Linear Optimization in a Nutshell
(from professor Boris Vexler's notes to LaTeX)

Lado Turnamidze

2024

Week 1

$$\min_{x \in \mathbb{R}^n} c^T x \text{ with } Ax \leq b \text{ and } Bx = d$$

A typical linear optimization problem involves minimizing a linear function $c^T x$ where $c, x \in \mathbb{R}^n$ subject to constraints of the form $Ax \leq b$(Component-wise inequality) and $Bx = d$. Here, $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $B \in \mathbb{R}^{p \times n}$ and $d \in \mathbb{R}^p$.

$$Ax \leq b = \begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \leq b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \leq b_m \end{cases}$$

The admissible set X is the set of all vectors $x \in \mathbb{R}^n$ that satisfy the constraints of the linear optimization problem: $X = \{x \in \mathbb{R}^n \mid Ax \leq b, Bx = d\}$.

A vector $\bar{x} \in \mathbb{R}^n$ is a global solution to the linear optimization problem if $\bar{x} \in X$ and $c^T \bar{x} \leq c^T x$ for all $x \in X$.

A set $X \subseteq \mathbb{R}^n$ is called convex if for any two points $x, y \in X$ and for any $\lambda \in [0, 1]$: $\lambda x + (1 - \lambda)y \in X$.

A convex function $f : X \to \mathbb{R}$ on a convex set $X \subset \mathbb{R}^n$ satisfies the inequality $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ for all $x, y \in X$ and $\lambda \in [0; 1]$.

A convex optimization problem is an optimization problem where the objective function f is convex and the admissible set X is also convex. Geometrically, for a given pair $x, y \in X$ the graph of f should be below a line connecting x and y (This line is $\lambda f(x) + (1 - \lambda)f(y)$).

Theorem: For a convex problem, a local minimizer is a global minimizer.

Theorem (Last Semester): Let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable and let X be convex. Then, f is convex on X if and only if $\nabla f(x)^T(y-x) \le f(y) - f(x)$ for all $x, y \in X$.

Theorem: Let $g : \mathbb{R}^n \to \mathbb{R}^m$, let every component of g be convex($g_i : \mathbb{R}^n \to \mathbb{R}$ for every $1 \le i \le m$) and let $h : \mathbb{R}^n \to \mathbb{R}^p$ be affine linear(i.e. $h(x) = Ax - b$). Then $X = \{g(x) \le 0 \mid h(x) = 0\}$ is a convex set.

Convex Optimization Problem: $\min f(x)$ subject to $g(x) \le 0$, $h(x) = 0$. $f : \mathbb{R}^n \to \mathbb{R}$ is convex, $g_i : \mathbb{R}^n \to \mathbb{R}^m$ is convex with $1 \le i \le n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ is affine linear. Linear optimization is a special case of convex optimization.

Theorem: Consider a convex problem $\min f(x), x \in X$. Let $X \subset \mathbb{R}^n$ be convex, and $f : X \to \mathbb{R}$ be convex and continuously differentiable. Then, $\bar{x} \in X$ is a global minimizer if and only if $\nabla f(\bar{x})^T(x - \bar{x}) \ge 0$ for all $x \in X$.

Week 2

Let V be a linear vector space(typically $V = \mathbb{R}^n$). Let $X \subset V$:

1. X is a (linear) subspace of V, if $X \neq \emptyset$ and $\sum_{i=1}^m \lambda_i x_i \in X$, $\forall m \in \mathbb{N}$, $\forall x_i \in X$, $\forall \lambda_i \in \mathbb{R}$.
2. X is an affine subspace of V, if $\sum_{i=1}^m \lambda_i x_i \in X$, $\forall m \in \mathbb{N}$, $\forall x_i \in X$, $\forall \lambda_i \in \mathbb{R}$ with $\sum_{i=1}^m \lambda_i = 1$.
3. X is a convex subset of V, if $\sum_{i=1}^m \lambda_i x_i \in X$, $\forall m \in \mathbb{N}$, $\forall x_i \in X$, $\forall \lambda_i \in \mathbb{R}$ with $\sum_{i=1}^m \lambda_i = 1$ and $\forall i, \lambda_i \ge 0$.

Let $X \subset V$ be a subset. Then,

1. $\operatorname{span}(X) = \bigcap_{X \subset L,\, L \text{ is a linear subspace of } V} L$ is called span of X or linear hull of X(So $\operatorname{span}(X)$ is the smallest linear subspace containing X).
2. $\operatorname{aff}(X) = \bigcap_{X \subset L,\, L \text{ is an affine subspace of } V} L$ is called affine hull of X or affine span of X.
3. $\operatorname{conv}(X) = \bigcap_{X \subset C, \text{ C is a convex set}} C$ is called convex hull of X.

1. if $X \neq \emptyset$, then

$$\operatorname{span}(X) = \left\{ \sum_{i=1}^m \lambda_i x_i \mid x_i \in X, \lambda_i \in \mathbb{R}, m \in \mathbb{N} \right\}$$

2.
$$\mathrm{aff}(X) = \left\{ \sum_{i=1}^{m} \lambda_i x_i \mid x_i \in X, \sum_{i=1}^{m} \lambda_i = 1, m \in \mathbb{N} \right\}$$

3.
$$\mathrm{conv}(X) = \left\{ \sum_{i=1}^{m} \lambda_i x_i \mid x_i \in X, \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1, m \in \mathbb{N} \right\}$$

Lemma:

1. X is an affine subspace if and only if $\exists a \in V$ such that $U = X - a = \{y | y = x - a, x \in X\}$ is a linear subspace.
2. Let $X \subset V$ be an affine subspace: X is a linear subspace iff $0 \in X$.
3. Let $X \subset V, a \in X$, then, $\mathrm{aff}(X) = a + \mathrm{span}(X - a)$.

Affine Dimension: Let $X \subset V$, $X \neq \emptyset$ be an affine subspace. We define the affine dimension $\dim(X) = \dim(X - a), a \in X \neq \emptyset$.

Let $\emptyset \neq X \subset V$ be an arbitrary subset. We define the affine dimension of X: $\dim(X) = \dim(\mathrm{aff}(X))$.

REMARK: Let X be an affine subspace with finite dimension. Then, there are 2 cases:

- $0 \in X \implies X$ is a linear subspace. $\dim(X) = \dim(\mathrm{span}(X))$
- $0 \notin X \implies \dim(X) = \dim(\mathrm{span}(X)) - 1$

Vectors $x_0, x_1, \cdots, x_m \in V$ are affine independent if $\dim(\mathrm{aff}(x_0, x_1, \cdots, x_m)) = m$.

Lemma:
Let $x_0, \cdots, x_m \in V$ be $(m+1)$ vectors. The following statements are equivalent:

1. $x_0, \cdots, x_m$ are affine independent.
2. $x_1 - x_0, \cdots, x_m - x_0$ are linearly independent.
3. $\begin{pmatrix} x_0 \\ 1 \end{pmatrix}, \begin{pmatrix} x_1 \\ 1 \end{pmatrix}, ..., \begin{pmatrix} x_m \\ 1 \end{pmatrix} \in V \times \mathbb{R}$ are linearly independent.

Let $x_0, x_1, ..., x_m$ be affine independent. For every $x \in \text{aff}(\{x_0, x_1, ..., x_m\})$ there is a unique representation $x = \sum_{i=0}^m \lambda_i x_i$ with $\sum_{i=0}^m \lambda_i = 1$. The coefficients λ_i are called barycentric coordinates.

Let $X \subset \mathbb{R}^n$ and $\tau : \mathbb{R}^n \to \mathbb{R}^m$ an be affine map: $\tau\left(\sum\limits_{i=1}^k \mu_i v_i\right) = \sum\limits_{i=1}^k \mu_i \tau(v_i)$ with $\sum\limits_{i=1}^k \mu_i = 1$. Then, $\tau(\text{conv}(X)) = \text{conv}(\tau(X))$.

$\text{aff}(\text{aff}(X)) = \text{aff}(X)$ and $\text{conv}(\text{conv}(X)) = \text{conv}(X)$.

Convex hull of a bounded set is bounded($a_1, a_2, \dots, a_k \in A$):

$$\|\sum_{i=1}^k \lambda_i a_i\| \leq \sum_{i=1}^k \lambda_i \|a_i\| \overset{C:=\sup_{x\in A}\|x\|<\infty}{\leq} C \cdot \sum_{i=1}^k \lambda_i = C$$

Convex hull of a closed set is not necessarily closed.

Week 3

Let $x_0, \cdots, x_m \in V$ be affine independent, then $\text{conv}(\{x_0, \cdots, x_m\})$ is called $m-$simplex and x_i are corners or vertices of this $m-$simplex.

Lemma:
A set $X \subset \mathbb{R}^n$ is an affine subspace iff $\exists m \in \mathbb{N}$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ such that $X = \{x \in \mathbb{R}^n \mid Ax = b\}$.

Orthogonal compliment: $U^\perp = \{y \in \mathbb{R}^n \mid y^T v = 0, \forall v \in U\}$ So we can rewrite $\mathbb{R}^n$ in terms of direct sum of space and its complement: $\mathbb{R}^n = U \oplus U^\perp$.

Theorem of **Caratheodory**:
Let $X \subset \mathbb{R}^n, X \neq \emptyset$. For every $x \in \text{conv}(X)$ there are affine independent points $x_1, \cdots, x_l \in X$ such that $x \in \text{conv}(\{x_1, \cdots, x_l\})$ with $l \leq n+1$.

Recall: $X \subset \mathbb{R}^n$

1. $\text{int}(X) = \{x \in X \mid \exists \varepsilon > 0, \mathbb{B}_\varepsilon(x) \subset X\}$
2. $\partial X = \{x \in \mathbb{R}^n \mid \forall \varepsilon > 0, \mathbb{B}_\varepsilon(x) \cap X \neq \emptyset, \mathbb{B}_\varepsilon(x) \cap (\mathbb{R}^n \setminus X) \neq \emptyset\}$
3. Closure of $X \to \bar{X} = X \cup \partial X$

Recall: Compact means closed and bounded.

$$\begin{cases} f : \mathbb{R}^n \to \mathbb{R}^m \text{ continuous,} \\ X \subset \mathbb{R}^n \text{ compact} \end{cases} \implies f(X) \text{ is compact.}$$

Theorem: Let $X \subset \mathbb{R}^n$ be compact. Then $\text{conv}(X)$ is compact.

Product of convex sets is convex: $X \subset V$ convex, $Y \subset W$ convex $\implies$ $X \times Y \subset V \times W$ convex.

Minkowski sum is convex: $X, Y \subset V : X + Y = \{x + y \mid x \in X, y \in Y\}$. X, Y are convex $\implies$ $X + Y$ is convex.

For a family X_α of convex sets ($\alpha \in A \leftarrow$ index set) $\implies$ $\bigcap\limits_{\alpha \in A} X_\alpha$ is convex.

$$\begin{cases} X \subset V \text{ is convex} \\ A : V \to W, b \in W \text{ is a linear mapping} \end{cases} \implies A(X)+b = \{A(x)+b \mid x \in X\} \text{ is convex}$$

Hyperplane in $\mathbb{R}^n$: $a \in \mathbb{R}^n \setminus \{0\}, b \in \mathbb{R}$

$$H(a, b) = \{x \in \mathbb{R}^n \mid a^T x = b\} \text{ is convex}$$

Half-Spaces: $a \in \mathbb{R}^n \setminus \{0\}, b \in \mathbb{R}$

$$H_{\leq}(a, b) = \{x \in \mathbb{R}^n \mid a^T x \leq b\}$$
$$H_{\geq}(a, b) = \{x \in \mathbb{R}^n \mid a^T x \geq b\}$$
$$H_{<}(a, b) = \{x \in \mathbb{R}^n \mid a^T x < b\}$$
$$H_{>}(a, b) = \{x \in \mathbb{R}^n \mid a^T x > b\}$$

Polyhedron: $X = \{x \in \mathbb{R}^n \mid Ax \leq b, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m\}$.

$X = \bigcap\limits_{i=1}^{m} H_{\leq}(a_i, b_i)$ is convex.

Lemma: Let $X \subset \mathbb{R}^n$ be convex. Then:

1. $\bar{X}$ is convex.
2. $\text{int}(X)$ is convex.

Week 4

Let $X \subset \mathbb{R}^n$.

$\text{relint}(X) = \{x \in X \mid \exists \varepsilon > 0, \mathbb{B}_\varepsilon(x) \cap \text{aff}(X) \subset X\}$.

$\text{relbd}(X) = \{x \in \text{aff}(X) \mid \forall \varepsilon > 0, X \cap \mathbb{B}_\varepsilon(x) \neq \emptyset, (\text{aff}(X) \setminus X) \cap \mathbb{B}_\varepsilon(x) \neq \emptyset\}$.

A set $O \subset W$ is called relative open in W iff $\exists \tilde{O} \subset V$ open: $O = \tilde{O} \cap W$.

$x \in \text{relint}(X) \Leftrightarrow \exists O \subset X$ relative open in $W = \text{aff}(X)$ with $x \in O$.

REMARK: if $a \in \mathbb{R}^n,\ \text{relint}(X) - a = \text{relint}$(X - a).

Lemma: Let $X \subset \mathbb{R}^n$

1. if $X \neq \emptyset$ and convex $\implies$ relint$(X) \neq \emptyset$
2. X is convex $\implies$ relint(X) is convex.

Recall: Preimages of open sets are open for continuous functions.

Consider a basis $x_1, \cdots, x_m$ of U, a linear mapping $T : Z \subset \mathbb{R}^m \to \mathbb{R}^n$

$$T : Z \mapsto \sum_{i=1}^{m} z_i x_i \in U \subset \mathbb{R}^n$$

Consider $T : \mathbb{R}^m \to U$(bijective) $\implies \exists T^{-1} : U \to \mathbb{R}^m$. T and T^{-1} are linear and continuous.

$M \subset \mathbb{R}^m$ open $\implies T(M) = (T^{-1})^{-1}(M)$ relative open in U.

$K \subset U$ relative open $\implies \exists O \subset \mathbb{R}^n : K = O \cap U \implies T^{-1}(O) = T^{-1}(O \cap U) = T^{-1}(K)$ open.

Claim: $\text{relint}(X) = T(\text{int}(T^{-1}(X)))$.

Projection on convex sets:

$X \subset \mathbb{R}^n, X \neq \emptyset$, closed, convex. $P_X : \mathbb{R}^n \to X, P_X(x) = \text{argmin}_{y \in X} \|y - x\|$. If $X = \text{span}(\{s^1, \cdots, s^k\})$ with vectors $s^1, \cdots, s^k$, orthonormal then it holds $P_X(y) = \sum_{i=1}^k s^i (s^i)^T y$.

Theorem: Let $X \subset \mathbb{R}^n, X \neq \emptyset$, closed, convex. Then $P_X : \mathbb{R}^n \to X$ is well defined.

Theorem: Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$, closed and convex. Let $x \in \mathbb{R}^n$. Then, $z = P_X(x) \Leftrightarrow z \in X$ and $(z-x)^T(y-z) \geq 0 \quad \forall y \in X$.

Theorem: Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$, closed, convex. Then $P_X : \mathbb{R}^n \to X$ is Lipschitz continuous with $L = 1$: $\|P_X(x_1) - P_X(x_2)\| \leq \|x_1 - x_2\|$.

Let $X_1 \subset \mathbb{R}^n$, $X_2 \subset \mathbb{R}^n$. Let $H(a,b) = \{x \in \mathbb{R}^n \mid a^T x = b, a \in \mathbb{R}^n \setminus \{0\}, b \in \mathbb{R}\}$ be a hyperplane.

1. $H(a,b)$ separates X_1 and X_2 if $X_1 \subset H_{\leq}(a,b)$ and $X_2 \subset H_{\geq}(a,b)$.
2. $H(a,b)$ separates strictly X_1 and X_2 if $X_1 \subset H_{<}(a,b)$ and $X_2 \subset H_{>}(a,b)$.
3. $H(a,b)$ separates strongly X_1 and X_2 if $\exists \epsilon > 0 : X_1 \subset H_{\leq}(a, b-\epsilon)$ and $X_2 \subset H_{\geq}(a, b+\epsilon)$.
4. $H(a,b)$ separates properly X_1 and X_2 if $H(a,b)$ separates X_1 and X_2 and $\exists x_1 \in X_1$ and $\exists x_2 \in X_2 : a^T(x_1 - x_2) \neq 0$.
5. $H(a,b)$ supports the set $X \subset \mathbb{R}^n$ at the point $x \in X$ if $x \in H(a,b)$ and $X \subset H_{\leq}(a,b)$ or $X \subset H_{\geq}(a,b)$.

Strong Separation $\implies$ Strict Separation $\implies$ Proper Separation $\implies$ Separation.

First Separation Theorem: Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$, closed, convex. Let $y \notin X$. Then $\exists$ a hyperplane $H(a,b)$ which separates strongly X and $\{y\}$.

Radon's Theorem: Every set of affinely dependent points in $\mathbb{R}^n$ (i.e. especially every set containing $n+2$ points) can be partitioned into two disjoint sets, whose convex hulls have at least one point in common.

Helly's Theorem: Let $K_1, \cdots, K_m \subset \mathbb{R}^n, m \geq n+1$ be convex such that the intersection $\cap_{i \in I} K_i \neq \emptyset$, for all $I \subset \{1, \cdots, m\}$ with $|I| = n+1$ then it holds $\cap_{i=1}^m K_i \neq \emptyset$.

Week 5

Second Separation Theorem: Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$ closed and convex. Let $y \in \partial X$ (boundary point). Then $\exists$ hyperplane $H(a,b)$ ($a \in \mathbb{R}^n \setminus \{0\}$, $b \in \mathbb{R}$) which supports X at y.

Bolzano-Weierstrass Theorem: Each bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Theorem: Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$, closed and convex. Then we can represent X as an intersection of half spaces: $X = \bigcap_{\substack{\text{hyperplane } H(a,b) \text{ supports } X \\ X \subset H_\leq(a,b)}} H_\leq(a,b)$.

$A, B \subset \mathbb{R}^n$ are closed and one of them is bounded $\implies A + B$ is closed.

Third Separation Theorem: Let X_1, $X_2 \subset \mathbb{R}^n$ non-empty, closed and convex. Let $X_1 \cap X_2 = \emptyset$. And one of these sets is bounded. Then $\exists$ hyperplane $H(a,b)$, which strongly separates X_1 and X_2. This means: $\exists a \in \mathbb{R}^n \backslash \{0\}$, $b_1, b_2 \in \mathbb{R}$, $b_2 > b_1$, $a^T x_1 \leq b_1\ \forall x_1 \in X_1, a^T x_2 \geq b_2\ \forall x_2 \in X_2$.

Let $X \subset \mathbb{R}^n$, $X \neq \emptyset$ and convex. Then:

1. $\text{int}(X) = \text{int}(\overline{X})$
2. $\text{relint}(X) = \text{relint}(\overline{X})$
3. $\overline{\text{relint}(X)} = \overline{X}$

Week 6

Remaining Separation Theorems:

Fourth Separation Theorem: Let $X \subset \mathbb{R}^n, X \neq \emptyset$ and convex. Let $y \notin \text{int}(X)$. Then, there is a separation hyperplane between X and $\{y\}$.

Fifth Separation Theorem: Let $X_1, X_2 \subset \mathbb{R}^n$ be non-empty and convex. Let $X_1 \cap X_2 = \emptyset$. Then, there is a separation hyperplane.

Theorem: Let X_1, X_2 be two non-empty convex sets. Then, X_1 and X_2 can be properly separated $\Leftrightarrow \text{relint}(X_1) \cap \text{relint}(X_2) = \emptyset$.

Cones:
A set $K \subset \mathbb{R}^n$ is called a cone if for $x \in K,\ \lambda \in \mathbb{R},\ \lambda > 0 \implies \lambda x \in K$.

Polar Cone of K is $K^0 = \{y \in \mathbb{R}^n \mid y^T x \leq 0,\ \forall x \in K\}$.

Lemma: Let K be a cone. Then, K^0 is a convex and closed cone.

Let $X \subset \mathbb{R}^n$ be convex, $x \in X$. We define Tangential Cone of X at x: $T(X,x) = \overline{R_+(X-x)} = \overline{\{\lambda(y-x) \mid y \in X,\ \lambda > 0\}}$.

Normal Cone of X at $x \in X$: $N(X,x) = T(X,x)^0$.

Week 7

$T(X,x)$ is a closed, convex cone.

if $x \in \text{int}(X) \implies T(X,x) = \mathbb{R}^n, \quad N(X,x) = \{0\}$.

if $x \in \text{relint}(X) \implies T(X,x) = \text{span}(X - x)$.

$\bar{x} \in X$, $\nabla f(\bar{x})^T(x - \bar{x}) \geq 0 \quad \forall x \in X \Leftrightarrow \bar{x} \in X$, $\nabla f(\bar{x})^T s \geq 0 \quad \forall s \in T(X,\bar{x}) \Leftrightarrow \bar{x} \in X$, $(-\nabla f(\bar{x}))^T s \leq 0 \quad \forall s \in T(X,\bar{x}) \Leftrightarrow \bar{x} \in X$, $-\nabla f(\bar{x}) \in N(X,\bar{x})$

Theorem:

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and continuously differentiable, $X \subset \mathbb{R}^n, X \neq \emptyset$ and convex. Then:
$\bar{x}$ is solution of convex optimization problem $\Leftrightarrow \bar{x} \in X$ and $-\nabla f(\bar{x}) \in N(X,\bar{x})$

Active index set for $x \in X$: $\mathcal{A}(x) = \{1 \leq j \leq m \mid a_j^T x = b_j\}$.

Inactive index set for $x \in X$: $I(x) = \{1 \leq j \leq m \mid a_j^T x < b_j\}$.

Lemma:

Let $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, D \in \mathbb{R}^{p \times n}, e \in \mathbb{R}^p$. Let $X = \{x \in \mathbb{R}^n \mid Ax \leq b, Dx = e\}$. Let $x \in X$. Then:

$$T(X,x) = \{s \in \mathbb{R}^n \mid A_{\mathcal{A}(x)} s \leq 0, \quad Ds = 0\}$$

Definition: $K^{00} = \{z \in \mathbb{R}^n \mid z^T y \leq 0 \quad \forall y \in K^0\}$

Theorem:

Let $K \subset \mathbb{R}^n$ be a non-empty cone. Then:

1. $K \subset K^{00}$
2. $K^{00} = \overline{\text{conv}(K)}$

Convex Conical Hull Definition 1:

$$\text{cone}(X) = \bigcap_{\substack{K \text{ convex cone} \\ X \cup \{0\} \subset K}} K$$

Convex Conical Hull Definition 2:

$$\operatorname{cone}(X) = \left\{ \sum_{i=1}^{m} \lambda_i x_i \mid m \in \mathbb{N},\, x_i \in X,\, \lambda_i \in \mathbb{R},\, \lambda_i \geq 0 \right\}$$

Caratheodory Theorem for convex conical hulls

Let $X \subset \mathbb{R}^n, X \neq \emptyset$. For every $v \in \operatorname{cone}(X)$ there are linearly independent vectors $x_1, x_2, \ldots, x_m \in X$ $(m \leq n)$ and $\lambda_i \geq 0$ and such that $v = \sum_{i=1}^{m} \lambda_i x_i$

Week 8

Theorem: Let $a_1, a_2, ..., a_n \in \mathbb{R}^l$. Then $K = \operatorname{cone}(\{a_1, a_2, ..., a_n\})$ is closed.

Lemma of Farkas:

Let $A \in \mathbb{R}^{m \times n}$ with $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix}$, $a_i \in \mathbb{R}^n$. We consider $K = \{s \in \mathbb{R}^n \mid As \leq 0\}$. Then, $K^0 = \operatorname{cone}(\{a_1, a_2, ..., a_m\})$.

Alternative Formulation of Lemma of Farkas:

Let $A \in \mathbb{R}^{m \times n}$, $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix}$, $a_i \in \mathbb{R}^n$. Let $c \in \mathbb{R}^n$. Then the following two statements are equivalent:

1. For every $s \in \mathbb{R}^n$ with $As \leq 0$ holds $c^T s \leq 0$.
2. $\exists \lambda \in \mathbb{R}^m, \lambda \geq 0$ with $c = A^T \lambda$.

Corollary(Lemma of Farkas in the case of inequality and equality constraints):

Let $A \in \mathbb{R}^{m \times n}$, $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_m^T \end{pmatrix}$, $a_i \in \mathbb{R}^n$. Let $B \in \mathbb{R}^{p \times n}$ and $c \in \mathbb{R}^n$. Then the following two statements are equivalent:

1. For every $s \in \mathbb{R}^n$ with $As \leq 0$ and $Bs = 0$ holds $c^T s \leq 0$.
2. $\exists \lambda \in \mathbb{R}^m, \lambda \geq 0, \exists \mu \in \mathbb{R}^p$ with $c = A^T \lambda + B^T \mu$.

$(P_{\text{canonical}}) : \min c^T x$ with $Ax \leq b$
$(P_{\text{standard}}) : \min c^T x$ with $Ax = b$ and $x \geq 0$

Karush-Kuhn-Tucker optimality conditions:
$\bar{x} \in \mathbb{R}^n$ is solution of the convex optimization problem

$$\min c^T x \text{ s.t. } Ax \leq b,\, Dx = e$$

$$\iff \exists \text{ Lagrange multipliers } \bar{\lambda} \in \mathbb{R}^m, \bar{\mu} \in \mathbb{R}^p \text{ such that } \begin{cases} c^T + A^T\bar{\lambda} + D^T\bar{\mu} = 0 \\ D\bar{x} = e \\ A\bar{x} \leq b,\, \bar{\lambda} \geq 0,\, \bar{\lambda}^T(A\bar{x} - b) = 0 \end{cases}$$

Definition: A triple $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ is called a KKT triple or KKT point if it fulfills the KKT system.

Theorem: $\bar{x}$ is a solution of the convex optimization problem

$$\min c^T x \text{ s.t. } Ax \leq b,\, Dx = e$$

$\Leftrightarrow \exists \bar{\lambda} \in \mathbb{R}^m$, $\bar{\mu} \in \mathbb{R}^p$ such that $(\bar{x}, \bar{\lambda}, \bar{\mu})$ is a KKT triple.

KKT-System for the canonical form ($P_{\text{canonical}}$):

$$(\bar{x}, \bar{\lambda}), \bar{\lambda} \in \mathbb{R}^n \text{ such that } \begin{cases} c + A^T\lambda = 0 \\ A\bar{x} \leq b,\, \bar{\lambda} \geq 0,\, \bar{\lambda}^T(A\bar{x} - b) = 0 \end{cases}$$

KKT-System for the standard form (P_{standard}):

$$\bar{\lambda} \in \mathbb{R}^n,\, \bar{\mu} \in \mathbb{R}^m \begin{cases} c - \bar{\lambda} + A^T\bar{\mu} = 0 \\ A\bar{x} = b \\ \bar{x} \geq 0,\, \bar{\lambda} \geq 0,\, \bar{\lambda}^T\bar{x} = 0 \end{cases}$$

For solution of $\min c^T x$ with $Ax \leq b$ we must require:

1. $X = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq \emptyset$
2. $\{c^T x \mid x \in X\} \subset \mathbb{R}$ has to be bounded from below.

Assume both holds. Then, $\exists f* \in \mathbb{R}$ such that $f* = \inf\{c^T x \mid x \in X\}$
$\bar{x}$ is a solution $\Leftrightarrow c^T\bar{x} = f* \implies$ if $X \neq \emptyset$ and bounded, then the existence is clear(since X turns out to be compact and $f(x) = c^T x$ is continuous guarantees existence of minimizer)

Week 9

Consider $\min c^T x$ such that $Ax \leq b$. Let $X = \{x \in \mathbb{R}^n \mid Ax \leq b\} \neq \emptyset$ and let $\{c^T x \mid x \in X\} \subset \mathbb{R}$ be bounded from below. Then, there exists a solution $\bar{x}$.

Duality
$(P) : \min\limits_{x\in\mathbb{R}^n} c^T x$ such that $Ax \le b$.
$(P*) : \max\limits_{\lambda\in\mathbb{R}^n} -b^T\lambda$ such that $A^T\lambda = -c, \lambda \ge 0$.

Definition: Lagrange functional for (P) is $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $L(x,\lambda) = c^T x + \lambda^T(Ax - b)$.

Lemma: $(P) \Leftrightarrow \inf\limits_{x} \sup\limits_{\lambda\ge 0} L(x,\lambda)$.

Lemma: $(P*) \Leftrightarrow \sup\limits_{\lambda\ge 0} \inf\limits_{x} L(x,\lambda)$.

Weak Duality Theorem: Let $x \in \mathbb{R}^n$ be admissible for (P) and $\lambda \in \mathbb{R}^n$ be admissible for $(P*)$. Then $c^T x \ge -b^T\lambda$.

Conclusions:

1. If (P) is NOT bounded from below $\implies$ $(P*)$ has no admissible points
2. If $(P*)$ is NOT bounded from above $\implies$ (P) has no admissible points
3. If (P) and $(P*)$ have admissible points $\implies$ for both problems solutions exist!
4. Let $\bar{x}$ be a solution of (P) and $\bar{\lambda}$ solution of $(P*)$ $\implies$ $c^T\bar{x} \ge -b^T\bar{\lambda}$

Corollary: Let $\bar{x} \in \mathbb{R}^n$ be admissible for (P), let $\bar{\lambda}$ be admissible for $(P*)$. Let, moreover, $c^T\bar{x} = -b^T\bar{\lambda}$. Then, $\bar{x}$ is a solution of (P) and $\bar{\lambda}$ is a solution of $(P*)$.

Strong Duality Theorem: Let $\bar{x}$ be a solution of (P) and $\bar{\lambda}$ be a solution of $(P*)$. Then:

1. $c^T\bar{x} = -b^T\bar{\lambda}$
2. $(\bar{x}, \bar{\lambda})$ fulfills KKT.

Goal: Simplex Method
$(P) \min c^T x$ s.t. $Ax \le b, Dx = e(A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, D \in \mathbb{R}^{p\times n}, e \in \mathbb{R}^p)$. Admissible set for (P) is called (convex) polytope.

Definition:
Point $x \in X$ is called vertex of polytope X if: $\operatorname{rank}\begin{pmatrix} A_{\mathcal{A}(x)} \\ D \end{pmatrix} = n$

Definition:
A vertex $x \in X$ is called *regular* if $\begin{pmatrix} A_{\mathcal{A}(x)} \\ D \end{pmatrix}$ has n rows$(\#\mathcal{A}(x) + p = n)$. (in this case $\begin{pmatrix} A_{\mathcal{A}(x)} \\ D \end{pmatrix} \in \mathbb{R}^{n\times n}$ is non-singular).

Definition: A vertex is called degenerate if $\#\mathcal{A}(x) + p > n$.

Definition: Let $F \subset X$. $\mathcal{A}(F) = \bigcap_{x \in F} \mathcal{A}(x) = \{j \mid a_j^T x = b_j, \forall x \in F\}$

Definition: A set $F \subset X$ is called an edge of X, if $\emptyset \neq F = \{x \in X \mid A_{\mathcal{A}(F)} x = b_{\mathcal{A}(F)}, Dx = e\}$ and rank $\begin{pmatrix} A_{\mathcal{A}(F)} \\ D \end{pmatrix} = n - 1$.

Definition: Vertices x, y of X are neighbors, if rank $\begin{pmatrix} A_{\mathcal{A}(x) \cap \mathcal{A}(y)} \\ D \end{pmatrix} = n - 1$.

Week 10

(P): $\min_{x \in \mathbb{R}^n} c^T x$ such that $Ax \leq b$, $Dx = e$

Theorem: Let $c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, D \in \mathbb{R}^{p \times n}, e \in \mathbb{R}^p$.

- Let $X = \{x \in \mathbb{R}^n \mid Ax \leq b, Dx = e\} \neq \emptyset$
- Let $f* = \inf\{c^T x \mid x \in X\} \in \mathbb{R}$
- Let rank $\begin{pmatrix} A \\ D \end{pmatrix} = n$

Then, there exists a vertex $\bar{x} \in X$ with $c^T \bar{x} \leq c^T x$, $\forall x \in X$.

Corollary 1 $\exists$ an optimal solution of (P), which is a vertex.
Corollary 2 If $X \neq \emptyset$ and rank $\begin{pmatrix} A \\ D \end{pmatrix} = n \implies \exists$ vertex.

Lemma: Let $\bar{x} \in X$ be a vertex, which is NOT optimal. The, there is an edge going out from $\bar{x}$ such that the cost functional decreases along this edge $K = \{\bar{x} + ts \mid 0 \leq t \leq t_+\} \subset X$, $t_+ > 0, s \in \mathbb{R}^n$ and $c^T s < 0$.

Week 11

Lemma: Let $\bar{x}$ be a vertex and $s \in \mathbb{R}^n \setminus \{0\}$ be a direction of an edge going out from $\bar{x}$: $K = \{\bar{x} + ts \mid 0 \leq t \leq t_+\} \subset X$, $t_+ > 0$. Then, there is a subset $\mathcal{A} \subset \mathcal{A}(x_k)$ with $|\mathcal{A}| = n$, $\text{rank}(A_\mathcal{A}) = n$ and $\exists j \in \mathcal{A}$, $\tau > 0$ such that $A_\mathcal{A} \cdot s = -\tau (e_j)_\mathcal{A}$

Instead of solving $\mathcal{A}_{\mathcal{A}_k} s_{k_j} = -(e_j)_{\mathcal{A}_k}$ for every $j \in \mathcal{A}_k$, solve $A_{\mathcal{A}_k}^T (\lambda_k)_{\mathcal{A}_k} = -c \implies c^T s_{k_j} = (-A_{\mathcal{A}_k}^T (\lambda_k)_{\mathcal{A}_k})^T s_{k_j} = (\lambda_k)_{\mathcal{A}_k}^T (e_j)_{\mathcal{A}_k} = (\lambda_k)_j$

if $(\lambda_k)_j < 0 \implies s_{k_j}$ is potentially a decreasing edge.
if $(\lambda_k)_j \geq 0 \implies s_{k_j}$ can not be a decreasing edge.

Dual Simplex Algorithm:

Given $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n$, minimize $c^T x$ s.t. $Ax \leq b$ (This is P). Assume rank$(A) = n$, a.k.a. existence of vertices.

0. Find one vertex $x_0 \in X$ and choose a working index set $\mathcal{A}_0 \subset \mathcal{A}(x_0)$ with $|\mathcal{A}_0| = n$ and rank$(A_{\mathcal{A}_0}) = n$.

Iterate for $k = 0, 1, 2, \cdots$

1. Compute $(\lambda_k)_{\mathcal{A}_k} \in \mathbb{R}^n$ as the solution of $A^T_{\mathcal{A}_k}(\lambda_k)_{\mathcal{A}_k} = -c$
2. if $(\lambda_k)_{\mathcal{A}_k} \geq 0 \implies$ STOP with x_k being an optimal solution!
3. Choose $j_k \in \mathcal{A}_k$ with $(\lambda_k)_{j_k} < 0$. Compute $s_k \in \mathbb{R}^n$ as solution to $A_{\mathcal{A}_k} s_k = -(e)_{j_k}$. If $A_{\mathcal{A}(x_k)} s_k \leq 0$, then go to step 5.
4. (We are here if s_k is not an edge) Change the working index set(staying in the same vertex). Choose $i_k \in \mathcal{A}(x_k) \backslash \mathcal{A}_k$ such that rank$(A_{(\mathcal{A}_k \setminus \{j_k\}) \cup \{i_k\}}) = n$. Set $x_{k+1} = x_k, \mathcal{A}_{k+1} = (\mathcal{A}_k \setminus \{j_k\}) \cup \{i_k\}$ (remove j_k from the working index set $\mathcal{A}_k$ and add i_k to it). Go to step 1(next iteration).
5. (We are here if s_k is a decreasing edge) If $As_k \leq 0 \implies$ STOP: The problem is not bounded below, so no solution exists.
6. (We are here if s_k is a decreasing edge and $As_k \not\leq 0$) We compute step size $\sigma_k > 0$ with $\sigma_k = \min\limits_{i: a_i^T s_k > 0} \frac{b_i - a_i^T x_k}{a_i^T s_k}$. Let i_k be argmin of this min.
7. $x_{k+1} = x_k + \sigma_k s_k$(new vertex), $\mathcal{A}_{k+1} = (\mathcal{A}_k \setminus \{j_k\}) \cup \{i_k\}$. Go to step 1.

Remark:

This version of the algorithm could potentially contain (infinite) cycles.

Convergence Theorem of the Dual Simplex Method:

Assume X has vertices and the above algorithm does not produce cycles. Then the algorithm stops after finitely many cycles with either of the following:

(a) x_k is an optimal vertex

(b) the problem is not bounded from below.

Week 12

Blend's rule to avoid cycles:

1. In step 3, we choose $j_k \in \mathcal{A}_k$ with $(\lambda_k)_{j_k} < 0$ as the smallest index with this property.
2. In step 4, we choose $i_k \in \mathcal{A}(x_k) \setminus \mathcal{A}_k$ with $a_{i_k}^T s_k > 0$ as the smallest index with this property.

Theorem: If we use Blend's rule then cycles are than avoided.

How to obtain a first vertex?

$A \in \mathbb{R}^{m\times n}$, $m \geq n, \operatorname{rank}(A) = n, b \in \mathbb{R}^m$, $X = \{x \in \mathbb{R}^n \mid Ax \leq b\}$.

Assume, we know $x_0 \in X \to$ admissible point $\to$ find a vertex!

Algorithm given x_0:

Iterate $k = 0, 1, 2, \cdots$
0. If $\operatorname{rank}(A_{\mathcal{A}(x_k)}) = n \implies$ STOP x_k is a vertex.

1. Compute $w_k \in \mathbb{R}^n \setminus \{0\} : A_{\mathcal{A}(x_k)} w_k = 0$(i.e. $w_k \in \ker\left(A_{\mathcal{A}(x_k)}\right)$)
2. If $c^T w_k = 0$, we choose $v_k \in \{\pm w_k\}$ such that $a_i^T v_k > 0$ for at least one i.
3. If $c^T w_k \neq 0$, then we choose $v_k \in \{\pm w_k\}$ such that $c^T v_k < 0$
   - If $a_j^T v_k \leq 0$ for all $j \in I(x_k) \implies$ STOP: the problem is not bounded from below $\implies$ no solution.
4. Compute $t_k = \min\left\{\frac{b_i - a_i^T x_k}{a_i^T v_k} \mid i \in I(x_k),\, a_i^T v_k > 0\right\}$, $x_{k+1} = x_k + t_k v_k$, go to the next iteration.

How to find an admissible point?

Consider the following linear problem $(\tilde{P})$ subject to $a_i^T x - t \leq b_i, t \geq 0$

$$\min_{x\in\mathbb{R}^n,\, t\in\mathbb{R},\, \begin{pmatrix} x \\ t \end{pmatrix} \in \mathbb{R}^{n+1}} t$$

Let $x_0 \in \mathbb{R}^n$ be arbitrary. Choose $t_0 = \max(0, \max\limits_{1\leq i\leq m}(a_i^T x_0 - b_i)) \implies t_0 \geq 0, t_0 \geq a_i^T x_0 - b_i, \forall i \implies (x_0, t_0)$ is an admissible point for $(\tilde{P}) \implies$ Find a vertex for $(\tilde{P}) \implies$ We solve $(\tilde{P})$ by the simplex method $\implies \begin{pmatrix} \bar{x} \\ \bar{t} \end{pmatrix} \in \mathbb{R}^{n+1}$ is a solution of $(\tilde{P})$. If $\bar{t} > 0$, then (P) has no admissible set. If $\bar{t} = 0$, $\bar{x} \in X$.

$$\tilde{A} = \begin{pmatrix} a_1^T & -1 \\ \vdots & \\ a_m^T & -1 \\ 0\cdots 0 & -1 \end{pmatrix} \text{ and } \tilde{b} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

Week 13

(P) $\min c^T x$ s.t. $a_i^T x \leq b_i$, $1 \leq i \leq m-p$ and $a_i^T x = b_i$, $m-p+1 \leq i \leq m$.

Set $G = \{m-p+1, \cdots, m\}$. Require $G \subset \mathcal{A}_k$ for all the iterations k.
We compute our $(\lambda_k)_{\mathcal{A}_k}$ in step 3; We choose $j_k \in \mathcal{A}_k \setminus G$ with $(\lambda_k)_{j_k} < 0$.
STOPPING Criterium: $(\lambda_k)_{\mathcal{A}_k \setminus G} \geq 0$.

Primal Simplex Method(in standard form):
(P) $\min c^T x$, such that $Ax = b, x \geq 0, A \in \mathbb{R}^{m \times n} : m < n$.

x is a vertex $\Leftrightarrow$ rank $\begin{pmatrix} -I_{\{x_i=0\}} \\ A \end{pmatrix} = n$

Notation: Row view of matrix $A_J = \begin{pmatrix} a_{j_1}^T \\ \vdots \\ a_{j_m}^T \end{pmatrix} \in \mathbb{R}^{\#J \times n}$, $a_i \in \mathbb{R}^n$

Column view of matrix $A_{\bullet,J} = \begin{pmatrix} a_{j_1} & a_{j_2} & \cdots & a_{j_k} \end{pmatrix} \in \mathbb{R}^{m \times \#J}$

Equivalent Characterization of a Vertex:
The point $\bar{x} \in X = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is a vertex of $X \Leftrightarrow A_{\bullet,\{\bar{x}_i > 0\}}$ has full column rank.

Definition:
A vector $\bar{x} \in X = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is called **admissible basis solution**, if there is a set $\mathcal{B} \subset \{1, 2, \ldots, n\}$ with

- $\#\mathcal{B} = m = \text{rank}(A_{\bullet,\mathcal{B}})$
- $\bar{x}_i = 0, \forall i \notin \mathcal{B}$

In this case, the index set $\mathcal{B}$ is called a basis index set, $A_{\bullet,\mathcal{B}}$ is called basis matrix.

Lemma:
Let $\text{rank}(A) = m$. Then, $\bar{x}$ is a vertex of $X \Leftrightarrow \bar{x}$ is an admissible basis solution.

Correspondence of Dual and Primal Simplex Algorithms:

$x \geq 0 \Leftrightarrow -Ix \leq 0$, So $x \geq 0 \wedge Ax \leq b$ can be rewritten as $\hat{A} = \begin{pmatrix} -I \\ A \end{pmatrix} \in \mathbb{R}^{(m+n)\times n}$.

In the logic of Dual Simplex method:

- inequality constraints $1, 2, \cdots, n$
- equality constraint $n+1, n+2, \cdots, n+m$(always active)

In iteration k of Dual Simplex, working index set $\hat{\mathcal{A}}_k$ with $\#\hat{\mathcal{A}}_k = n$, such that $\{n+1, n+2, \ldots, n+m\} \subset \hat{\mathcal{A}}_k$ and $\text{rank}(A_{\hat{\mathcal{A}}_k}) = n$.

We can uniquely describe $\hat{\mathcal{A}}_k$ by identification of indices i from $\{1, 2, \ldots, n\}$ which are not in $\hat{\mathcal{A}}_k$: we set $\mathcal{B}_k = \{1 \leq i \leq n \mid i \notin \hat{\mathcal{A}}_k\} = \{1, 2, \ldots, n\} \setminus \hat{\mathcal{A}}_k$ and $\mathcal{N}_k = \{1, 2, \ldots, n\} \setminus \mathcal{B}_k$.

Notation: $B_k = A_{\bullet, \mathcal{B}_k} \in \mathbb{R}^{m\times m}$ and $N_k = A_{\bullet, \mathcal{N}_k} \in \mathbb{R}^{m\times(n-m)}$

Lemma: In this notation the following statements are equivalent:

1. x_k is a vertex of X and $\hat{\mathcal{A}}_k = \{1, \ldots, n+m\} \setminus \mathcal{B}_k$ is a working index set.
2. x_k is an admissible solution with $\mathcal{B}_k$ being a basis set.

Choice of edge: $B_k s_{k_{j,\mathcal{B}_k}} = -a_j$ with $B_k = A_{\bullet, \mathcal{B}_k} \in \mathbb{R}^{m\times m}$.

We are only interested in edges, which are descent directions: $c^T s_{k_j} < 0$.

$$\begin{cases} -\lambda_{k,\mathcal{N}_k} + N_k^T \mu_k = -c_{\mathcal{N}_k} \\ B_k^T \mu_k = -c_{\mathcal{B}_k} \text{ (we need to solve this)} \end{cases}$$

Then $\lambda_{k,\mathcal{N}_k} = N_k^T \mu_k + c_{\mathcal{N}_k}$. If $\lambda_{k,\mathcal{N}_k} \geq 0$, then x_k is an optimal solution, so STOP.

If not, For every $j \in \mathcal{N}_k$ with $(\lambda_k)_j < 0$ the corresponding s_{k_j} would be a descent direction.

Week 14

Primal Simplex method (P):

$\min c^T x$ such that $Ax = b, x \geq 0, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ and $m < n$ $(\text{rank}(A) = m)$, $X = \{x \in \mathbb{R}^n \mid Ax - b,\, x \geq 0\}$.

Algorithm:

0. $x_0 \in X$ is basis solution with basis set $\mathcal{B}_0$. We set $\mathcal{N}_0 = \{1, 2, \cdots, n\} \setminus \mathcal{B}_0$. Iterate $k = 0, 1, \cdots$

1. Compute $\mu_k \in \mathbb{R}^m, \lambda_{k_{\mathcal{N}_k}} \in \mathbb{R}^{n-m}$ by $B_k^T \mu_k = -c_{\mathcal{B}_k}, \lambda_{k_{\mathcal{N}_k}} = c_{\mathcal{N}_k} + N_k^T \mu_k$
2. if $\lambda_{k_{\mathcal{N}_k}} \geq 0 \implies$ STOP, x_k is optimal (no decreasing edges)
3. Choose $j_k \in \mathcal{N}_k$ with $\lambda_{j_k} < 0$
4. Compute $s_{k_{\mathcal{B}_k}}$ by solving $B_k s_{k_{\mathcal{B}_k}} = -a_{j_k}$ ($m \times m$ system)
5. if $s_{k_{\mathcal{B}_k}} \geq 0 \implies$ STOP, the problem has no solution(not bounded from below).
6. $\sigma_k = \min\limits_{i \in \mathcal{B}_k \wedge s_{k_i} < 0} \left(-\frac{x_{k_i}}{s_{k_i}}\right)$ and i_k is the argmin.
7.

$$x_{k+1,\mathcal{B}_k} = x_{k,\mathcal{B}_k} + \sigma_k s_{\mathcal{B}_k} \tag{1}$$

$$x_{k+1,j_k} = \sigma_k \tag{2}$$

$$x_{k+1,\mathcal{N}_k \setminus \{j_k\}} = 0 \tag{3}$$

$$\mathcal{B}_{k+1} = (\mathcal{B}_k \setminus \{i_k\}) \cup \{j_k\} \tag{4}$$

$$\mathcal{N}_{k+1} = (\mathcal{N}_k \setminus \{j_k\}) \cup \{i_k\} \tag{5}$$

   Go to next iteration

Cycles are still possible, but avoidable with Blend's rules.

How to find x_0?

Assume $b \geq 0$. $\min\limits_{x \in \mathbb{R}^n, y \in \mathbb{R}^m} \ell^T y$ such that $Ax + y = b, x \geq 0, y \geq 0$ with $\ell = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^m$. We take $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$ and this is an admissible basis solution(vertex) of admissible set. We can start primal simplex method with $\begin{pmatrix} 0 \\ b \end{pmatrix}$. Since the problem is bounded from below ($y \geq 0 \implies \ell^T y \geq 0$), the algorithm will produce (after finitely many steps) a solution $\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix}$.

Worst case time complexity of simplex method is $O(2^n)$, when algorithm has to go through all vertices; However, average runtime is $O(n^{86})$ (this average complexity comes from "Beyond Hirsch Conjecture: Walks On Random Polytopes And Smoothed Complexity Of The Simplex Method" by Roman Vershynin).

Extreme Points(Sets) of Convex Sets

Definition:
Let $X \subset \mathbb{R}^n$, non-empty and convex. A point $\bar{x} \in X$ is called extreme point of X, if $x \in X \wedge x \neq x_1 + t(x_2 - x_1)$, $\forall x_1, x_2 \in X$ $(x_1 \neq x_2 \wedge t \in (0,1))$.

The set of all extreme points of X is called $\text{ext}(X)$.

Lemma: Let $X = \{x \mid Ax \leq b\}$, $x_1, \cdots, x_p \in X$, $t_1, \cdots, t_p$, $t_i > 0$ with $\sum\limits_{i=1}^{p} t_i = 1$. Let $x = \sum\limits_{i=1}^{p} t_i x_i$. Then, $\mathcal{A}(x) = \bigcap\limits_{i=1}^{p} \mathcal{A}(x_i)$

Theorem: $X = \{x \mid Ax \leq b\}$: $x \in \text{ext}(X) \Leftrightarrow x$ is a vertex of X.

Lemma: Let $\emptyset \neq C \in \mathbb{R}^n$ be convex. Then:

$$x \in \text{ext}(C) \Leftrightarrow \left(y, z \in C,\, x = \frac{1}{2}(y+z) \implies x = y = z\right)$$

Week 15

Lemma: Let $C \subset \mathbb{R}^n$ be non-empty and convex. Let $H = H(a,b)$ with $0 \neq a \in \mathbb{R}^n$, $b \in \mathbb{R}$ be a supporting hyperplane. Then:

$$\text{ext}(C \cap H) = \text{ext}(C) \cap H$$

Minkowski Theorem(special case of Krein-Milman Theorem):
Let $\emptyset \neq K \subset \mathbb{R}^n$ be compact and convex set. Then, $K = \text{conv}(\text{ext}(K))$.
Especially for $K = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ (compact and non-empty), K is a convex hull of the set of its vertices.

Maximum Principle:
Let $\emptyset \neq K \subset \mathbb{R}^n$ be convex and compact. Let $f : K \to \mathbb{R}$ be a convex function. If f attains a maximum on K, then it attains this maximum also in an extreme point of K.

$$f(\bar{x}) = \max_{x \in K} f(x) \implies \exists \tilde{x} \in \text{ext}(K) : f(\tilde{x}) = f(\bar{x}) = \max_{x \in K} f(x)$$

Corollary:

Let $\emptyset \neq K \subset \mathbb{R}^n$ be convex and compact. Let $f : K \to \mathbb{R}$ be convex and continuous. Then, there exists $\bar{x} \in \text{ext}(K)$ s.t. $f(\bar{x}) = \max_{x \in K} f(x)$

Corollary:

Let $f(x) = c^T x$, $c \in \mathbb{R}$ and $K \neq \emptyset$ convex and compact. Then, $\max c^T x$ and $\min c^T x$ with $x \in K$ have solutions in $\text{ext}(K)$.

Theorem:

Let $\emptyset \neq K$ convex and compact. Let $A \subset K$. Then the following two statements are equivalent:

1. $\text{ext}(K) \subset A$
2. $K = \text{conv}(A)$