other words, the argmax of the expression whose sup defines the conjugate is the set of points at which ϕ is a subgradient of f.
Next, if $\partial f(x) \neq \emptyset$, then

$$x \in \partial f^*(y) \iff y \in \partial f(x). \tag{3.6.2}$$

First, suppose $x \in \partial f^*(y)$. Then, $\forall y',\ \langle x, y' - y\rangle \leq f^*(y') - f^*(y)$. Because of (3.6.1) and $\partial f(x) \neq \emptyset$, we can take $y' = \bar{y} \in \partial f(x)$, giving $\forall x',\ \langle x, \bar{y} - y\rangle \leq$

$\langle \bar{y}, x\rangle - f(x) - \langle y, x'\rangle + f(x')$. So, $\forall x'$, $\langle y, x' - x\rangle \leq f(x') - f(x)$, i.e. $y \in \partial f(x)$. On the other hand, if $y \in \partial f(x)$, then since for any y', $\langle x, y'\rangle \leq f^*(y') + f(x)$ and by (3.6.1), $-\langle x, y\rangle = -f^*(y) - f(x)$, we have $\langle x, y' - y\rangle \leq f^*(y') - f^*(y)$, i.e. $x \in \partial f^*(y)$.

Now we can compute $\partial\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x)$. We have

$$\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x) = \sup_{y\in C}\langle x, y\rangle - \frac{\|y\|^2}{2} = \frac{\|x\|^2}{2} - \frac{\|x - P_C(x)\|^2}{2}.$$

i.e. $y = P_C(x)$ uniquely achieves the sup (see Section 2.1, Exercise 8, which proves the nearest point in a closed convex set exists and is unique). By (3.6.1), $x \in \partial\left(\frac{\|\cdot\|^2}{2} + \delta_C(\cdot)\right)(P_C(x))$. By (3.6.2), $\{P_C(x)\} = \partial\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x)$, where $P_C(x)$ is the only point in the subgradient because it uniquely achieves the above sup. Therefore, having a unique subgradient, $-\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*$ is differentiable, and

$$\partial\frac{d_C^2}{2}(x) = \nabla\frac{\|x\|^2}{2} - \nabla\left(\frac{\|\cdot\|^2}{2} + \delta_C\right)^*(x) = x - P_C(x).$$

I bashed out the solution manually previously. This approach, which applies seemingly general facts of subgradients/conjugates, seems much better. Although it might be useful to write down the bashy approach to compare.

(e) Define the *Lambert W-function* $W : \mathbb{R}_+ \to \mathbb{R}_+$ as the inverse of $y \in \mathbb{R}_+ \mapsto ye^y$. Prove the conjugate of the function

$$x \in \mathbb{R} \mapsto \exp^*(x) + \frac{x^2}{2}$$

is the function

$$y \in \mathbb{R} \mapsto W(e^y) + \frac{(W(e^y))^2}{2}.$$

It's really not hard to solve this by calculating $\sup_{x\in\mathbb{R}_+} yx - x\log x + x - \frac{x^2}{2}$ by taking the derivative, but in the interest of the theme of this question and trying to use general theory we'll use a different particular approach.

We notice that by part (b), $\exp^*(y) + \frac{y^2}{2} = \left(\exp \odot \frac{(\cdot)^2}{2}\right)^*(x)$. Let us compute $\left(\exp \odot \frac{(\cdot)^2}{2}\right)(y) =: f(y)$.

$$f(y) = \sup_{x\in\mathbb{R}}\left\{\exp(x) + \frac{(y-x)^2}{2}\right\}.$$

Solving for the sup by setting the derivative to 0, $\exp(x) + x - y = 0 \implies \exp(x) = y - x$. Therefore, if x is optimal for the sup, then $f(y) = \exp(x) + \frac{\exp(x)^2}{2}$. By $\exp(x)\exp(\exp(x)) = \exp(y)$, we obtain $\exp(x) = W(\exp(y))$. Thus, $f(y) = W(\exp(y)) + \frac{W(\exp(y))^2}{2}$.

If we knew that $\left(\exp \odot \frac{(\cdot)^2}{2}\right)^{**} = \exp \odot \frac{(\cdot)^2}{2}$, we would be done. We'll prove this by showing that for any convex $f : \mathbb{E} \to (-\infty, +\infty]$ with $x \in \operatorname{core}(\operatorname{dom} f)$, $f(x) = f^{**}(x)$.

$$f^{**}(x) = \sup_{y \in \mathbb{E}} \langle x, y \rangle - f^*(y) = \sup_{y \in \mathbb{E}} \langle x, y \rangle - \sup_{z \in \mathbb{E}} \langle y, z \rangle - f(z).$$

Recall (3.6.2); since $\partial f(x)$ is nonempty by Theorem 3.1.8 (Max formula), if $y \in \partial f(x)$, then applying (3.6.1) to both sups,

$$f^{**}(x) = \langle x, y \rangle - \langle y, x \rangle + f(x) = f(x).$$

Since the function $W(e^y) + \frac{W(e^y)^2}{2}$ is finite everywhere, its biconjugate equals itself.

13 * (Applications of Fenchel duality).

(a) **(Sandwich theorem).** Let the functions $f : \mathbb{E} \to (-\infty, +\infty]$ and $g : \mathbb{Y} \to (-\infty, +\infty]$ be convex and the map $A : \mathbb{E} \to \mathbb{Y}$ be linear. Suppose $f \geq -g \circ A$ and $0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$ (or $A \operatorname{dom} f \cap \operatorname{cont} g \neq \emptyset$). Prove there is an affine function $\alpha : \mathbb{E} \to \mathbb{R}$ satisfying $f \geq \alpha \geq -g \circ A$.
By Fenchel duality (Theorem 3.3.5),

$$\inf_{x \in \mathbb{E}} f(x) + g(Ax) = \sup_{\phi \in \mathbb{Y}} -f^*(A^*\phi) - g^*(-\phi).$$

Since $f \geq -g \circ A$, we have $f(x) + g(Ax) \geq 0$ for all $x \in \mathbb{E}$. Therefore, both sides of the equation are finite and in particular at least 0, which again by Fenchel duality implies the supremum is attained by some $\phi \in \mathbb{Y}$. In other words,

$$\inf_{x \in \mathbb{E}} f(x) + g(Ax) = \inf_{x \in \mathbb{E}} \{f(x) - \langle A^*\phi, x \rangle\} + \inf_{y \in \mathbb{Y}} \{g(y) + \langle \phi, y \rangle\} \geq 0.$$

Thus, for any $\bar{x} \in \mathbb{E}$,

$$\begin{aligned} &\langle A^*\phi, \bar{x} \rangle + \inf_{x \in \mathbb{E}} \{f(x) - \langle A^*\phi, x \rangle\} \leq f(\bar{x}) \\ &\langle -\phi, A\bar{x} \rangle + \inf_{y \in \mathbb{Y}} \{g(y) + \langle \phi, y \rangle\} \leq \langle -\phi, A\bar{x} \rangle + \inf_{x \in \mathbb{E}} \{g(Ax) + \langle \phi, Ax \rangle\} \leq g(A\bar{x}). \end{aligned}$$

Negating the second inequality gives

$$\langle A^*\phi, \bar{x} \rangle - \inf_{y \in \mathbb{Y}} \{g(y) + \langle \phi, y \rangle\} \geq -g(A\bar{x}).$$

Since $\inf_x \{f(x) - \langle A^*\phi, x \rangle\} \geq -\inf_y \{g(y) + \langle \phi, y \rangle\}$, with $C = \inf_x \{f(x) - \langle A^*\phi, x \rangle\}$, we have for any $\bar{x} \in \mathbb{E}$

$$-g(A\bar{x}) \leq \langle A^*\phi, \cdot \rangle + C \leq f(\bar{x}),$$

which proves the desired statement.

(b) Interpret the Sandwich theorem geometrically in the case when A is the identity.
It means we can fit a hyperplane between the graphs of a convex function and a concave function, assuming the former upper bounds the latter.

(c) **(Pshenichii-Rockafellar conditions [159]).** If the convex set C in $\mathbb{E}$ satisfies the condition $C \cap \operatorname{cont} f \neq \emptyset$ (or the condition $\operatorname{int} C \cap \operatorname{dom} f \neq \emptyset$), and if f is bounded below on C, use part (a) to prove there is an affine function $\alpha \leq f$ with $\inf_C f = \inf_C \alpha$. Deduce that a point $\bar{x}$ minimizes f on C if and only if it satisfies $0 \in \partial f(\bar{x}) + N_C(\bar{x})$.
Define $M = \inf_{x \in C} f(x)$. $M \in \mathbb{R}$, because f is bounded below on C and $C \cap \operatorname{dom} f \neq 0$. Notice for all $x \in C$, $f(x) - M \geq 0$, and $\inf_{x \in C}\{f(x) - M\} = 0$. By the former, the convex function δ_C upper bounds $M - f$. By the condition $C \cap \operatorname{cont} f \neq \emptyset$, we may apply part (a) to deduce the existence of an affine function α such that

$$\delta_C \geq \alpha \geq M - f.$$

Therefore

$$0 = \sup_{x \in C} \delta_C(x) \geq \sup_{x \in C} \alpha(x) \geq \sup_{x \in C}\{M - f\} = 0.$$

So, the affine function $M - \alpha$ satisfies $M - \alpha \leq f$ and $\inf_{x \in C}\{M - \alpha(x)\} = M - \sup_{x \in C} \alpha(x) = M = \inf_{x \in C} f(x)$, as desired.
Let $\alpha(x) = \langle \phi, x \rangle + b$ with $\inf_C \alpha = \inf_C f$. If $\bar{x}$ is a minimizer of f over C, then $\langle \phi, \bar{x} \rangle + b = f(\bar{x})$, and thus $b = f(\bar{x}) - \langle \phi, \bar{x} \rangle$. As a result, $\phi \in \partial f(\bar{x})$. Furthermore, $x' \in C$ implies $\alpha(x') - \alpha(\bar{x}) \geq 0$, i.e. $\langle \phi, x' - \bar{x} \rangle \geq 0$. Therefore, $-\phi \in N_C(\bar{x})$. As a result, $0 \in \partial f(\bar{x}) + N_C(\bar{x})$.

(d) Apply part (c) to the following two cases:

(i) C a single point $\{x^0\} \subset \mathbb{E}$
Part (c) for this case simply says that if $x^0 \in \operatorname{cont} f$, then $\bar{f}(x^0) \neq \emptyset$.

(ii) C a polyhedron $\{x \mid Ax \leq b\}$, where $b \in \mathbb{R}^n = \mathbb{Y}$
If $\operatorname{cont} f \cap C \neq \emptyset$, then there exist $c \in \mathbb{R}$, $\phi \in \mathbb{E}$ such that

$$\inf_{x : Ax \leq b} f(x) = \inf_{x : Ax \leq b} \langle \phi, w \rangle + c.$$

Is the conclusion something like $\phi = -A^\top \lambda$ for some $\lambda \in \mathbb{R}^n_+$ with nonzero entries in the entries where $Ax = b$? I am not sure this is the conclusion I was supposed to derive.

(e) **(Hahn-Banach extension)** If the function $f : \mathbb{E} \to \mathbb{R}$ is everywhere finite and sublinear, and for some linear subspace L of $\mathbb{E}$ the function $h : L \to \mathbb{R}$ is linear and *dominated* by f (in other words $f \geq h$ on L), prove there is a linear function $\alpha : \mathbb{E} \to \mathbb{R}$, dominated by f, which agrees with h on L.
Define the function $h_{-\infty} : \mathbb{E} \to \mathbb{R}$ by $h_{-\infty} = h$ on L and otherwise $h_{-infty} = -\infty$. This function is concave and dominated by f. Furthermore, being sublinear, f is continuous everywhere (examine the maximum absolute value of f over a basis). Therefore, by the sandwich theorem (part (a)), there exists an affine α such that

$$f \geq \alpha \geq h_{-\infty}.$$

Using sublinearity of f and linearity of h, $f(0) = h_{-\infty} = 0$, so we have $\alpha = 0$, i.e. α is linear. Furthermore, for $x \in L$, $\alpha(x) \geq h(x)$ and $-\alpha(x) = \alpha(-x) \geq h(-x) = -h(x)$, implying $\alpha = h$ on L. This completes the proof.

14. Fill in the details of the proof of the Krein-Rutman calculus (3.17).

Proof. Note that for any cone K, $K^- = \partial\delta_K(0)$. Further note that $\delta_K + \delta_H \circ A = \delta_{K\cap A^{-1}H}$, and that $K \cap A^{-1}H$ is a cone. Finally, note that if $H - AK = \mathbb{Y}$, then $0 \in \operatorname{core}(H - AK) = \operatorname{core}(\operatorname{dom}\delta_H - A\operatorname{dom}\delta_K)$, so that we can apply the subgradient equality in Theorem 3.3.5 (Fenchel duality):

$$\partial\delta_{K\cap A^{-1}H}(0) = \partial(\delta_K + \delta_H \circ A)(0) = \partial\delta_K(0) + A^*\partial\delta_H(0).$$

Since for any cone $K, K^- = \delta_K(0)$, we have

$$(K \cap A^{-1}H)^- = A^*H^- + K^-.$$

□

15 * (Bipolar theorem) For any nonempty set $K \subset \mathbb{E}$, prove the set $\operatorname{cl}(\operatorname{conv}(\mathbb{R}_+K))$ is the smallest closed convex cone containing K. Deduce Theorem 3.3.14 (Bipolar cones) (3.18).
Clearly, any cone containing K contains $\mathbb{R}_+K$. Any convex cone containing K contains $\mathbb{R}_+K$ and is convex, and thus contains the smallest convex set containing $\mathbb{R}_+K$, $\operatorname{conv}(\mathbb{R}_+K)$ (see Section 1.1 Exercise 2). Therefore, any closed convex cone containing K contains $\operatorname{conv}(\mathbb{R}_+K)$ and is closed. Therefore, it contains the smallest closed set containing $\operatorname{conv}(\mathbb{R}_+K)$, which is $\operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$.
We have shown that any closed convex cone containing K contains $\operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$. It remains to show that $\operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$ is a closed convex cone. By Section 1.1, Exercise 3, this set is closed and convex. It is a cone, first because $\mathbb{R}_+K$ is a cone, then because $\operatorname{conv}(\mathbb{R}_+K)$ is a cone ($c \in \mathbb{R}_+,\ \sum_i \lambda_i c_i k^i \in \operatorname{conv}(\mathbb{R}_+K) \implies c\sum_i \lambda_i c_i k^i = \sum_i \lambda_i c c_i k^i \in \operatorname{conv}(\mathbb{R}_+K)$), then because the closure of a cone is a cone ($y^i \to y \in C \implies c^i y^i \to cy \in C$ for $c \in \mathbb{R}_+$).
Finally we deduce Theorem 3.3.14 (Bipolar cones) (3.18). Let $y \notin \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$. We'll show $y \notin K^{--}$, which means $K^{--} \subset \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$. By basic separation, there exists $a \in \mathbb{E}$ not equal to 0 and $b \in \mathbb{R}$ such that for all $x \in \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$, $\langle a, y\rangle > b \geq \langle a, x\rangle$. Let's prove we can take $b = 0$. If any $x \in \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$ satisfies $\langle a, x\rangle > 0$, then by scaling x by a large enough positive constant, we contradict $b \geq \langle a, x\rangle$. So, if $b > 0$, then we can safely set it to 0. If $b < 0$, then setting $x = 0$ gives another contradiction. Therefore, we have $\langle a, y\rangle > 0 \geq \langle a, x\rangle$ for each $x \in \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$. The second inequality implies $a \in K^-$, after which the first implies $y \notin K^{--}$. The polar cone is a closed and convex cone ($c, \mu \geq 0,\ \langle \phi_1, x\rangle \leq 0,\ \langle \phi_2, x\rangle \leq 0 \implies \langle c\phi_1 + \mu\phi_2, x\rangle \leq 0$, and if $\phi^i \to \phi$, $\langle \phi^i, x\rangle \leq 0$, then $\langle \phi, x\rangle \leq 0$). Therefore, $K^{--} = \operatorname{cl}\operatorname{conv}(\mathbb{R}_+K)$, which is the content of Theorem 3.3.14 (Bipolar cones)(3.18).
16 * (Sums of closed cones)

(a) Prove that any cones $H, K \subset \mathbb{E}$ satisfy $(H + K)^- = H^- \cap K^-$.
Suppose $x \in H^- \cap K^-$. Then, for any $h \in H,\ k \in K$, we have $\langle x, h + k\rangle = \langle x, h\rangle + \langle x, k\rangle \leq 0$, so that $x \in (H+K)^-$. Conversely, if $x \in (H+K)^-$, we must have $\langle x, h\rangle \leq 0$ and $\langle x, k\rangle \leq 0$ for every $h \in H, k \in K$, since $0 \in H \cap K$. Thus, $x \in H^- \cap K^-$.

(b) Deduce that if H and K are closed convex cones then they satisfy $(H \cap K)^- = \operatorname{cl}(H^- + K^-)$, and prove that the closure can be omitted under the condition $K \cap \operatorname{int} H \neq \emptyset$.

By part (a) applied to H^-, K^-, $H^{--} \cap K^{--} = (H^- + K^-)^-$. Taking polars on both sides and applying Theorem 3.3.14 (3.18), we get $(H^{--} \cap K^{--})^- = \mathrm{cl}(H^- + K^-)$ (since $H^- + K^-$ is a convex cone already). By Theorem 3.3.14, $H^{--} = H$ and $K^{--} = K$. This completes the proof.
If $K \cap \mathrm{int}\, H \neq \emptyset$, then $H - K = \mathbb{E}$, as if $k \in K \cap \mathrm{int}\, H$ then for any $d \in \mathbb{E}$ there exists $h \in H$ such that $h - k = td$ for some $t > 0$, then $h/t - k/t = d$. By the polar calculus (Corollary 3.3.13 (3.17)), $(H \cap K)^- = H^- + K^-$.

In $\mathbb{R}^3$, define sets

$$H = \{x \mid x_1^2 + x_2^2 \leq x_3^4,\ x_3 \leq 0\} \text{ and}$$
$$K = \{x \mid x_2 = -x_3\}.$$

(c) Prove H and K are closed convex cones.
H is a cone, since $x_1^2 + x_2^2 \leq x_3^2 \implies c^2x_1^2 + c^2x_2^2 \leq c^2x_3^2$ for any $c \geq 0$, and $cx_3 \leq 0$ if $x_3 \leq 0$. It is closed since $f(x) = x_1^2 + x_2^2 - x_3^2$ is a continuous function, and thus $f^{-1}((-\infty, 0])$ is closed. It is convex because of the triangle inequality, which follows from Cauchy Schwarz:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2x^\top y \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$$
$$\implies \sqrt{(x_1+y_1)^2 + (x_2+y_2)^2} = \|x+y\| \leq \|x\| + \|y\| = \sqrt{x_1^2 + x_2^2} + \sqrt{y_1^2 + y_2^2}.$$

K is a linear subspace, which is closed, convex, and a cone.

(d) Calculate the polar cones H^-, K^-, and $(H \cap K)^-$.
Let $y \in -H$. If $x \in H$, then

$$\langle y, x\rangle \leq \sqrt{y_1^2 + y_2^2}\sqrt{x_1^2 + x_2^2} - |y_3||x_3| \leq 0,$$

which holds because $\sqrt{y_1^2 + y_2^2} \leq |y_3|$ and $\sqrt{x_1 + x_2^2} \leq |x_3|$ and these quantities are nonnegative. Thus, $-H \in H^-$. Now let $y \in H^-$. We have

$$\langle y, (y_1, y_2, -\|(y_1, y_2)\|)\rangle = \|(y_1, y_2)\|^2 - y_3\|(y_1, y_2)\| \leq 0$$

which if $(y_1, y_2) \neq 0$ implies $\|(y_1, y_2)\| \leq y_3$, i.e. $y \in -H$. If $(y_1, y_2) = 0$, we need $y_3 \geq 0$ or else we can take any $x \in H$ to get a positive dot product. Thus, $y \in H^- \implies y \in -H$. We conclude $-H = H^-$.
Since K is a linear subspace and $K^- = N_K(0)$, by Section 2.1, Exercise 2(c), $K^- = K^\perp$. By definition, $K = \mathrm{Span}\{(0,1,1)\}^\perp$. Since V a subspace of a finite dimensional linear space satisfies $(V^\perp)^\perp = V$, we have $K^- = \mathrm{Span}\{(0,1,1)\}$.
$H \cap K = \mathbb{R}_+(0,1,-1)$, since if $y_1 \neq 0$ and $y_2 = -y_3$ then $y_1^2 + y_2^2 > y_2^2 = y_3^2$. The polar is $\{x \in \mathbb{R}^3 : x_2 - x_3 \leq 0\}$, i.e. the half space passing through 0 with normal vector $(0,1,-1)$.

(e) Prove $(1,1,1) \in (H \cap K)^- \setminus (H^- + K^-)$, and deduce that the sum of two closed convex cones is not necessarily closed.

$(1,1,1)^\top(0,1,-1)=0$, so $(1,1,1)\in(H\cap K)^-$. Suppose $h^-\in H^-$ and $k^-=c(0,1,1)\in K^-$. Thus, if $h^-+k^-=(1,1,1)$, then $h^-=(1,1-c,1-c)$. But this is a contradiction, since $1+(1-c)^2>(1-c)^2$. Thus, $(1,1,1)\notin H^-+K^-$.
If H^-+K^- were closed, then by part (b) $(H\cap K)^-=\mathrm{cl}(H^-+K^-)=H^-+K^-$. Thus, H^-+K^- is not closed, despite H^- and K^- being closed convex cones.

17 * (Subdifferential of a max-function). With the notation of Section 3.2, Exercise 13, suppose

$$\operatorname{dom} g_j \cap \bigcap_{i\in I\setminus\{j\}} \operatorname{cont} g_i \neq \emptyset$$

for some index j in I. Prove

$$\partial(\max_i g_i)(\bar{x}) = \operatorname{conv}\bigcup_{i\in I}\partial g_i(\bar{x}).$$

From Section 3.2, Exercise 13, we have

$$\partial g(\bar{x}) = \bigcup\left\{\partial\left(\sum_{i\in I}\lambda_i g_i\right)(\bar{x})\,\middle|\,\lambda\in\mathbb{R}^I_+,\ \sum_{i\in I}\lambda_i=1\right\}.$$

Note that for any $c>0$ and function f, $\partial(cf)=c\partial f$. This is because for $\bar{x}\in\mathbb{E}$, if $c>0$,

$$\forall x,\ \langle\phi, x-\bar{x}\rangle\le cf(x)-f(\bar{x}) \iff \forall x,\ \langle\phi/c, x-\bar{x}\rangle\le f(x)-f(\bar{x}).$$

By the condition $\operatorname{dom} g_j\cap\bigcap_{i\in I\setminus\{j\}}\operatorname{cont} g_i$ and inducting with Theorem 3.3.5 (3.14), for any $I'\subset I$ such that $\forall i\in I'$, $\lambda_i>0, \sum_{i\in I'}\lambda_i=1$,

$$\partial\left(\sum_{i\in I'}\lambda_i g_i\right)(\bar{x})=\sum_{i\in I'}\partial(\lambda_i g_i)(\bar{x})=\sum_{i\in I'}\lambda_i\partial g_i(\bar{x}).$$

We obtain

$$\partial\max g_i(\bar{x})=\bigcup\left\{\sum_{i\in I'}\lambda_i\partial g_i(\bar{x})\,\middle|\, I'\subset I, \lambda\in\mathbb{R}^{I'}_{++},\ \sum_{i\in I'}\lambda_i=1\right\}=\operatorname{conv}\bigcup\partial g_i(\bar{x}).$$

18 * (Order convexity). Given a Euclidean space $\mathbb{Y}$ and a closed convex cone $S\subset\mathbb{Y}$, we write $u\le_S v$ for points u and v in $\mathbb{Y}$ if $v-u$ lies in S.

(a) Identify the partial order $\le_S$ in the following cases:

(i) $S=\{0\}$ $\le_S ==$.

(ii) $S=\mathbb{Y}$ $\le_S$ returns true for every pair of elements.

(iii) $\mathbb{Y}=\mathbb{R}^n$ and $S=\mathbb{R}^n_+$. $u\le_S v \iff u\le v$ entrywise.

Given a convex set $C\subset\mathbb{E}$, we say a function $F: C\to\mathbb{Y}$ is S-*convex* if it satisfies

$$F(\lambda x+\mu z)\le_S \lambda F(x)+\mu F(z)$$

for all points x and z in $\mathbb{E}$ and nonnegative reals λ and μ satisfying $\lambda+\mu=1$. If, furthermore, C is a cone and this inequality holds for all λ and μ in $\mathbb{R}_+$, then we say F is S-*sublinear*.

(b) Identify S-convexity in the cases listed in part (a).

(i) $F(\lambda x + (1-\lambda)y) = \lambda F(x) + (1-\lambda)F(y)$ for all $x, y \in C$. For $\mathbb{Y} = \mathbb{R}$, I think this is called being affine. Since if $F(0) = 0$, $F(cx) = cF(x)$ if $c \in [0,1]$ and $F(x) = F(cx/c) = F(cx)/c \implies cF(x) = F(cx)$. Further, $0 = F(0) = F(x/2) + F(-x/2) = (F(x) + F(-x))/2$ implies that $F(-x) = F(x)$. Finally $F(x+y) = F(1/2(2x+2y)) = F(2x)/2 + F(2y)/2 = F(x) + F(y)$ proves that F is linear.

(ii) This is just any ordinary function.

(iii) We have $F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y)$ entrywise, i.e. each component of F is convex.

(c) Prove F is S-convex if and only if the function $\langle \phi, F(\cdot)\rangle$ is convex for all elements ϕ of $-S^-$.

Suppose F is S-convex. Then, given $\lambda \in [0,1]$ and $\phi \in -S^-$,

$$\begin{aligned} &\langle \phi, \lambda F(x) + (1-\lambda)F(y) - F(\lambda x + (1-\lambda)y)\rangle \geq 0 \\ \implies\ &\lambda\langle \phi, F(x)\rangle + (1-\lambda)\langle \phi, F(y)\rangle \geq \langle \phi, F(\lambda x + (1-\lambda)y)\rangle. \end{aligned}$$

If $\langle \phi, F(\cdot)\rangle$ is convex for any $\phi \in -S^-$, then given $x, y \in C$ and $\lambda \in [0,1]$, we have

$$\begin{aligned} &\forall \phi \in -S^-,\ \langle \phi, \lambda F(x) + (1-\lambda)F(y) - F(\lambda x + (1-\lambda)y)\rangle \geq 0 \\ \implies\ &\lambda F(x) + (1-\lambda)F(y) - F(\lambda x + (1-\lambda)y) \in S^{--} = S, \end{aligned}$$

using the fact that S is a closed convex cone so $S^{--} = S$.

(d) Prove the following functions are $\mathbb{S}^n_+$-convex:

(i) $X \in \mathbb{S}^n \mapsto X^2$

(ii) $X \in \mathbb{S}^n_{++} \mapsto X^{-1}$

(iii) $X \in \mathbb{S}^n_+ \mapsto -X^{1/2}$

Hint: Use Exercise 25 in Section 3.1.

By Section 3.1 Exercise 25 (which I have corrected per Sinho's suggestion to look at the coefficient of the quadratic term in the power series), if F is any of these functions, $\langle C, F(\cdot)\rangle \geq 0$ for all $C \in \mathbb{S}^n_+$. Therefore, F is $-(\mathbb{S}^n_+)^- = \mathbb{S}^n_+$ convex, using the fact that $\mathbb{S}^n_+$ is self-dual in the symmetric matrices (Proposition 3.3.12, (3.16)).

(e) Prove the function $X \in \mathbb{S}^2 \mapsto X^4$ is not $\mathbb{S}^2_+$-convex. Hint: Consider the matrices

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}.$$

```
In [1]: import numpy as np
```

```
In [5]: X = np.array([[4,2],[2,1]])
        Y = np.array([[4,0],[0,8]])
```

```
In [6]: def f(X):
            return X@X@X@X
```

```
In [10]: print(f(X), '\n', f(Y))

[[500 250]
 [250 125]]
 [[ 256    0]
 [   0 4096]]
```

```
In [11]: print(f(1/2*X + 1/2*Y))

[[361.25    325.125 ]
 [325.125   523.8125]]
```

```
In [18]: (1/2*f(X) + 1/2*f(Y) - f(1/2*X+1/2*Y))

Out[18]: array([[  16.75   , -200.125 ],
                [-200.125 , 1586.6875]])
```

```
In [19]: np.linalg.det((1/2*f(X) + 1/2*f(Y) - f(1/2*X+1/2*Y)))

Out[19]: -13472.99999999999
```

19 (Order convexity of inversion). For any matrix A in $\mathbb{S}^n_{++}$, define a function $q_A : \mathbb{R}^n \to \mathbb{R}$ by $q_A(x) = x^\top Ax/2$.

(a) Prove $q_A^* = q_{A^{-1}}$.
$q_A^*(y) = \sup_{x\in\mathbb{R}^n} \langle y, x\rangle - x^\top Ax/2$. Setting the gradient to 0, $y = Ax \implies x = A^{-1}y$. This returns $q_A^*(y) = y^\top A^{-1}y/2 = q_{A^{-1}}(y)$.

(b) For any other matrix B in $\mathbb{S}^n_{++}$, prove $2(q_A \odot q_B) \le q_{(A+B)/2}$ (See Exercise 12).

$$2(q_A \odot q_B)(y) = \inf_{x\in\mathbb{E}} x^\top Ax + (y-x)^\top B(y-x) \le \frac{y^\top Ay}{4} + \frac{y^\top By}{4} = q_{(A+B)/2}.$$

(c) Deduce $(A^{-1} + B^{-1})/2 \succeq ((A+B)/2)^{-1}$.
If $f \le g$, then $f^* \le g^*$. Moreover, if $c > 0$, then $(cf)^*(x) = cf^*(x/c)$. Thus,

$$2(q_A \odot q_B)^*(\cdot/2) = (2(q_A \odot q_B))^* \ge q^*_{(A+B)/2} = q_{((A+B)/2)^{-1}}.$$

Using Exercise 12, part (b),

$$2(q_A \odot q_B)^*(\cdot/2) = 2q_{A^{-1}}(\odot/2) + 2q_{B^{-1}}(\odot/2).$$

Combining this with the above, we obtain for all $x \in \mathbb{E}$,

$$x^\top\left(\frac{A^{-1}+B^{-1}}{2}\right)x \ge x^\top\left(\frac{A+B}{2}\right)^{-1}x.$$

20 ** (Pointed cones and bases). Consider a closed convex cone K in $\mathbb{E}$. A *base* for K is a convex set C with $0 \notin \operatorname{cl} C$ and $K = \mathbb{R}_+ C$. Using Exercise 16, prove the following properties are equivalent by showing the implications

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (f) \Rightarrow (a).$$

(a) K is pointed.
$(a) \Rightarrow (b)$: By Exercise 16 part (b), since K and $-K$ are closed convex cones, $(K \cap -K)^- = \operatorname{cl}(K^- - K^-)$ (note $(-K)^- = -K^-$). But $(K \cap -K)^- = (\{0\})^- = \mathbb{E}$, which proves $(a) \Rightarrow (b)$.

(b) $\operatorname{cl}(K^- - K^-) = \mathbb{E}$.
$(b) \Rightarrow (c)$: If C is convex and $\operatorname{cl} C = \mathbb{E}$, then $C = \mathbb{E}$. Applying this to the convex set $K^- - K^-$ gives $(b) \Rightarrow (c)$. The reason why $\operatorname{cl} C = \mathbb{E}$ is, if $\{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathbb{E}$, for any $M > 0$ and $\epsilon > 0$, C contains vectors of the form $\{\pm M e_i + \eta_{i,\pm}\}_{i=1}^n$, where each $\|\eta_{i,\pm}\| \le \epsilon$. We can assume that $\epsilon < M/2\sqrt{n}$. For arbitrary $j \in \{1, \dots, n\}$, consider the set of vectors $\{a_{i,\pm}\}_{i=1}^n = \{\pm M e_i + \eta_{i,\pm} - M e_j/2\}_{i=1}^n$. Take any vector $v \in \mathbb{E}$. Since $\|v\|^2 = \sum_{i=1}^n (\langle v, e_i \rangle)^2$, there exists $i \in [n]$ such that $|\langle v, e_i \rangle| \ge \|v\|/\sqrt{n}$. Now,

$$\begin{aligned} &v^\top (M \operatorname{sgn}(\langle v, e_i \rangle) e_i + \eta_{i,\operatorname{sgn}(\langle v, e_i \rangle)} - M/2 e_j) \\ &\quad \ge \frac{M|\langle v, e_i \rangle|}{2} - \|v\| \|\eta_{i,\operatorname{sgn}(\langle v, e_i \rangle)}\| \ge \frac{M\|v\|}{2\sqrt{n}} - \frac{M\|v\|}{2\sqrt{n}} = 0. \end{aligned}$$

Therefore, there does not exist a solution to the system $\langle v, a_{i,\operatorname{sgn}} \rangle < 0$ for every $i \in [n], \operatorname{sgn} \in \{\pm\}$. Thus, by Gordan's Theorem (Theorem 2.2.1), there exists a solution to the system

$$\sum_{i=1, \operatorname{sgn} \in \{\pm\}}^{n} \lambda_{i,\operatorname{sgn}} a_{i,\operatorname{sgn}} = 0, \quad \sum_{i,\operatorname{sgn}} \lambda_{i,\operatorname{sgn}} = 1, \quad \lambda_{i,\operatorname{sgn}} \ge 0.$$

In other words, $M e_j/2$ is in the convex hull of $\{\operatorname{sgn} M e_i + \eta_{i,\operatorname{sgn}}\}_{i \in [n], \operatorname{sgn} \in \{\pm\}} \subset C$. j and M were arbitrary. Thus, any multiple of any orthogonal basis vector is contained in C. This easily implies that $C = \mathbb{E}$.
Is there a less black magicky way? XD

(c) $K^- - K^- = \mathbb{E}$.
$(c) \Rightarrow (d)$. We'll show $\operatorname{aff} K^- = \mathbb{E}$, which because $\operatorname{relint} K^-$ is nonempty proves that $\operatorname{int} K^- = \operatorname{relint} K^-$ is nonempty.
$k_1^- - k_2^- = k_1^- - k_2^- + 0$, and $0 \in K^-$. This proves $\mathbb{E} = K^- - K^- \subset \operatorname{aff} K^-$, implying $\operatorname{aff} K^- = \mathbb{E}$.

(d) K^- has nonempty interior. (Here you may use the fact that K^- has nonempty relative interior–see Section 1.1, Exercise 13.)
$(d) \Rightarrow (e)$: Let $k^- \in \operatorname{int} K^-$. Then if $y = -k^-$, there exists $\epsilon > 0$ such that for any $d \in \mathbb{E}$ and $x \in K$,

$$\langle y - \epsilon d, x \rangle \ge 0 \implies \langle y, x \rangle \ge \epsilon \langle d, x \rangle.$$

We can choose $d = x/\|x\|$ for nonzero x to obtain the result. If $x = 0$ then the result automatically follows.

(e) There exists a vector y in $\mathbb{E}$ and real $\epsilon > 0$ with $\langle y, x\rangle \geq \epsilon\|x\|$ for all points $x \in K$.
(e) $\Rightarrow$ *(f)*: Let y be such that for some $\epsilon > 0$, $\langle y, x\rangle \geq \epsilon\|x\|$ for all $x \in K$. Consider the set $C = \{x \in K \mid \langle y, x\rangle\} = 1$. This set is convex, because it is the intersection of the convex sets K and $\{x \mid \langle y, x\rangle = 1\}$. It is a base, since $x \in K, x \neq 0$ implies $\langle y, x\rangle > 0$, which implies $\langle y, cx\rangle = 1$ for some $c \in \mathbb{R}_+$. Then $cx/c = x$. It is bounded, since $\langle y, x\rangle = 1 \geq \epsilon\|x\|$ implies $\|x\| \leq 1/\epsilon$.

(f) K has a bounded base.
(f) $\Rightarrow$ *(a)*: Let C be a base for K. Then, $-C$ is a base for $-K$. Thus, $K \cap -K = \mathbb{R}_+C \cap -\mathbb{R}_+C$. Suppose $0 \neq x \in \mathbb{R}_+C \cap -\mathbb{R}_+C$. Then, for some $\mu, \lambda \in \mathbb{R}_+$, we have $x = \mu c_1, c_1 \in C$ and $x = \lambda c_2$, $c_2 \in C$. μ, λ cannot equal 0 because $x \neq 0$. Thus, $x/\mu \in C$ and $-x/\lambda \in C$. Then

$$\frac{\mu}{\mu+\lambda}\frac{x}{\mu} + \frac{\lambda}{\mu+\lambda}\left(-\frac{x}{\lambda}\right) = 0 \in C,$$

contradicting the fact that C is a base. Therefore, $K \cap K^- = \{0\}$.

21 ** (Order-subgradients). This exercise uses the terminology of Exercise 18, and we assume the cone $S \subset \mathbb{Y}$ is pointed: $S \cap -S = \{0\}$. An element y of $\mathbb{Y}$ is the S-infimum of a set $D \subset \mathbb{Y}$ (written $y = \inf_S D$) if the conditions

(i) $D \subset y + S$ and

(ii) $D \subset z + S$ for some z in Y implies $y \in z + S$

both hold.

(a) Verify that this notion corresponds to the usual infimum when $\mathbb{Y} = \mathbb{R}$ and $S = \mathbb{R}_+$.
In this case, $D \subset y + S$ means $D \subset y + \mathbb{R}_+$, i.e. $d \geq y$ for all $d \in D$. The condition $D \subset z + \mathbb{R}_+ \implies y \subset z + \mathbb{R}_+$ means $z \leq d$ for all $d \in D$ implies $z \leq y$. Put together, this means $y = \inf D$.

(b) Prove every subset of $\mathbb{Y}$ has at most one S-infimum.
Suppose y_1 and y_2 are both S-infima of D. Then, by (i), $D \subset y_1 + S$ and $D \subset y_2 + S$, so that by (ii), $y_1 - y_2 \subset S$ and $y_2 - y_1 \subset S$. Since S is pointed, i.e. $S \cap -S = \{0\}$, this implies $y_1 - y_2 = 0$, i.e. $y_1 = y_2$.

(c) Prove *decreasing* sequences in S converge:

$$x_0 \geq_S x_1 \geq_S x_2 \geq_S \ldots \geq_S 0$$

implies $\lim_n x_n$ exists and equals $\inf_S(x_n)$. (Hint: Prove first that $S \cap (x_0 - S)$ is compact using Section 1.1, Exercise 6 (Recession cones).)
We prove that $S = 0^+(S)$. As a reminder, $0^+(S) = \{d \in \mathbb{E} : S + \mathbb{R}_+d \subset S\}$. If $d \in S$, then $S + \mathbb{R}_+d \subset S$ because convex cones are closed under addition of nonnegative

multiples. Thus, $S \subset 0^+(S)$. If $d \in 0^+(S)$, then $\mathbb{R}_+ d \subset S$ (add $\mathbb{R}_+ d$ to $0 \in S$). In particular, $d \in S$. Thus, $0^+(S) \subset S$.

Shifting by a constant does not change the recession cone; therefore, $0^+(x_0 - S) = 0^+(-S) = -S$. Now we have $0^+(S) \cap 0^+(x_0 - S) = S \cap -S = \{0\}$. Therefore, by Section 1.1, Exercise 6 (c), $0^+(S \cap (x_0 - S)) = 0^+(S) \cap 0^+(x_0 - S) = \{0\}$. By Section 1.1, Exercise 6 (d), $S \cap (x_0 - S)$ is bounded (for closed convex C, $0^+(C)$ is nontrivial iff C is unbounded) and closed since it is the intersection of closed sets. Thus, it is compact.

Observe that the sequence lies in $S \cap (x_0 - S)$. By definition, each $x_i \geq_S 0$, so $x_i \in S$. The transitive property holds: $x_0 \geq_S x_i$, $x_i \geq_S x_j$ implies $x_0 \geq_S x_j$, since $x_0 - x_j = x_0 - x_i + x_i - x_j \in S$. Thus, each x_i satisfies $x_0 - x_i \in S \implies x_i \in x_0 - S$. Thus, the sequence contains a convergent subsequence $(x_{j_k})_k$ with limit x. For any x_i in the sequence, for all k large enough, $x_i - x_{j_k} \in S$. By closedness of S, $\lim_{k\to\infty} x_i - x_{j_k} = x_i - x \in S$. We have proven that $x_i \geq_S x$ for all $i \in \mathbb{N}$. Thus, $\{x_n\} \subset x + S$. If $\{x_n\} \subset z + S$, then $x_{j_k} - z \in S$ for all $k \in \mathbb{N}$, implying $x - z \in S$, i.e. $x \in S + z$. Therefore, $x = \inf_S(x_n)$.

Now we show $\lim_n x_n = x$. I am open to alternate solutions for this part XD. By Exercise 20 (e), there exists $a \in \mathbb{E}$, $\varepsilon > 0$ such that for all $x \in S$, $\langle a, x\rangle \geq \varepsilon\|x\|$. Now suppose $x, y \in S$ and $x \neq 0$.

$$\langle a, x + y\rangle = \langle a, x\rangle + \langle a, y\rangle \geq \varepsilon(\|x\| + \|y\|).$$

Now since for any $v \in \mathbb{E}$, $\arg\max_{x:\|x\|=\|v\|}\langle x, v\rangle = v$,

$$\|x+y\|^2 \geq \frac{\|x+y\|}{\|a\|}\langle a, x+y\rangle \geq \frac{\|x+y\|}{\|a\|}\left(\varepsilon\|x\| + \varepsilon\|y\|\right).$$

Now $x + y \neq 0$, since otherwise $x \neq 0 \in S \cap -S$. Thus,

$$\|x+y\| \geq \frac{\varepsilon}{\|a\|}\|x\| \quad \forall x, y \in S,\ x \neq 0. \tag{3.6.3}$$

Returning to showing $\lim_n x_n = x$, given $\epsilon > 0$, $\|x_{j_k} - x\| \leq \frac{\varepsilon\epsilon}{\|a\|}$ for all k large enough. Consider i greater than some j_k with k large enough. We have $x_{j_k} \geq_S x_i \geq_S x$. If $x_i = x$, then clearly $\|x_i - x\| \leq \epsilon$. Otherwise, we can apply (3.6.3) to $x_{j_k} - x_i \in S$ and $x_i - x \in S$:

$$\|x_{j_k} - x_i + x_i - x\| \geq \frac{\varepsilon}{\|a\|}\|x_i - x\|$$
$$\implies \|x_i - x\| \leq \frac{\|a\|}{\varepsilon}\|x_{j_k} - x\| \leq \epsilon.$$

An S-subgradient of F at a point x in C is a linear map $T : \mathbb{E} \to \mathbb{Y}$ satisfying

$$T(z - x) \leq_S F(z) - F(x) \quad \text{for all } z \in C.$$

The set of S-subgradients is denoted $\partial_S F(x)$. Suppose now $x \in \operatorname{core} C$. Generalize the arguments of Section 3.1 in the following steps.

(d) For any direction $h \in \mathbb{E}$, prove

$$\nabla_S F(x;h) = \inf_S \{t^{-1}(F(x+th) - F(x)) \mid t > 0, x+th \in C\}$$

exists and, as a function of h, is S-sublinear.

First we show the function $G : \mathbb{R} \to \mathbb{Y}$, $G(t) = F(x+th) - F(x)$ is S-convex. For $\lambda \in [0,1]$,

$$\begin{aligned} G(\lambda t + (1-\lambda)s) &= F(x + (\lambda t + (1-\lambda)s)h) - F(x) \\ &\in \lambda F(x+th) + (1-\lambda)F(x+sh) - S - F(x) \\ &= \lambda G(t) + (1-\lambda)G(s) - S. \end{aligned}$$

Therefore, $\lambda G(t) + (1-\lambda)G(s) - G(\lambda t + (1-\lambda)s) \in S$. Therefore, G is S-convex. Consider $0 < t < s$. Note $G(0) = 0$. By S-convexity,

$$\frac{t}{s}G(s) = \frac{t}{s}G(s) + \frac{s-t}{s}G(0) \geq_S G(t).$$

Thus, $tG(s)/s - G(t) = s$ for some $s \in S$. Then $G(s)/s - G(t)/t = s/t \in S$, using the fact that S is a cone. Thus, $G(s)/s \geq_S G(t)/t$.

We can also show $G(\cdot)/\cdot$ is increasing for $s < t < 0$. $\frac{t}{s}G(s) \geq_S G(t)$ similarly, so $tG(s)/s - G(t) = s \in S$, so $G(t)/t - G(s)/s = -s/t \in S$ ($-1/t > 0$ because $t < 0$).

Now we show if $s < 0 < t$, then $G(s) \leq_S G(t)$. For $0 < \lambda < 1$ equal to $\lambda = \frac{-s/2}{t-s}$, we have $\lambda s + (1-\lambda)t = \frac{s}{2}$. Thus, using the observable facts $a \leq_S b + c \iff a - c \leq_S b$ and $r > 0 \implies a \leq_S rb \iff a/r \leq_S b$,

$$\begin{gathered} G(s/2) \leq_S \lambda G(t) + (1-\lambda)G(s) \\ \lambda G(s/2) + (1-\lambda)(G(s/2) - G(s)) \leq_S \lambda G(t) \\ G(s/2) \leq_S G(s/2) + \frac{1-\lambda}{\lambda}(G(s/2) - G(s)) \leq_S G(t), \end{gathered}$$

using the fact that $G(s/2) \geq_S G(s)$. By the transitive property of $\leq_S$, $G(s) \leq_S G(t)$.

Since $x \in \operatorname{core} C$, $G(s)$ is defined for some $s < 0$. Then, for any $t > 0$, $G(t) = t^{-1}(F(x+th) - F(x)) \geq_S G(s)$.

Now any decreasing sequence $t_n > 0$, $t_n \to 0$ satisfies $G(t_0) \geq_S G(t_1) \geq_S \ldots \geq_S G(s)$. i.e., $G(t_0) - G(s) \geq_S \ldots \geq_S 0$. By part (c), $\inf_S G(t_n)$ exists and is the limit of the sequence, which we denote as y. Thus, $\lim_n G(t_n) = \lim(G(t_n) - G(s)) + G(s) = y + G(s) =: z$. Finally, we have $z = \inf_S\{G(t) : x + th \in C, t > 0\}$, because (i) for any $t > 0$, $G(t) \geq_S G(t_n) \geq_S z$ for n large enough, and $\{G(t)\} \subset z' + S$ implies $\{G(t_n)\} \subset z' + S \implies z \subset z' + S$. Thus, $z = \nabla_S F(x;h)$.

Since for any sequence $t_n \downarrow 0$, $\lim G(t_n) = \nabla_S F(x;h)$, by analysis, $\lim_{t \downarrow 0} G(t) = \nabla_S F(x;h)$.

(e) For any S-subgradient $T \in \partial_S F(x)$ and direction $h \in \mathbb{E}$, prove $Th \leq_S \nabla_S F(x;h)$.

$$\forall t > 0,\ Th = T(x + th - x)/t \leq_S \frac{F(x+th) - F(x)}{t}.$$

By definition of the S-infimum $\inf_S\{t^{-1}(F(x+th) - F(x)) \mid t > 0\} = \nabla_S F(x;h)$, we have $Th \leq_S \nabla_S F(x;h)$.

(f) Given $h \in \mathbb{E}$, prove there exists $T \in \partial_S F(x)$ satisfying $Th = \nabla_S F(x; h)$. Deduce the max formula

$$\nabla_S F(x; h) = \max\{Th \mid T \in \partial_S F(x)\}$$

and, in particular, that $\partial_S F(x)$ is nonempty. (You should interpret the "max" in the formula.)

Notice that since $S = S^{--}$, i.e. if a point in $\mathbb{Y}$ has nonnegative inner product with all of $-S^-$, then it is in S,

$$T(z-y) \leq_S F(z) - F(y) \iff \forall \phi \in -S^-,\ \langle \phi, T(z-y)\rangle \leq \langle \phi, F(z) - F(y)\rangle.$$

Okay I attempted to solve the problem by using this fact along with the fact that for any $\phi \in -S^-$, $\langle \phi, F(\cdot)\rangle$ is convex, but I failed :(. Instead here is translating each step of the original proof to S-convexity.

Define a function $F : \mathbb{E} \to \mathbb{Y}$ as S-sublinear if for all $x, y \in \mathbb{E}$ and $c, \mu \in \mathbb{R}_+$, $F(cx + \mu y) \leq_S cF(x) + \mu F(y)$. It can be shown that F is S-sublinear iff it is positively homogeneous: $F(cx) = cF(x)$, and S-subadditive: $F(x + y) \leq_S F(x) + F(y)$.

We show that $\nabla_S F(x; \cdot)$ is sublinear. Given $h, k \in \mathbb{E}$,

$$\frac{F(x + t(h+k)/2) - F(x)}{t/2} \leq_S \frac{F(x+th) - F(x)}{t} + \frac{F(x+tk) - F(x)}{t}.$$

Taking $t \to 0$ shows that $\nabla_S F(x; h+k) \leq_S \nabla_S F(x; h) + F(x; k)$. Furthermore,

$$\nabla_S F(x; ch) = \lim_{t \downarrow 0} \frac{F(x + tch) - F(x)}{t} = c \lim_{t\downarrow 0} \frac{F(x+tch) - F(x)}{tc} = c\nabla_S F(x; h).$$

Thus, $\nabla_S F(x; \cdot)$ is S-sublinear.

Now we show that given an S-sublinear G, the function $P(\cdot) = G(h; \cdot)$ satisfies $P(h) = G(h)$ and $\operatorname{lin} P \supset \operatorname{lin} G + \operatorname{Span}\{h\}$. Furthermore, $P \leq G$.

$P(h) = \lim_{t\downarrow 0} \frac{G(h+th)-G(h)}{t} = \lim_t \frac{(1+t)G(h)-G(h)}{t} = G(h)$. $P(-h) = \lim_{t\downarrow 0} \frac{G(h-th)-G(h)}{t} = \lim_t \frac{(1-t)G(h)-G(h)}{t} = -G(h)$. Coupled with the fact that P is positively homogenous, this proves $\operatorname{Span}\{h\} \subset \operatorname{lin} P$. $P(x) \leq_S \lim_{t\downarrow 0} \frac{G(h+tx)-G(h)}{t} \leq_S G(x)$, so $P \leq_S G$. Now if $y \in \operatorname{lin} G$, then $P(-y) \leq_S G(-y) = -G(y)$. Further, $P(-y) \geq_S -P(y) \geq_S -G(y)$, so that $P(-y) = -G(y) = G(-y) = -P(y)$. Thus, $\operatorname{lin} G \subset \operatorname{lin} P$. Since $\operatorname{lin} P$ is linear ($P(x+y) \leq_S P(x) + P(y)$, $-P(x+y) \leq_S P(-x-y) \leq_S -P(x) - P(y)$ implies $P(x+y) = P(x) + P(y)$), $\operatorname{lin} G + \operatorname{Span}\{h\} \subset \operatorname{lin} P$.

The proof of the max formula closely follows the case for normal convexity. Define $p_0(\cdot) = \nabla_S F(x; \cdot)$ and the basis $\{v_1 = h, v_2, \ldots, v_n\}$ of $\mathbb{E}$. If $h = 0$, then any subgradient T satisfies $Th = \nabla_s F(x; h) = 0$. We'll show a subgradient exists by consider $h \neq 0$. Define $p_i(\cdot) = p_{i-1}(v_{i-1}; \cdot)$ for every $i = 1, \ldots, n$. Note $p_0, \ldots, p_n$ are all sublinear. We have $\operatorname{lin} p_n = \mathbb{E}$, so that $p_n(\cdot) = T\cdot$ for some $T : \mathbb{E} \to \mathbb{Y}$. For any $z \in C$,

$$T(z-x) \leq_S p_{n-1}(z-x) \leq_S \ldots \leq_S p_0(z-x) = \nabla_S F(x; z-x) \leq_S F(z) - F(x).$$

Thus, T is an S-subgradient. Finally, we have

$$-Th = T(-h) \leq_S p_{n-1}(-h) \leq_S \ldots \leq_S p_1(-h) = -p_1(h) = -p_0(h) = -\nabla_S F(x; h).$$

Thus, $Th = \nabla_S F(x; h)$. This proves

$$\nabla_S F(x; h) = \max_{T \in \partial_S F(x)} Th,$$

where the max means that $Th \geq_S T'h$ for all $T' \in \partial_S F(x)$.

(g) The function F is *Gateaux differentiable* at x (with derivative the linear map $\nabla F(x) : \mathbb{E} \to \mathbb{Y}$) if

$$\lim_{t \to 0} t^{-1}(F(x + th) - F(x)) = (\nabla F(x))h$$

holds for all $h \in \mathbb{E}$. Prove this is the case if and only if $\partial_S F(x)$ is a singleton.
If $\partial_S F(x) = \{T\}$, then the max formula shows that $\nabla_S F(x; h) = Th$ for all $h \in \mathbb{E}$. Thus, $Th = \nabla_S F(x; h)$.
Conversely, if T is the gradient, then if A is an S-subgradient,

$$\forall h \in \mathbb{E},\ Ah \leq_S \nabla_S F(x; h) = Th \implies \forall h \in \mathbb{E},\ (T - A)h \in S.$$

$(T - A)\mathbb{E}$ is a linear subspace that is a subset of S. Therefore, it equals 0. Thus, $T = A$, i.e. the only subgradient is T.

Now fix an element ϕ of $-\operatorname{int}(S^-)$.

(h) Prove $\langle \phi, F(\cdot) \rangle'(x; h) = \langle \phi, \nabla_S F(x; h) \rangle$.

$$\lim_{t \downarrow 0} \frac{\langle \phi, F(x + th) - F(x) \rangle}{t} = \langle \phi, \lim_{t \downarrow 0} \frac{F(x + th) - F(x)}{t} \rangle = \langle \phi, \nabla_S F(x; h) \rangle.$$

(i) Prove F is Gateaux differentiable at x if and only if $\langle \phi, F(\cdot) \rangle$ is likewise.
If F is Gateaux differentiable at x, then

$$\forall h \in \mathbb{E},\ \langle \phi, F(\cdot) \rangle'(x; h) = \langle \phi, \nabla_S F(x; h) \rangle = \langle \phi, \nabla F(x)h \rangle = \langle \nabla F(x)^* \phi, h \rangle.$$

Therefore, $(\nabla F(x))^* \phi = \nabla \langle \phi, F(\cdot) \rangle (x)$. On the other hand, if $\langle \phi, F(\cdot) \rangle$ is differentiable at x, then there exists a $\psi \in \mathbb{E}$ such that $\langle \phi, F(\cdot) \rangle'(x; h) = \langle \psi, h \rangle$. Moreover, $\partial \langle \phi, F(\cdot) \rangle (x) = \{\psi\}$ (See Exercise 18 (c) for the fact that $\langle \phi, F(\cdot) \rangle$ is convex whenever $\phi \in -S^-$). Now $\phi \in -\operatorname{int}(S^-)$ implies there exists $\epsilon > 0$ such that for all $d \in \mathbb{E}$, $\|d\| \leq 1$, $\phi + \epsilon d \in -S^-$. Consider $\phi + \epsilon d$ and $\phi - \epsilon d$. By convexity of $\langle \phi \pm \epsilon d, F(\cdot) \rangle$, there exist subgradients ψ_1, ψ_2 such that

$$\langle \psi_1, z - x \rangle \leq \langle \phi + \epsilon d, F(z) - F(x) \rangle$$
$$\langle \psi_2, z - x \rangle \leq \langle \phi - \epsilon d, F(z) - f(x) \rangle.$$

Notice that by adding these, $\frac{\psi_1 + \psi_2}{2}$ is a subgradient for $\langle \phi, F(\cdot) \rangle$. But the subgradient of $\langle \phi, F(\cdot) \rangle$ is a singleton. Therefore, the subgradients of $\langle \phi \pm \epsilon d, F(\cdot) \rangle$ are also singletons, i.e. these functions are differentiable. WLOG, an element of $\phi + \epsilon B$ can be expressed as $\phi + \epsilon d$ for some $\|d\| \leq 1$. Thus, all the functions defined by the ball are differentiable.

Choose a basis of $\mathbb{Y}$ from the ball, $\{\phi_1, \ldots, \phi_m\}$. We have $\{\psi_1, \ldots, \psi_m\}$ as the gradients $\{\nabla\langle\phi_1, F(\cdot)\rangle(x), \ldots, \nabla\langle\phi_m, F(\cdot)\rangle(x)\}$. For any $\phi = \sum_{i=1}^n c_i\phi_i \in \mathbb{Y}$,

$$\nabla\langle\phi, F(\cdot)\rangle(x) = \nabla\sum_{i=1}^m c_i\langle\phi_i, F(\cdot)\rangle(x) = \sum_{i=1}^m c_i\nabla\langle\phi_i, F(\cdot)\rangle(x) = \sum_{i=1}^m c_i\psi_i.$$

So, if we define the linear map $A : \mathbb{Y} \to \mathbb{E}$ by mapping the basis $\{\phi_1, \ldots, \phi_m\}$ to $\{\psi_1, \ldots, \psi_m\}$,

$$\langle\phi, A^*(z-x)\rangle = \langle A\phi, z-x\rangle = \langle\sum_i c_i\psi_i, z-x\rangle = \langle\phi, F(\cdot)\rangle'(x; z-x) = \langle\phi, \nabla_S F(x; z-x)\rangle.$$

Since ϕ was arbitrary, $A^*(z-x) = \nabla_S F(x; z-x)$ for all $z \in C$. Thus, $A^* = \nabla F(x)$.

22 ** (Linearly constrained examples). Prove Corollary 3.3.11 (Fenchel duality for linear constraints). Deduce duality theorems for the following problems.

Proof.

$$\begin{aligned}\inf_{x\in\mathbb{E}}\{f(x) \mid Ax = b\} &= \inf_{x\in\mathbb{E}}\{f(x) + \delta_{\{b\}}(Ax)\} \\ &\geq \sup_{\phi\in\mathbb{Y}}\{-f^*(A^*\phi) - \delta^*_{\{b\}}(-\phi)\} = \sup_{\phi\in\mathbb{Y}}\{\langle\phi, b\rangle - f^*(A^*\phi)\}.\end{aligned}$$

By Fenchel duality (Theorem 3.3.5 (3.14)), we have the above, and equality if $0 \in \operatorname{core}(\{b\} - A\operatorname{dom} f)$, i.e. $b \in \operatorname{core}(A\operatorname{dom} f)$, where the sup is attained if finite. □

(a) **Separable problems**

$$\inf\left\{\sum_{i=1}^n p(x_i) \,\middle|\, Ax = b\right\},$$

where the map $A : \mathbb{R}^n \to \mathbb{R}^m$ is linear, $b \in \mathbb{R}^m$, and the function $p : \mathbb{R} \to (-\infty, +\infty]$ is convex, defined as follows:

(i) **(Nearest points in polyhedrons).** $p(t) = t^2/2$ with domain $\mathbb{R}_+$.

$$\inf_{x\in\mathbb{R}^n_+}\left\{\frac{\|x\|^2}{2} \,\middle|\, Ax = b\right\} \geq \sup_{y\in\mathbb{R}^m}\left\{\langle b, y\rangle - \frac{\|(A^\top y)^+\|^2}{2}\right\}$$

We used Exercise 1, 6.: $(\frac{|x|^2}{2}, \mathbb{R}_+) \leftrightarrow (\frac{|y^+|^2}{2}, \mathbb{R})$. Equality occurs if $b \in \operatorname{core}(A\mathbb{R}^n_+)$. Also,

$$\left(\sum_i p(x_i)\right)^*(\phi) = \sup_{x\in\mathbb{R}^n}\sum_i(\phi_i x_i - p(x_i)) = \sum_i p^*(\phi_i).$$

(ii) **(Analytic center).** $p(t) = -\log t$ with domain $\mathbb{R}_{++}$.

$$\inf_{x\in\mathbb{R}^n_{++}} \left\{ \sum_{i=1}^n -\log x_i \,\middle|\, Ax = b \right\} \geq \sup_{y\in -(A^\top)^{-1}\mathbb{R}^n_{++}} \left\{ \langle b, y\rangle + 1 + \sum_{i=1}^n \log(-(A^\top y)_i) \right\}.$$

Equality occurs if $b \in \operatorname{core}(A\mathbb{R}_{++})$. We used $(-\log)^*(s) = \sup_{t>0} st + \log t,\ t = -1/s \implies (-\log)^*(s) = -1 - \log(-s)$ with domain $-\mathbb{R}_{++}$.

(iii) **(Maximum entropy).** $p = \exp^*$.

$$\inf_{x\in\mathbb{R}^n} \left\{ \sum_{i=1}^n \exp^*(x_i) \,\middle|\, Ax = b \right\} \geq \sup_{y\in\mathbb{R}^m} \left\{ \langle b, y\rangle - \sum_{i=1}^n \exp((A^\top y)_i) \right\}.$$

Equality is obtained if $b \in \operatorname{core}(A\mathbb{R}^n_+)$.

What happens if the objective function is replaced by $\sum_i p_i(x_i)$?

(b) The **BFGS update** problem in Section 2.1, Exercise 13.

$$\inf_{X\in\mathbb{S}^n_{++}} \left\{ \langle C, X\rangle - \log\det X \,\middle|\, Xs = y \right\} \geq \sup_{z\in(s^*)^{-1}(C-\mathbb{S}^n_{++})} \left\{ \langle y, z\rangle + n + \log\det(C - \frac{1}{2}(zs^\top + sz^\top)) \right\}.$$

Equality occurs if $y \in \operatorname{core}(\mathbb{S}^n_{++}s)$. To compute $\sup_{X\in\mathbb{S}_{++}} \langle D - C, X\rangle + \log\det X$, set $D - C + X^{-1} = 0$, i.e. $X = (C-D)^{-1}$. We get the function $-n - \log\det(C - D)$, with domain $C - \mathbb{S}^n_{++}$. Further, recall that $\langle z, Xs\rangle = \langle \frac{1}{2}(zs^\top + sz^\top), X\rangle$, so that $(\cdot s)^*(z) = \frac{1}{2}(zs^\top + sz^\top)$.

(c) The **DAD problem** in Section 3.1, Exercise 28.
Define $G \in \mathbb{R}^{2k\times Z}$ (Z is the set of (i,j) where A is nonzero) by setting $G_{i,(i,j)} = G_{j,(i,j)} = 1$ for all $(i,j) \in Z$ and the rest of the entries to 0. Then, the DAD problem can be expressed as

$$\inf_{x\in\mathbb{R}^Z} \left\{ \sum_{(i,j)\in Z} (\exp^*(x_{ij}) - x_{ij}\log A_{ij}) \,\middle|\, Gx = \mathbf{1} \right\} \geq \sup_{y\in\mathbb{R}^{2k}} \left\{ \langle y, \mathbf{1}\rangle - \sum_{(i,j)\in Z} A_{ij}\exp(y_i + y_j) \right\}.$$

Equality occurs if $\mathbf{1} \in G\mathbb{R}^Z_+$. $\sup_x yx - \exp^*(x) + x_{ij}\log A_{ij}$ is obtained from $y - \log x + \log A_{ij} = 0 \implies x = \exp(y + \log A_{ij})$.

(d) Example (3.3.1).
Denote $A = [a^0\, a^1\, \ldots\, a^m]$.

$$\inf_{x\in\mathbb{R}^{m+1}} \left\{ \sum_{i=0}^m \exp^*(x_i) \,\middle|\, \sum_i x_i = 1,\ Ax = z \right\} \geq \sup_{y\in\mathbb{R}^{n+1}} \left\{ [1\ z^\top]y - \sum_{i=0}^m \exp([\mathbf{1}\ A^\top]y)_i \right\}.$$

Equality occurs if $\begin{bmatrix} 1 \\ z \end{bmatrix} \in \operatorname{core}(\begin{bmatrix} \mathbf{1}^\top \\ A \end{bmatrix} \mathbb{R}^{m+1}_+)$.

23 * (Linear inequalities). What does Corollary 3.3.11 (Fenchel duality for linear constraints) (3.15) become if we replace the constraint $Ax = b$ by $Ax \in b + K$ where $K \subset \mathbb{Y}$ is a convex cone? Write down the dual problem for Section 3.2, Exercise 2, part (a), solve it, and verify the duality theorem.

Proof. First we compute δ^*_{b+K}.

$$\delta^*_{b+K}(y) = \sup_{x \in b+K} \langle x, y \rangle = \langle b, y \rangle + \delta_{K^-}(y).$$

If $y \notin K^-$, then exists $k \in K$ where $\langle k, y \rangle > 0$, and so $\langle b + ck, y \rangle \to +\infty$ as $c \to +\infty$, giving $\delta^*_{b+K}(y) = +\infty$. Otherwise, $\langle b + k, y \rangle \leq \langle b, y \rangle$ with equality at $k = 0 \in K$. Therefore, the duality theorem becomes

$$\inf_{x \in \mathbb{E}} \{f(x) \mid Ax \in b + K\} \geq \sup_{\phi \in \mathbb{Y}} \{-\delta^*_{b+K}(-\phi) - f^*(A^*\phi)\} = \sup_{\phi \in \mathbb{Y}} \{\langle b, \phi \rangle - f^*(A^*\phi) \mid \phi \in -K^-\}$$

with equality if $0 \in \operatorname{core}(b + K - A \operatorname{dom} f)$, i.e. $b \in \operatorname{core}(A \operatorname{dom} f - K)$.
Section 3.2 Exercise 2 part (a) features the problem

$$\begin{array}{ll} \inf & x_1^2 + x_2^2 - 6x_1 - 2x_2 + 10 \\ \text{subject to} & 2x_1 + x_2 - 2 \leq 0 \\ & x_2 - 1 \leq 0 \\ & x \in \mathbb{R}^2. \end{array}$$

The optimal solution found in that exercise was $\bar{x} = (1, 0)$ with value 5. Let us compute the dual, first by computing the conjugate of the objective. As a shortcut, we subtract the gradient of the objective from an arbitrary $y \in \mathbb{R}^2$:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} 2x_1 - 6 \\ 2x_2 - 2 \end{bmatrix} = 0 \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{y_1}{2} + 3 \\ \frac{y_2}{2} + 1 \end{bmatrix}.$$

Thus the conjugate of the objective is

$$\begin{aligned} g(y) &= \langle y, x \rangle - x_1^2 - x_2^2 + 6x_1 + 2x_2 - 10 \\ &= \frac{y_1^2 + y_2^2}{2} + 3y_1 + y_2 - (\frac{y_1}{2} + 3)^2 - (\frac{y_2}{2} + 1)^2 + 3y_1 + 18 + y_2 + 2 - 10 \\ &= \frac{y_1^2 + y_2^2}{4} + 3y_1 + y_2 + 10 - 10 = \frac{y_1^2 + y_2^2}{4} + 3y_1 + y_2. \end{aligned}$$

The problem constraints may be formulated as

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} x = Ax \leq \begin{bmatrix} 2 \\ 1 \end{bmatrix} = b, \qquad \text{i.e.} \qquad Ax \in b - \mathbb{R}^2_+.$$

We can now write the dual:

$$\begin{aligned} &\sup_{\phi \in \mathbb{R}^2} \{\langle b, \phi \rangle - g(A^\top \phi) : \phi \in -(-\mathbb{R}^2_+)^- = -\mathbb{R}^2_+\} \\ &\quad = \sup_{\phi \in -\mathbb{R}^2_+} \left\{ 2\phi_1 + \phi_2 - \phi_1^2 - \frac{(\phi_1 + \phi_2)^2}{4} - 6\phi_1 - \phi_1 - \phi_2 \right\} \\ &\quad = \sup_{\phi \in -\mathbb{R}^2_+} \left\{ -\phi_1^2 - \frac{(\phi_1 + \phi_2)^2}{4} - 5\phi_1 \right\} = - \inf_{\phi \in -\mathbb{R}^2_+} \left\{ \phi_1^2 + \frac{(\phi_1 + \phi_2)^2}{4} + 5\phi_1 \right\}. \end{aligned}$$

We can look for a Lagrange multiplier vector (the constraints are $\phi_1 \le 0,\ \phi_2 \le 0$). The gradient of the Lagrangian is

$$\begin{bmatrix} \frac{5\phi_1}{2} + \frac{\phi_2}{2} + 5 + \lambda_1 \\ \frac{\phi_1+\phi_2}{2} + \lambda_2 \end{bmatrix} = 0. \tag{3.6.4}$$

If $\lambda_2 = 0$, then $\phi_1 + \phi_2 = 0$, which by feasibility implies $\phi_1 = \phi_2 = 0$, implying $\lambda_1 = -5$, a contradiction. Therefore, $\phi_2 = 0$. Set $\lambda_1 = 0$. (3.6.4) becomes

$$\begin{bmatrix} \frac{5\phi_1}{2} + 5 \\ \frac{\phi_1}{2} + \lambda_2 \end{bmatrix} = 0 \implies \phi_1 = -2,\ \lambda_2 = 1.$$

The feasible solution $\phi = (-2, 0)$ has the Lagrange multiplier vector $\lambda = (0, 1)$. The optimal value is

$$-\left((-2)^2 + \frac{(-2)^2}{4} - 10\right) = 5,$$

which matches the optimal value of the original problem. □

24 (Symmetric Fenchel duality). For functions $f, g : \mathbb{E} \to [-\infty, +\infty]$, define the *concave conjugate* $g_* : \mathbb{E} \to [-\infty, +\infty]$ by

$$g_*(\phi) = \inf_{x \in \mathbb{E}} \{\langle \phi, x\rangle - g(x)\}.$$

Prove

$$\inf(f - g) \ge \sup(g_* - f^*),$$

with equality if f is convex, g is concave, and

$$0 \in \operatorname{core}(\operatorname{dom} f - \operatorname{dom}(-g)).$$

Proof. By Theorem 3.3.5 (Fenchel Duality, (3.14)),

$$\inf_{x \in \mathbb{E}} \{f(x) - g(x)\} \ge \sup_{\phi \in \mathbb{E}} \{-f^*(\phi) - (-g)^*(-\phi)\}$$

with equality if f is convex, g is concave, and $0 \in \operatorname{core}(\operatorname{dom}(-g) - \operatorname{dom} f)$. We have

$$-(-g)^*(-\phi) = -\sup_{x \in \mathbb{E}} \{-\langle \phi, x\rangle + g(x)\} = \inf_{x \in \mathbb{E}} \{\langle \phi, x\rangle - g(x)\} = g_*(\phi).$$

Therefore, we have

$$\inf_{x \in \mathbb{E}} \{f(x) - g(x)\} \ge \sup_{\phi \in \mathbb{E}} \{g_*(\phi) - f^*(\phi)\}$$

with equality if f is convex, g is concave, and $0 \in \operatorname{core}(\operatorname{dom}(-g) - \operatorname{dom} f)$. □

25 ** (Divergence bounds [135]).

(a) Prove the function

$$t \in \mathbb{R} \mapsto 2(2+t)(\exp^* t + 1) - 3(t-1)^2$$

is convex and minimized when $t = 1$.
For $t > 0$, the derivative is

$$2(\exp^* t + 1) + 2(2+t)\log t - 6(t-1).$$

The second derivative is

$$2\log t + 2\log t + \frac{2(2+t)}{t} - 6 = 4\log t + \frac{4}{t} - 4.$$

The only critical point of $4\log t + \frac{4}{t} - 4$ is, taking the derivative again which equals $\frac{4}{t} - \frac{4}{t^2} = (4/t)(1-1/t)$, $t = 1$. At this point, the function equals 0. Clearly, the function goes to $+\infty$ as $t \to +\infty$, and in fact also as $t \to 0$ since the derivative goes to $-\infty$. If it dropped below 0, it would equal 0 again, leading to another critical point. Therefore, $4\log t + \frac{4}{t} - 4 \geq 0$ for all $t > 0$, proving that $2(2+t)(\exp^* t + 1) - 3(t-1)^2$ is convex (use continuity to conclude for $t = 0$). Also, the derivative equals 0 at $t = 1$.

(b) For $v \in \mathbb{R}_{++}$ and $u \in \mathbb{R}_+$, deduce the inequality

$$3(u-v)^2 \leq 2(u+2v)\left(u\log\left(\frac{u}{v}\right) - u + v\right).$$

The minimum of the function at $t = 1$ equals 0; thus the function is nonnegative. Plugging in $t = u/v$, we obtain:

$$2(2+\frac{u}{v})\left(\frac{u}{v}\log\left(\frac{u}{v}\right) - \frac{u}{v} + 1\right) - 3\left(\frac{u}{v} - 1\right)^2 \geq 0.$$

Moving the squared term to the other side and multiplying both sides by v^2 gives the desired inequality.
Now suppose the vector $p \in \mathbb{R}^n_{++}$ satisfies $\sum_{i=1}^n p_i = 1$.

(c) If the vector $q \in \mathbb{R}^n_{++}$ satisfies $\sum_{i=1}^n q_i = 1$, use the Cauchy-Schwarz inequality to prove the inequality

$$\left(\sum_{i=1}^n |p_i - q_i|\right)^2 \leq 3\sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + 2q_i},$$

and deduce the inequality

$$\sum_{i=1}^n p_i \log\left(\frac{p_i}{q_i}\right) \geq \frac{1}{2}\left(\sum_{i=1}^n |p_i - q_i|\right)^2.$$

Define $u \in \mathbb{R}^n_+$ and $v \in \mathbb{R}^n_{++}$ by $u_i = \frac{|p_i - q_i|}{\sqrt{p_i + 2q_i}}$ and $v_i = \sqrt{p_i + 2q_i}$. We obtain

$$\left(\sum_{i=1}^n |p_i - q_i|\right)^2 = (u^\top v)^2 \leq \|v\|^2 \|u\|^2 = \left(\sum_{i=1}^n p_i + 2q_i\right)\left(\sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + 2q_i}\right).$$

Then we note that $\sum_{i=1}^n p_i + 2\sum_{i=1}^n q_i = 3$ since p and q both sum to 1. Now using part (b) with $u = p_i, v = q_i$,

$$\begin{aligned}\left(\sum_{i=1}^n |p_i - q_i|^2\right) &\leq 3\sum_{i=1}^n \frac{(p_i - q_i)^2}{p_i + 2q_i} \leq 2\sum_{i=1}^n \frac{p_i + 2q_i}{p_i + 2q_i}\left(p_i \log\left(\frac{p_i}{q_i}\right) - p_i + q_i\right) \\ &= 2\sum_{i=1}^n \log\left(\frac{p_i}{q_i}\right) - \sum_{i=1}^n p_i - q_i = 2\sum_{i=1}^n \log\left(\frac{p_i}{q_i}\right),\end{aligned}$$

using the fact that $\sum_i p_i = \sum_i q_i = 1$. Divide by 2 on both sides to get the desired result. I believe this is Pinsker's inequality, wow, what a completely obtuse way of deriving it!

(d) Hence show the inequality

$$\log n + \sum_{i=1}^n p_i \log p_i \geq \frac{1}{2}\left(\sum_{i=1}^n \left|p_i - \frac{1}{n}\right|\right)^2.$$

$$\frac{1}{2}\left(\sum_{i=1}^n \left|p_i - \frac{1}{n}\right|\right)^2 \leq \sum_{i=1}^n p_i \log\left(\frac{p_i}{1/n}\right) = \sum_{i=1}^n p_i(\log p_i + \log n) = \sum_{i=1}^n p_i \log p_i + \log n.$$

(e) Use convexity to prove the inequality

$$\sum_{i=1}^n p_i \log p_i \leq \log \sum_{i=1}^n p_i^2.$$

This is immediate from the concavity of log.

(f) Deduce the bound

$$\log n + \sum_{i=1}^n p_i \log p_i \leq \frac{\max p_i}{\min p_i} - 1.$$

$$\begin{aligned}\log n + \sum_{i=1}^n p_i \log p_i &\leq \log n + \log \sum_{i=1}^n p_i^2 = \log \sum_{i=1}^n p_i \frac{p_i}{1/n} \\ &\leq \log \sum_{i=1}^n p_i \frac{\max_i p_i}{\min_i p_i} = \log \frac{\max_i p_i}{\min_i p_i} \leq \frac{\max_i p_i}{\min_i p_i} - 1.\end{aligned}$$

I found deducing this pretty tricky actually.