Example class 1

1. Prove the separating hyperplane theorem (Theorem 1.1) when C is not necessarily closed.

   Solution: Let C be a convex set (not necessarily closed) and let $y \notin C$. If $y \notin \mathrm{cl}(C)$ (where $\mathrm{cl}(C)$ is the closure of C) then we can use the proof done in class to separate y from the closure of C (which is convex). Assume then that $y \in \mathrm{cl}(C) \setminus C$. We can find a sequence $(y_k) \notin \mathrm{cl}(C)$ such that $y_k \to y$. By applying the separating hyperplane theorem to $y_k \notin \mathrm{cl}(C)$ we know that there exist $a_k \in \mathbb{R}^n \setminus \{0\}$ and $b_k \in \mathbb{R}$ such that $\langle a_k, y_k \rangle = b_k$ and $\langle a_k, x \rangle \leq b_k$ for all $x \in \mathrm{cl}(C)$. We can also (by adjusting b_k) assume that $\|a_k\| = 1$ for all k. Thus we can extract from (a_k) a subsequence that converges to $a \in \mathbb{R}^n \setminus \{0\}$. Let $b = \lim b_k = \langle a, y \rangle$. Then we get that $\langle a, y \rangle = b$ and $\langle a, x \rangle \leq b$ for all $x \in C$ which is what we wanted.

2. Let C, D be two disjoint convex sets in $\mathbb{R}^n$. Show that there is $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$ such that $\langle a, x \rangle \leq b$ for all $x \in C$ and $\langle a, y \rangle \geq b$ for all $y \in D$. If we assume C and D closed, can we always find a *strict* separating hyperplane?

   Solution: If C and D are convex then the set $C - D = \{x - y : x \in C, y \in D\}$ is also convex. Also if $C \cap D = \emptyset$ then $0 \notin C - D$. Applying the separating hyperplane theorem gives the desired result. One can get a strict separating hyperplane if we assume for example that C and D are closed and at least one is bounded. Merely assuming that C and D are closed is not enough to guarantee a strict separating hyperplane (e.g., take $C = \{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } y \geq 1/x\}$ and $D = \{(x, y) \in \mathbb{R}^2 : y \leq 0\}$).

3. (Carathéodory theorem) Let $S \subset \mathbb{R}^n$. Show that any point in $\mathrm{conv}(S)$ can be written as a convex combination of at most $n+1$ points of S (hint: any k points $s_1, \ldots, s_k$ with $k \geq n+2$ are affinely dependent i.e., there exist $\mu_1, \ldots, \mu_k$ such that $\sum_{i=1}^k \mu_i s_i = 0$ and $\sum_{i=1}^k \mu_i = 0$).

   Solution: Assume $x \in \mathrm{conv}(S)$ is a convex combination of k elements in S, i.e., $x = \sum_{i=1}^k \lambda_i s_i$ where $s_i \in S, \lambda_i > 0$ and $\sum_{i=1}^k \lambda_i = 1$. If $k \geq n+2$ we will show that one can express x using $k-1$ elements of S. As indicated in the hint, if $k \geq n+2$ there exist $\mu_1, \ldots, \mu_k$ such that $\sum_{i=1}^k \mu_i s_i = 0$ and $\sum_{i=1}^k \mu_i = 0$. Observe that for any $t \in \mathbb{R}$ we have $x = \sum_{i=1}^k (\lambda_i + t\mu_i) s_i$ and that $\sum_{i=1}^k (\lambda_i + t\mu_i) = 1$. One can choose a value of t such that one of the coefficients $\lambda_i + t\mu_i$ is 0, while the rest remains ≥ 0 (take for example $t = \max\{-\lambda_i/\mu_i : i \text{ s.t. } \mu_i > 0\}$). For this value of t the expression $x = \sum_{i=1}^k (\lambda_i + t\mu_i) s_i$ gives x as a convex combination of $k-1$ elements of S, which is what we wanted.

4. Let C be a closed and bounded convex set in $\mathbb{R}^n$ and consider the minimization problem $\min\{\langle c, x \rangle : x \in C\}$ where $c \in \mathbb{R}^n$. Show that there is at least one optimal solution that is an extreme point of C.

   Solution: This is a direct consequence of Minkowski's theorem. Since C is compact the minimization problem $\min\{\langle c, x \rangle : x \in C\}$ is attained. Let x^* be an optimal point. By Minkowski theorem we can write x^* as a convex combination of extreme points $x^* = \sum_i \lambda_i x_i$ where each x_i is an extreme point of C. But then $\langle c, x_i \rangle = \langle c, x^* \rangle$ for all i since $\langle c, x^* \rangle = \sum_i \lambda_i \langle c, x_i \rangle \geq \sum_i \lambda_i \langle c, x^* \rangle = \langle c, x^* \rangle$. This implies that all the x_i are also optimal solutions of the problem $\min\{\langle c, x \rangle : x \in C\}$ and since they are extreme points of C this completes the proof.

5. Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ and the inequality $Ax \leq b$ is interpreted componentwise. Show that if x is an extreme point of P then there

exists $I \subseteq \{1, \dots, m\}$ such that $\ker(A_I) = \{0\}$ (where A_I is the matrix obtained from A by keeping only the rows in I), $b \in \operatorname{im}(A_I)$ and $x = A_I^{-1} b_I$. Deduce that P has a finite number of extreme points and give an upper bound for the number of extreme points in terms of m and n. [*Note:* an extreme point of a polyhedron is sometimes called a *basic feasible point*, especially in the context of the simplex algorithm for linear programming. The set I that identifies the extreme point is known as the *basis* of the basic feasible point].

Solution: We start by proving the claim.

Claim 1. *If x is an extreme point of P then there must exist $I \subseteq \{1, \dots, m\}$ such that $\ker(A_I) = \{0\}$ (where A_I is the matrix obtained from A by keeping only the rows in A), $b \in \operatorname{im}(A_I)$ and $x = A_I^{-1} b_I$.*

Proof. Let x be an extreme point of P and let $I = \{i \in \{1, \dots, m\} : \langle a_i, x \rangle = b_i\}$ (where a_i is the i'th row of A; the set I is called the *active set* at x; it is the set of linear inequalities that are *active* at x). We will show that $\ker(A_I) = \{0\}$. Let $v \in \ker(A_I)$. Since $\langle a_i, x \rangle < b_i$ for all $i \in I^c$ (the complement of I), we can find $\epsilon > 0$ small enough such that $A(x + \epsilon v) \leq b$ and $A(x - \epsilon v) \leq b$. But then $x \pm \epsilon v \in P$ and, unless $v = 0$, this contradicts that x is an extreme point (since then $x = \frac{1}{2}(x + \epsilon v) + \frac{1}{2}(x - \epsilon v)$). So we have shown that $\ker(A_I) = \{0\}$. That $b_I \in \operatorname{im}(A_I)$ is clear since $A_I x = b_I$ by definition of I. It thus follows that $x = A_I^{-1} b_I$. □

The number of extreme points of P is thus at most the number of subsets $I \subseteq \{1, \dots, m\}$ that satisfy $\ker(A_I) = \{0\}$ and $b_I \in \operatorname{im}(A_I)$. There are at most $\binom{m}{n}$ such subsets and so this gives us an upper bound on the number of extreme points of P.

6. Let $P = \operatorname{conv}(v_1, \dots, v_N)$ where $v_1, \dots, v_N \in \mathbb{R}^n$. Show that P is the intersection of a finite number of halfspaces. Give an upper bound on the number of halfspaces needed. (*Hint:* Assume (without loss of generality) that P has nonempty interior and $0 \in \operatorname{int}(P)$, and apply the result of the previous question to the polyhedron $P^o = \{y \in \mathbb{R}^n : \langle y, v_i \rangle \leq 1, \ \forall i = 1, \dots, N\}$; P^o is known as the *polar* of P).

   Solution: We can assume without loss of generality that P has nonempty interior and that $0 \in \operatorname{int}(P)$. Define $P^o = \{y \in \mathbb{R}^n : \langle y, v_i \rangle \leq 1 \ \forall i = 1, \dots, N\}$. Note that P^o is a polyhedron. Thus by the previous question it has a finite number of extreme points $y_1, \dots, y_m$. We claim that $P = \{x \in \mathbb{R}^n : \langle y_i, x \rangle \leq 1, \ \forall i = 1, \dots, m\}$, i.e., an intersection of m halfspaces. The inclusion $\subseteq$ is trivial by the definition of P^o. To prove $\supseteq$ assume that $x \notin P$. We will show that there is at least one $i \in \{1, \dots, m\}$ such that $\langle y_i, x \rangle > 1$. Since $x \notin P$, there exists $y \in \mathbb{R}^n \setminus \{0\}$ such that $\langle y, v_i \rangle < 1$ for all $i = 1, \dots, N$ and $\langle y, x \rangle > 1$ (separating hyperplane). By the first condition we get that $y \in P^o$ and so y is a convex combination of the y_i (by Minkowski theorem, since P^o is closed and bounded). But then we get that $1 < \langle y, x \rangle = \sum_{i=1}^m \lambda_i \langle y_i, x \rangle$ and so at least one of the $\langle y_i, x \rangle$ is greater than 1. This completes the proof.

7. (Linear image of a polyhedron) Let $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron and assume that P is bounded. Let $\pi : \mathbb{R}^n \to \mathbb{R}^k$ be a linear map. Show that $\pi(P)$ is also a polyhedron, i.e., it is the intersection of a finite number of halfspaces. (*Hint:* Use questions 2 and 3 above. *Note:* the assumption P is bounded is for simplicity; the result is still true if P is not bounded).

   Solution: We know from question 5 that P has a finite number of extreme points $v_1, \dots, v_N$. Also since P is bounded (and closed) we have $P = \operatorname{conv}(v_1, \dots, v_N)$. Thus $\pi(P) = \operatorname{conv}(\pi(v_1), \dots, \pi(v_N))$ is a (bounded) polyhedron by question 6.

8. (Farkas' lemma) Let $K = \{x \in \mathbb{R}^n : Ax \geq 0\}$ be a polyhedral cone. Show that $K^* = \operatorname{cone}(a_1, \dots, a_m)$ where $a_1, \dots, a_m$ are the rows of A. Deduce Farkas' lemma (conic version): if $\langle y, x \rangle \geq 0$ for all $x \in K$ then there exists $\lambda \geq 0$ such that $y^T = \lambda^T A$.

Solution: Define $K' = \text{cone}(a_1, \ldots, a_m)$. Clearly $K' \subseteq K^*$. For the reverse inclusion we will use a separating hyperplane argument and we first want to show that K' is closed. To show that K' is closed, using a similar argument as in the proof of Carathéodory theorem one can show that any $y \in K'$ can be written as a conic combination of *linearly independent vectors* $\{a_i : i \in I\}$ where $I \subseteq \{1, \ldots, m\}$. Thus this means that

$$K' = \bigcup_{\substack{I \subseteq \{1,\ldots,m\} \text{ s.t.} \\ \{a_i\}_{i \in I} \text{ linearly independent}}} \text{cone}(\{a_i : i \in I\}). \tag{1}$$

Each $\text{cone}(\{a_i : i \in I\})$ is closed, because it is the image of $\mathbb{R}^I_+$ via an *injective* linear map (the map that sends $\lambda \in \mathbb{R}^I$ to $\sum_{i \in I} \lambda_i a_i$). Thus (1) shows that K' is closed since it is a union of a *finite* number of closed sets. We are now ready to show that $K^* \subseteq K'$. Assume $y \notin K'$. By the separating hyperplane theorem there exist $x \neq 0$ such that $\langle a_i, x \rangle \geq 0$ and $\langle y, x \rangle < 0$. The first condition means that $x \in K$. The latter condition thus implies that $y \notin K^*$. We have thus shown that $K' = K^*$. Farkas' lemma is then an immediate consequence: assume $\langle y, x \rangle \geq 0$ for all $x \in K$. Then $y \in K^* = K'$. This means that y is a conic combination of the a_is, i.e., there exists $\lambda \geq 0$ such that $y = \sum_{i=1}^m \lambda_i a_i$, in other words $y^T = \lambda^T A$.

9. Let K be a closed convex cone in $\mathbb{R}^n$.

   (a) Show that the following conditions are equivalent:

      (i) K has nonempty interior

      (ii) $\text{span}(K) = \mathbb{R}^n$

      (iii) For any $w \in \mathbb{R}^n \setminus \{0\}$ there exists $x \in K$ such that $\langle w, x \rangle \neq 0$.

   (b) Show that K is pointed if and only K^* has nonempty interior.

   (c) Show that $y \in \text{int}(K^*)$ if and only if $\langle y, x \rangle > 0$ for all $x \in K \setminus \{0\}$.

   Solution:

   (a) The implication $(i) \Rightarrow (ii)$ is clear. We prove $(ii) \Rightarrow (i)$. Let $a_1, \ldots, a_n$ be elements of K that span $\mathbb{R}^n$ and let $x_0 = a_1 + \cdots + a_n$. We claim that $x_0 \in \text{int}(K)$. For this let A be the $n \times n$ matrix whose columns are $a_1, \ldots, a_n$ and define a norm N on $\mathbb{R}^n$ as follows: $N(x) = \|A^{-1}x\|_\infty$. That N is a norm is easy to establish. Consider the ball $B = \{x \in \mathbb{R}^n : N(x - x_0) \leq 1/2\}$ around x_0. We will show that $B \subseteq K$ which will establish that $x_0 \in \text{int}(K)$. Note that $N(x - x_0) \leq 1/2$ if and only if $x - x_0 = \lambda_1 a_1 + \cdots + \lambda_n a_n$ where $|\lambda_i| \leq 1/2$. By definition of x_0 we thus have that $x \in B$ if and only $x = \mu_1 a_1 + \cdots + \mu_n a_n$ where $1/2 \leq \mu_i \leq 3/2$. Thus each $x \in B$ is a conic combination of $a_1, \ldots, a_n$ and thus lies in K. This shows that $B \subseteq K$ which is what we wanted.
   The equivalence $(ii) \iff (iii)$ is simple linear algebra (observe that the negation of (iii) means that K lies in a hyperplane).

   (b) If K is not pointed, there exists $x \neq 0$ such that $x \in K$ and $x \in -K$. This implies that for any $y \in K^*$ we have $\langle x, y \rangle \geq 0$ and $\langle -x, y \rangle \geq 0$ i.e., $\langle x, y \rangle = 0$. This means that K^* lies in a hyperplane and so has empty interior (cf. previous question).
   For the converse, assume K^* has empty interior. Then (by the previous question applied to K^*) there exists $w \in \mathbb{R}^n \setminus \{0\}$ such that $\langle w, y \rangle = 0$ for all $y \in K^*$. This means that w and $-w$ lie in $(K^*)^* = K$ (since K is closed and convex) which in turns means that K is not pointed.

   (c) Let $y \in \text{int}(K^*)$ and $x \in K \setminus \{0\}$. We want to show that $\langle y, x \rangle > 0$. Since $y \in \text{int}(K^*)$ we know that for small enough $\epsilon > 0$ we have $y - \epsilon x \in K^*$. This means that $\langle y - \epsilon x, x \rangle \geq 0$ i.e., $\langle y, x \rangle \geq \epsilon \|x\|_2^2 > 0$.
   We now show the converse. Assume $y \in \mathbb{R}^n$ is such that $\langle y, x \rangle > 0$ for all $x \in K \setminus \{0\}$. We will show that $y \in \text{int}(K^*)$. Let $\epsilon = \min_{x \in K, \|x\|=1} \langle y, x \rangle$ and note that $\epsilon > 0$ by our assumption and

the fact that $\{x \in K, \|x\| = 1\}$ is compact. We will now prove that $y \in \text{int}(K^*)$ by proving that $y + r \in K^*$ for any r with $\|r\|_2 \leq \epsilon$. Let $x \in K$ and note that for any such r we have

$$\langle y + r, x/\|x\|\rangle = \langle y, x/\|x\|\rangle + \langle r, x/\|x\|\rangle \overset{(*)}{\geq} \epsilon - \epsilon \geq 0$$

where we used the Cauchy-Schwarz inequality in $(*)$. Thus this shows that $\langle y+r, x\rangle \geq 0$ for any $x \in K$ and thus that $y + r \in K^*$. This is valid for any r with $\|r\|_2 \leq \epsilon$ and thus shows that $y \in \text{int}(K^*)$.

10. In this exercise we will prove Minkowski's theorem for closed convex pointed cones (Theorem 2.2 in lecture 2). Let K be a closed convex cone.

    (a) Assume that there exists $y \in \mathbb{R}^n$ such that $\langle y, x\rangle > 0$ for all $x \in K \setminus \{0\}$. Show how to prove the theorem in this case. *(hint: define $C = \{x \in K$ s.t. $\langle y, x\rangle = 1\}$; show that C is a compact convex set and apply Minkowski's theorem for compact convex sets to C).*

    (b) Use Question 9 to show that when K is a closed *pointed* convex cone, there exists $y \in \mathbb{R}^n$ verifying $\langle y, x\rangle > 0$ for all $x \in K \setminus \{0\}$. Use part (a) to conclude proof of Theorem 2.2.

    Solution:

    (a) Define $C = \{x \in K \text{ s.t. } \langle y, x\rangle = 1\}$. Note that C is closed since it is the intersection of two closed sets. We are going to show that C is bounded. Assume x_n is a sequence in C such that $\|x_n\| \to \infty$. Note that $\langle y, x_n/\|x_n\|\rangle = 1/\|x_n\| \to 0$ as $n \to \infty$. We can assume (after extracting a convergent subsequence) that $x_n/\|x_n\| \to z \in K$ and so we get $\langle y, z\rangle = 0$. This contradicts our assumption on y (since $z \in K \setminus \{0\}$). We have thus shown that C is closed and bounded.

    The following lemma is straightforward and shows that extreme points of C span extreme rays of K:

    Lemma 1. *Assume v is an extreme point of C. Then the ray $S = \{\lambda v : \lambda \geq 0\}$ is an extreme ray for K.*

    Proof. For simplicity of notation we will denote ℓ the linear form $\ell(x) := \langle y, x\rangle$. Assume $x, y \in K \setminus \{0\}$ such that $x + y \in S$. We need to show that $x, y \in S$. Note that we have:

    $$\left(\frac{\ell(x)}{\ell(x) + \ell(x)}\right)\frac{x}{\ell(x)} + \left(\frac{\ell(y)}{\ell(x) + \ell(y)}\right)\frac{y}{\ell(y)} = \frac{x + y}{\ell(x) + \ell(y)} = v. \tag{2}$$

    The last equality is because we have (by assumption) $x + y = \lambda v$ for some $\lambda \geq 0$ and $\ell(v) = 1$ which imply $\lambda = \ell(x) + \ell(y)$. Since $x/\ell(x)$ and $y/\ell(y)$ are in C, the assumption that v is an extreme point, together with the fact that $\ell(x), \ell(y) > 0$ shows that $x/\ell(x) = y/\ell(y) = v$. Thus this shows that $x, y \in \mathbb{R}_+ v$ as desired. □

    We now complete the proof of Minkowski's theorem for K. Let $x \in K$; we need to show that x is a conic combination of extreme rays of K. Since $x/\ell(x)$ (where $\ell(x) = \langle y, x\rangle$) lies in the compact convex set C, Minkowski's theorem for compact convex sets says that $x/\ell(x)$ is a convex combination of extreme points of C, i.e., $x/\ell(x) = \sum_i \lambda_i v_i$ where each v_i is an extreme point of C and the λ_i satisfy $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$. From the lemma we just proved we know that the ray spanned by each v_i is an extreme ray of K. We finally get that $x = \sum_i (\ell(x)\lambda_i) v_i$ which expresses x is a conic combination of extreme rays of K. This completes the proof.

    (b) We know from Question 9 that since K is pointed closed convex cone then K^* has nonempty interior. From that same exercise we also know that any element $y \in \text{int}(K^*)$ satisfies the desired requirement $\langle y, x\rangle > 0$ for all $x \in K \setminus \{0\}$.

11. For each of the following sets: show that it is a closed convex pointed cone with nonempty interior, identify the extreme rays and give a simple expression for the dual cone:

(a) $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0 \; \forall i = 1, \ldots, n\}$

(b) $\mathbf{Q}^3 = \{(x,t) \in \mathbb{R}^2 \times \mathbb{R}_+ : \|x\|_2 \leq t\}$

(c) $K = \{(x,y,z) \in \mathbb{R}^2_+ \times \mathbb{R} : \sqrt{xy} \geq |z|\}$

Show that there is a linear invertible map $A : \mathbb{R}^3 \to \mathbb{R}^3$ such that $A(\mathbf{Q}^3) = K$.

Solution:

(b) Convexity of $\mathbf{Q}^3$ follows directly from convexity of the function $x \mapsto \|x\|_2 = \sqrt{x_1^2 + x_2^2}$. Closedness of $\mathbf{Q}^3$ follows from continuity of the ℓ_2 norm. It is easy to verify that the point $(0,0,1)$ lies in the interior of $\mathbf{Q}^3$. Pointedness of $\mathbf{Q}^3$ can also be easily verified. For the standard inner product on $\mathbb{R}^3$ the cone $\mathbf{Q}^3$ is self-dual, i.e., $(\mathbf{Q}^3)^* = \mathbf{Q}^3$. This can be proved using Cauchy-Schwarz inequality:

- Let $(a_1, a_2, b) \in \mathbf{Q}^3$. We want to show that it lies in the dual of $\mathbf{Q}^3$, namely that for any $(x_1, x_2, t) \in \mathbf{Q}^3$ it holds $a_1x_1 + a_2x_2 + bt \geq 0$. By Cauchy-Schwarz inequality we know that $a_1x_1 + a_2x_2 \geq -\|a\|_2\|x\|_2$. Since $\|a\|_2 \leq b$ and $\|x\|_2 \leq t$ we get that $a_1x_1 + a_2x_2 + bt \geq 0$. This is valid for any $(x_1, x_2, t) \in \mathbf{Q}^3$ and thus shows that $(a_1, a_2, b) \in (\mathbf{Q}^3)^*$.
- We now show the reverse inclusion. Assume $(a_1, a_2, b) \in (\mathbf{Q}^3)^*$. We want to show that $(a_1, a_2, b) \in \mathbf{Q}^3$. Let $x = -a \in \mathbb{R}^2$ and $t = \|a\|_2$. Since $(x,t) \in \mathbf{Q}^3$ we know by definition of duality that $\langle a, x\rangle + bt \geq 0$, i.e., $-\|a\|_2^2 + b\|a\|_2 \geq 0$ and so $\|a\|_2 \leq b$. This shows that $(a_1, a_2, b) \in \mathbf{Q}^3$ which is what we wanted.

The extreme rays of $\mathbf{Q}^3$ are those spanned by $(a, 1)$ where $\|a\|_2 = 1$. Proof:

- We first prove that any such ray is extreme. Assume $(x,t) \in \mathbf{Q}^3$ and $(x', t') \in \mathbf{Q}^3$ are such that $(x + x', t + t') = (a, 1)$. We want to show that (x,t) and (x', t') are a nonnegative multiple of $(a, 1)$. Note that we have $1 = \|a\| = \|x + x'\| \leq \|x\| + \|x'\| \leq t + t' = 1$ and thus all the intermediate inequalities must be equalities. One can easily finish the proof by noting that the equality case for the triangle inequality says that x and x' must be collinear.
- We now show that any other ray spanned by (a,t) where $\|a\| < t$ cannot be extreme. Indeed note that we can write $(a,t) = (a, \|a\|_2) + (0, t - \|a\|_2)$ and that the two summands do not lie on the ray spanned by (a,t). Thus this shows that the ray spanned by (a,t) is not extreme.

(c) Let $A(x_1, x_2, t) = (t - x_1, t + x_1, x_2)$. It is easy to see that $A(\mathbf{Q}^3) = K$. Properties of K then follow easily from the properties proved for $\mathbf{Q}^3$. Note also that $(x,y,z) \in K \Leftrightarrow \left(\begin{smallmatrix} x & z \\ z & y \end{smallmatrix}\right) \succeq 0$.