Chapter 1
Linear Programming

Paragraph 3
Mathematical Foundations - Geometry of the Solution Set

What we did so far

- By combining ideas of a specialized algorithm with a geometrical view on the problem, we developed an algorithm idea:

> Find a feasible corner (somehow).
> Check neighboring corners and see if one is better.
> Move over to the next corner until no better neighboring solution exists.

- We have an intuitive understanding how our geometrical view generalizes to more dimensions:
  - Corners correspond to solutions of equation systems.
  - Inequalities partition restrict the solution space to halfspaces.

Analysis

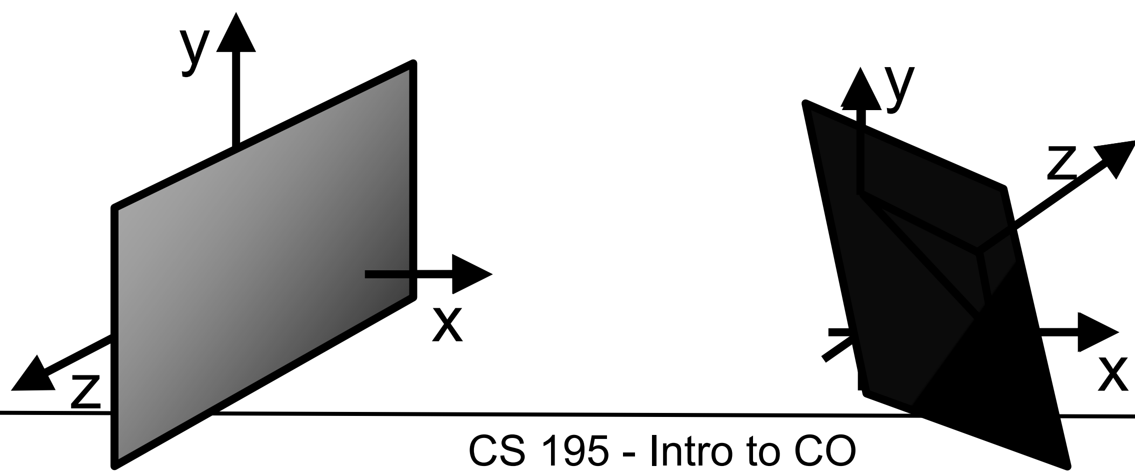
- Algorithm Idea

> Find a feasible corner (somehow).
> Check neighboring corners and see if one is better.
> Move over to the next corner until no better neighboring solution exists.

- Open Questions
  – Can we find a mathematical formalization of linear optimization problems for which we can define what "corner" and "neighboring corner" means?
  – Can we prove optimality?
  – How do we find a feasible starting solution?

What is a corner?

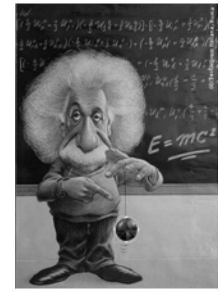
- A corner in our geometrical view is defined by the intersection of two lines. And a line is defined by an equation ⇨ a corner is a solution to an equation-system!
- What if there are more than two variables? How does an inequality look like then?
  – Given x,y,z, how does $x \geq 1$ look like?
  – Given x,y,z, how does $x+y+z = 1$ look like?



Equation Systems

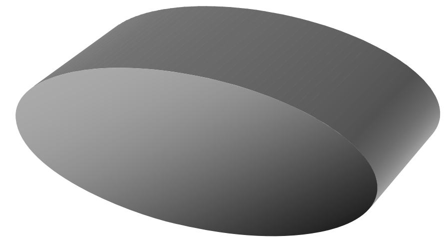
- So an equation in an n-dimensional space defines an n-1-dimensional hyperplane!
  - n = 2: equations define lines
  - n = 3: equations define planes
- Every inequality divides the space in two halfspaces!
- A corner in an n-dimensional space is defined by the intersection of n hyperplanes. Therefore, a corner defines a solution to an equation system and vice versa.

Mathematical Foundations



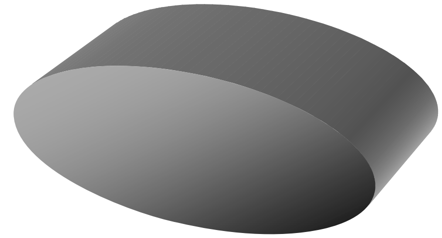
- We shall now define and study formally
  - what corners are,
  - how they correspond to basic solutions of equation systems,
  - what neighboring corners are, and
  - how optimal solutions to LPs are characterized.

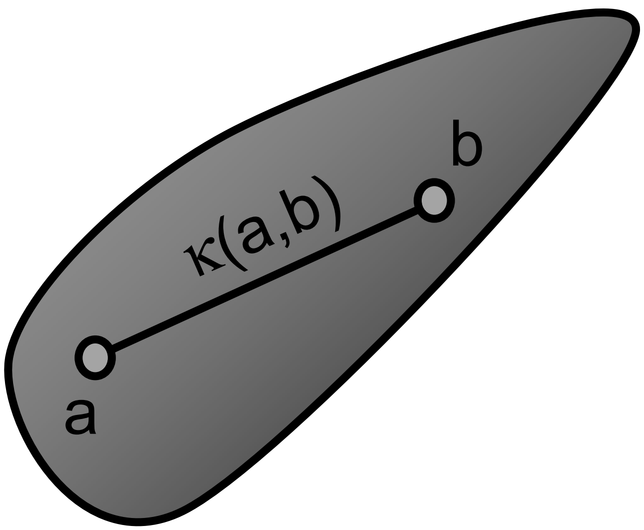
Convex Sets



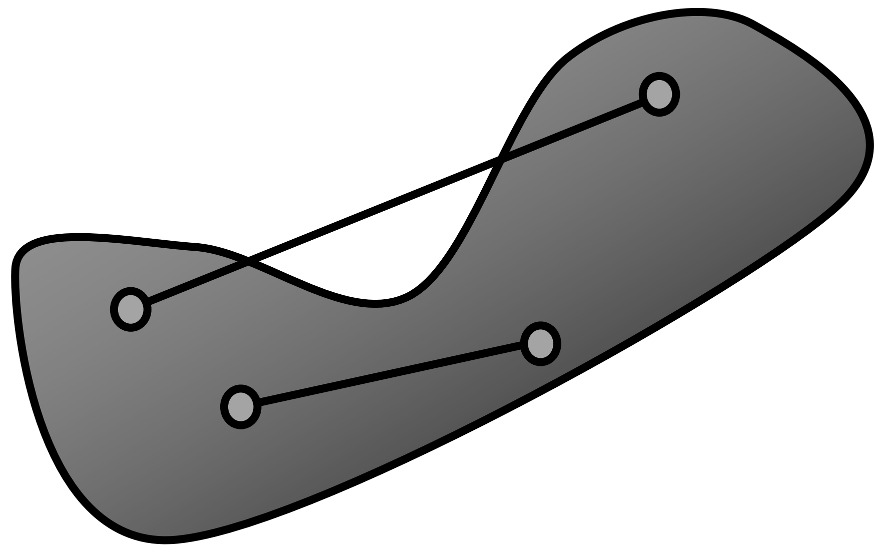
- Definition
  - Assume that $\{a^1,\ldots,a^m\} \subseteq \mathbb{R}^n$ and $\alpha^1,\ldots,\alpha^m \in \mathbb{R}_{\geq 0}$.
  - $\Sigma_i\ \alpha^i\ a^i$ is called a non-negative linear combination.
  - In case that $\Sigma_i\ \alpha^i = 1$, we call $\Sigma_i\ \alpha^i a^i$ a convex combination. It is called true convex combination iff $\alpha^1,\ldots,\alpha^m \in \mathbb{R}_{>0}$.
  - We define $\kappa(a^1,\ldots,a^m)$ as the set of all convex combinations of $a^1,\ldots,a^m$.
  - For $a,b \in \mathbb{R}^n$, $\kappa(a,b)$ is called the line between a and b.
  - A set $K \subseteq \mathbb{R}^n$ is called convex, iff for all $a,b \in K$: $\kappa(a,b) \subseteq K$.

Convex Sets



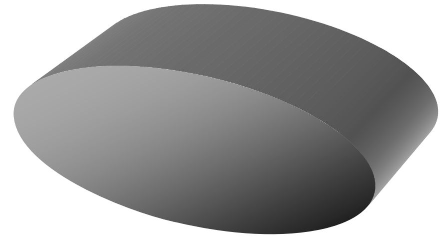


Convex



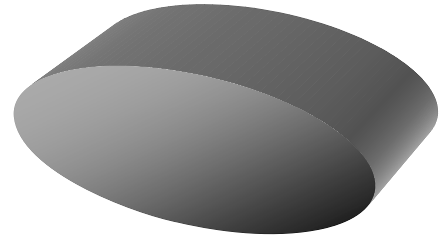
Not Convex

Convex Sets



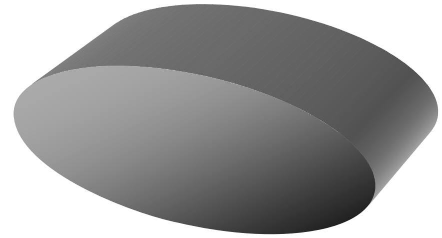
- Remark
  - If $K \subseteq \mathbb{R}^n$ is convex and $a^1,\ldots,a^m \in K$, then $\kappa(a^1,\ldots,a^m) \subseteq K$.
  - The intersection of convex sets is convex.
- Proof:

Convex Sets



- Examples
  - Let $a \in \mathbb{R}^n$, $a \neq 0$, and $\alpha \in \mathbb{R}$. $H = \{x \in \mathbb{R}^n \mid a^T x = \alpha\}$ is called a hyperplane and is convex.
  - Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. $V = \{x \in \mathbb{R}^n \mid Ax = b\}$ is called affine vector space and is convex.
  - Let $a \in \mathbb{R}^n$, $a \neq 0$, and $\alpha \in \mathbb{R}$. $H^{\geq} = \{x \in \mathbb{R}^n \mid a^T x \geq \alpha\}$ is called halfspace and is convex.
  - $P = \{x \in \mathbb{R}^n \mid Ax = b \text{ and } x \geq 0\}$ is convex.
  - $D = \{x \in \mathbb{R}^n \mid ||x||_2 \leq 1\}$ is convex.

Convex Sets



- Definition
  - For $M \subseteq \mathbb{R}^n$, we define $\kappa(M) := \bigcup_{k \in \mathbb{N},\, a1,\ldots,ak \in M} \kappa(a^1,\ldots,a^k)$.
  - $\kappa(M)$ is called the convex hull of M.
- Remark
  - It holds that $\kappa(M)$ equals the intersection of all convex sets that contain M. Thus, $\kappa(M)$ is the smallest convex set that contains M.

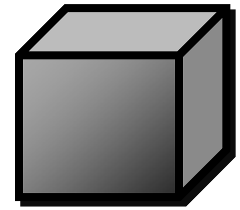
Extreme Points

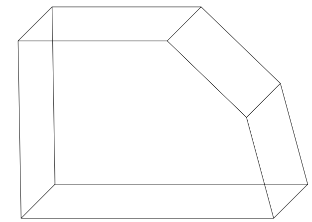
- Definition
  - A point $x \in K$, $K \subseteq \mathbb{R}^n$ convex, is called extreme point, iff it cannot be represented as a true convex combination of two points in K. We denote the set of all extreme points of K with $\varepsilon(K)$.
- Remark: Given $K \subseteq \mathbb{R}^n$ convex, the following statements are equivalent:
  - x^0 is an extreme point of K.
  - For all $a,b \in K$ with $x^0 \in \kappa(a,b)$ it is $x^0=a$ or $x^0=b$.
  - For all $y \in \mathbb{R}^n$, $y \neq 0$ it is $x^0 + y \notin K$ or $x^0 - y \notin K$.
  - $K \setminus \{x^0\}$ is convex.

Extreme Points

- Examples
  - $\varepsilon(\{x \in \mathbb{R}^n \mid x \geq 0\}) = \{0\}$
  - $\varepsilon(H^{\geq}) = \emptyset$
  - $\varepsilon(D) = \{x \in \mathbb{R}^n \mid ||x||_2 = 1\}$
  - $\varepsilon(\{ x \mid ||x||_2 < 1\}) = \emptyset$
  - $\varepsilon(\{1\}) = \{1\}$

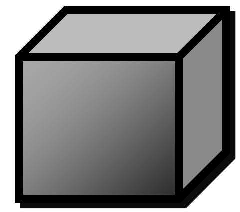
Convex Polyeders

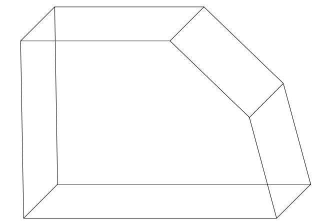




- Definition
  - A convex polyeder (convex polyhedron) P is defined as the finite intersection of halfspaces, i.e. $P = \cap H^{\geq} = \{x \mid Ax \geq b\}$.
  - A convex polyeder $P \neq \emptyset$ that is bounded is called (convex) polytope.
  - A hyperplane H is called supporting plane of P iff $H \cap P \neq \emptyset$ and $P \subseteq H^{\geq}$.
  - If $H \cap P = \{x^0\}$, then x^0 is called a corner of P.
  - If $H \cap P = \kappa(a,b)$ for some $a,b \in P$, then $\kappa(a,b)$ is called an edge of P.

Theorem

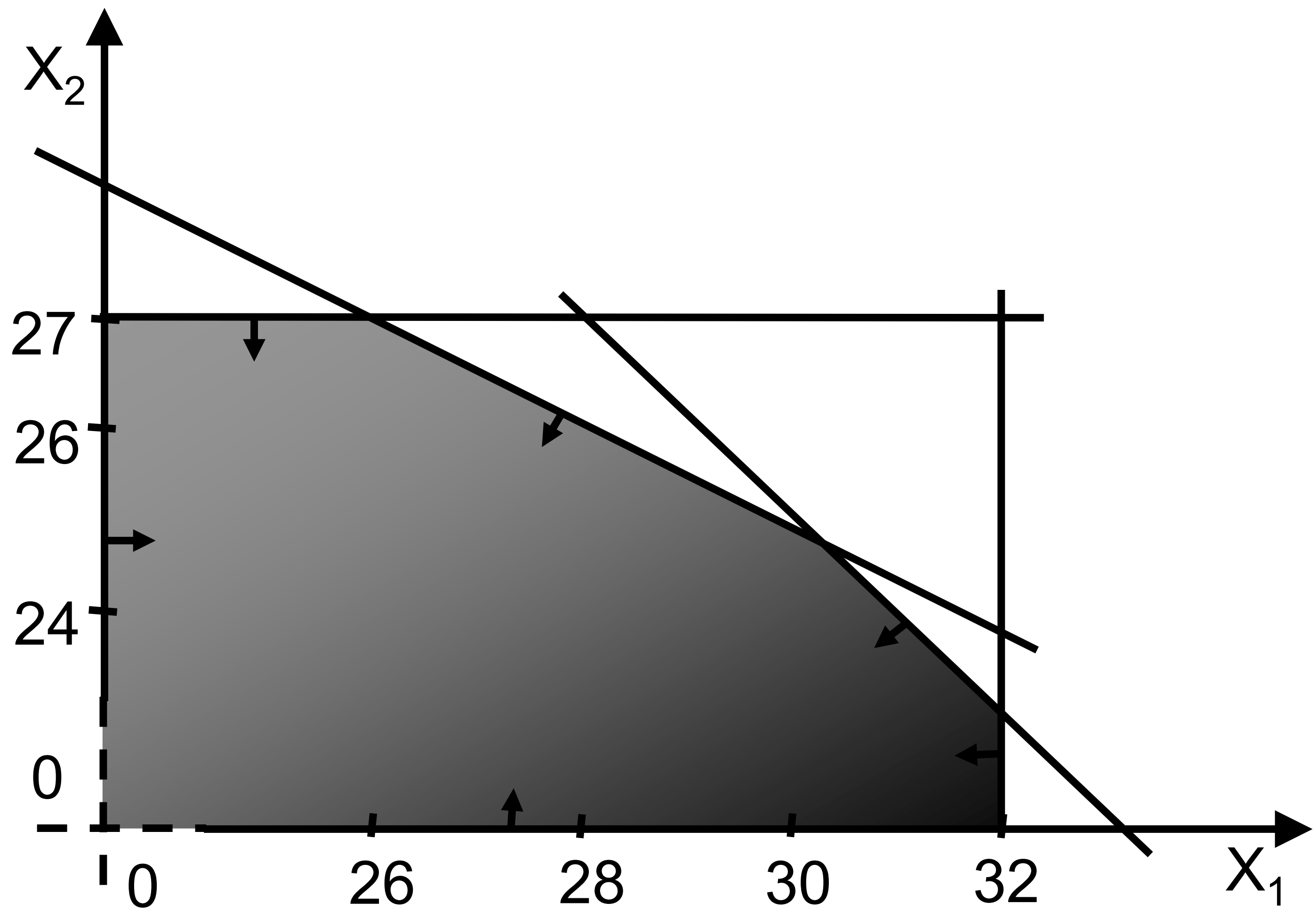




- Let P denote a convex polyeder.
  - Every corner of P is an extreme point.
  - P is the convex hull of its corners.

- Proof:

Basic Solutions and Corners



Basic Solutions

- Definition
  - Given $A \in \mathbb{R}^{m \times n}$ with rank(A) = m ≤ n, $b \in \mathbb{R}^m$, let $B : \mathbb{N}_m \rightarrow \mathbb{N}_n$, $N : \mathbb{N}_{n-m} \rightarrow \mathbb{N}_n$ injective such that $B(\mathbb{N}_m) \cup N(\mathbb{N}_{n-m}) = \mathbb{N}_n$ and $A_B = (a^{B(1)},\ldots,a^{B(m)})$ with $\text{rank}(A_B) = \text{rank}(A) = m$.
  - When setting $x_B := A_B^{-1} b$ and $x_N := 0$, we call $x^T := (x_B^T, x_N^T)$ basic (feasible) solution of $Ax = b$. x is called feasible, iff $x \geq 0$.

Basic Solutions

- Remark

$$\boxed{Ax = (A_B, A_N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} = A_B x_B + A_N x_N = A_B A_B^{-1} b + A_N 0 = b}$$

- Theorem
  - Let $A \in \mathbb{R}^{m \times n}$ with rank(A) = m ≤ n and $b \in \mathbb{R}^m$.
    For $x^0 \in P := \{x \in \mathbb{R}^n \mid Ax=b \text{ and } x \geq 0\}$ it is equivalent to say:
    - x^0 is extreme point of P.
    - $\{a_j \mid x^0_j > 0\}$ is linear independent.
    - x^0 is basic feasible solution
    - x^0 is a corner of P.

Basic Solutions

- Corollary
  - $0 \in P \Rightarrow 0 \in \varepsilon(P)$
  - Every corner has at most m entries that differ from 0!
  - S has at most $\boxed{\binom{n}{m}}$ corners!
- Definition
  - A corner is called degenerated iff $|\{ j \mid x_j > 0\}| < m$.
- Remark
  - If x^0 is not degenerated, the corresponding basis is uniquely defined!

Basic Solutions - Degeneracy

$X_1 + 2X_2 \le 80$

$X_1 + X_2 \le 55$

$X_1 \le 35$

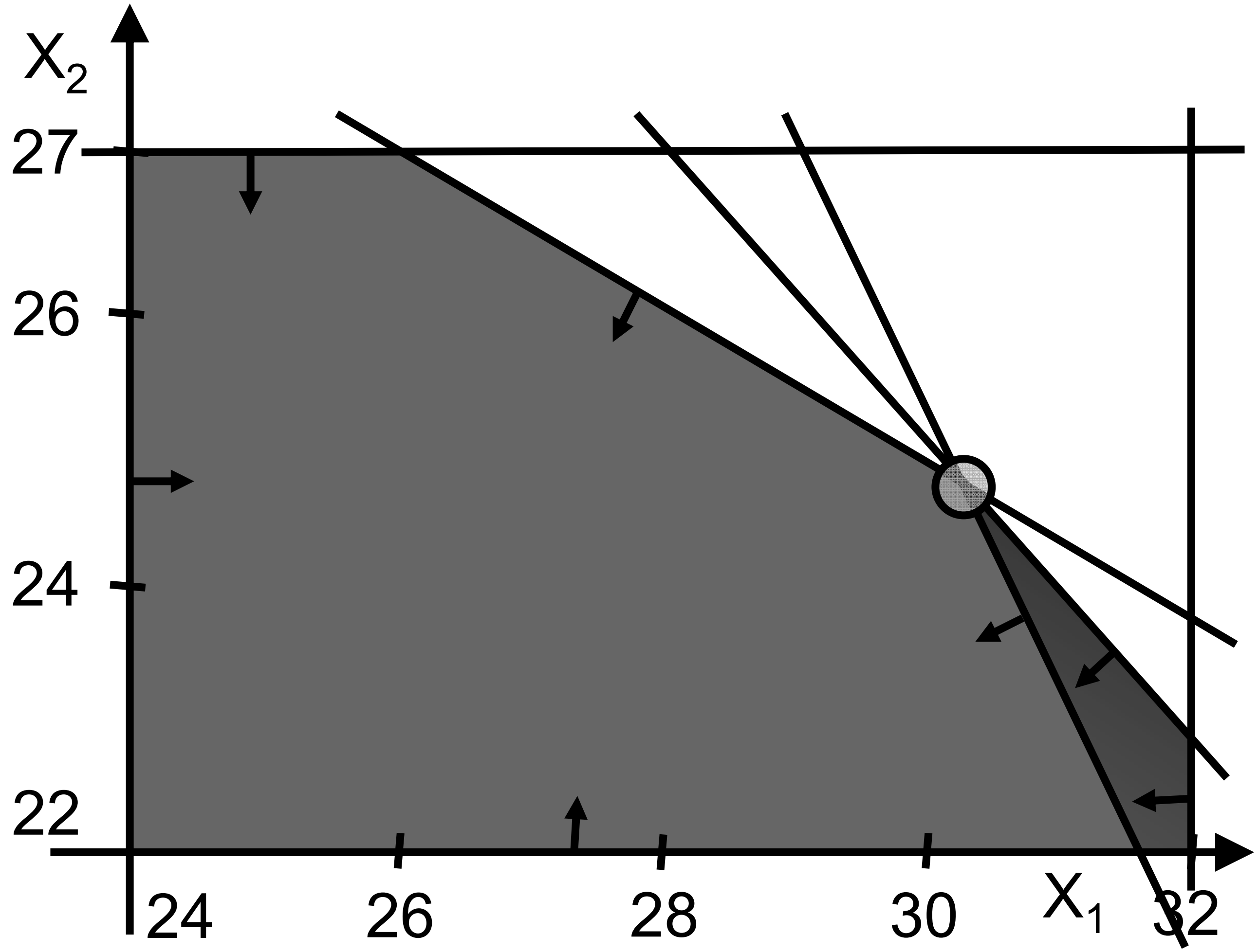
$X_2 \le 30$

$X_2 \le 27$

$2X_1 + X_2 \le 85$

$X_1 \le 32$

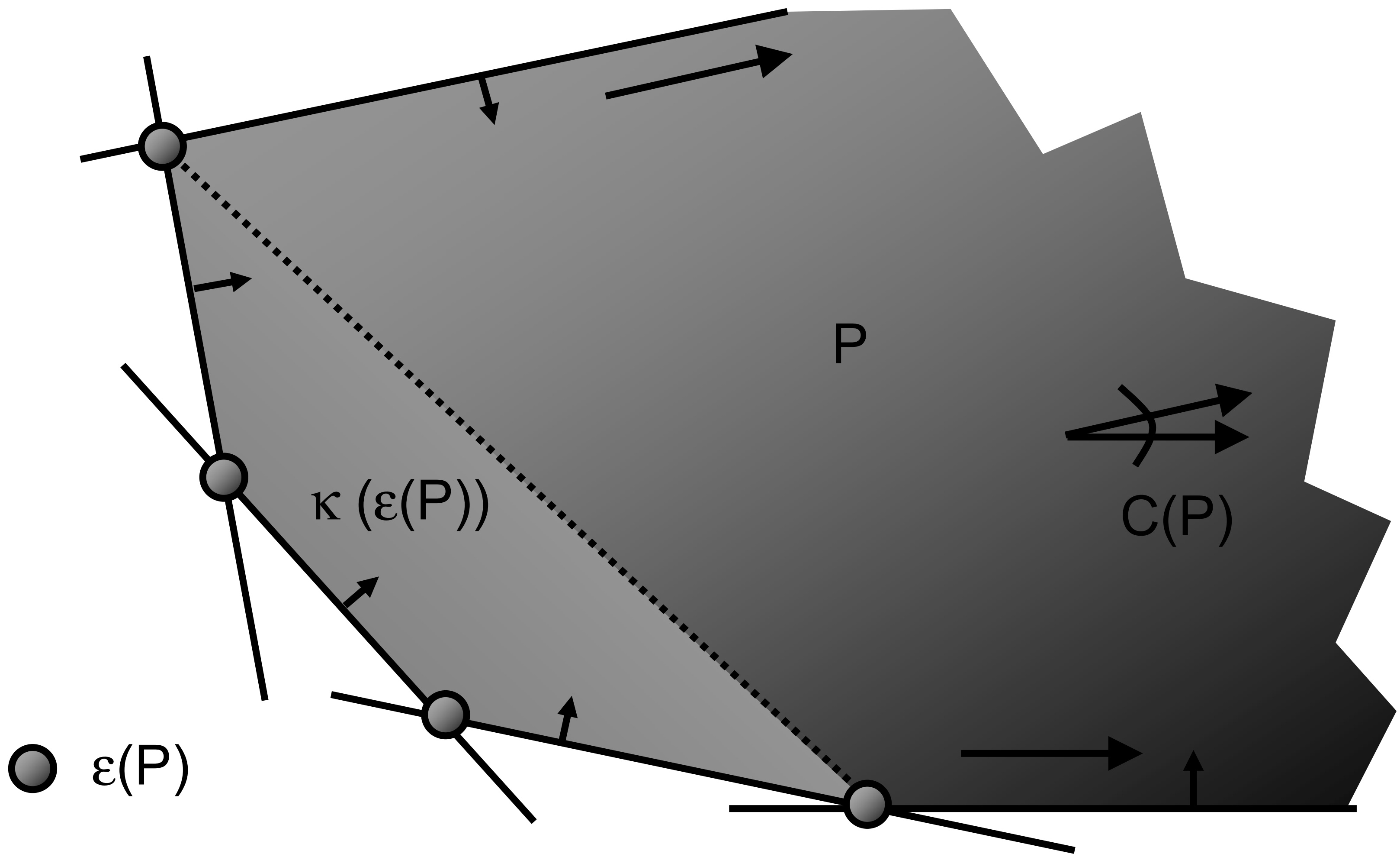
$X_1, X_2 \ge 0$



Optimal Solutions

- Definition
  - For $P \neq \emptyset$ we define $C(P) := \{ y \in \mathbb{R}^n \mid \forall\, x \in P, \lambda > 0 : x + \lambda y \in P\}$, and we say that $C(P)$ is the set of directions of P.
- Remark
  - $C(P) = \{ y \in \mathbb{R}^n \mid Ay = 0 \text{ and } y \geq 0\}$
- Theorem
  - $P = \kappa\,(\varepsilon(P)) + C(P)$

Optimal Solutions



Optimal Solutions

- Corollary
  - $P \neq \emptyset \Rightarrow$ P has corners!
  - If P contains an optimal solution, then there exists a corner with optimal objective value!
  - If $P \neq \emptyset$ and P has no optimal solution, then there exists $y \in C(P)$ such that $c^T y < 0$.
  - If $P \neq \emptyset$ and P is bounded, then $P = \kappa(\varepsilon(P))$.

Optimal Solutions

- Remark
  - The previous corollary yields an algorithm: Determine all basic solutions, eliminate all that are infeasible, and pick the one with the best objective function value.
  - What is the runtime of that algorithm?

Basis Changes

- Given $x^{0T} = (x^0{}_B{}^T, x^0{}_N{}^T)$ a basic feasible solution (i.e. $x^0{}_B = A_B^{-1} b$, $x^0{}_N = 0$, and $x^0 \geq 0$) and $x \in \mathbb{R}^n$ such that $Ax = b$, assume that for $y := x - x^0 \in \mathbb{R}^n$ it holds that $Ay = 0$.
- It holds $y_N = x_N$. Because of

$$\boxed{0 = Ay = A_B y_B + A_N y_N}$$

we have that

$$\boxed{y_B = -A_B^{-1} A_N x_N \quad \Rightarrow \quad y = \begin{pmatrix} -A_B^{-1} A_N x_N \\ x_N \end{pmatrix}}$$

Basis Changes

- Assume that we set specifically $x_N = e_k \in \mathbb{R}^{n-m}$, $t := N(k)$, and $\bar{A} := A_B^{-1} A$.
- Then, we have that $y_B = -\bar{a}^t := - A_B^{-1} a^t$. And therefore, $y^T = (-\bar{a}^{t\,T}_{(B)}, e_k^T{}_{(N)})$.
- For $\lambda \in \mathbb{R}$ and $x^\lambda := x^0 + \lambda y$, we thus have $Ax^\lambda = b$.
- $\boxed{0 \leq \lambda \leq \min \left\{ \frac{x^0_{B(j)}}{-y_{B(j)}} \mid y_{B(j)} < 0 \right\}}$ $\Rightarrow x^\lambda$ is feasible.

Basis Changes

- Theorem
  - If there exists $y_{B(j)} < 0$, we choose λ as large as possible ($\lambda < \infty$). Then, $x^1 := x^\lambda$ is a basic feasible solution. The corresponding basis is given by $B^*(i) := B(i)$ if $i \neq r$, and $B^*(r) := t$, whereby r such that $\lambda = x^0_{B(r)} / -y_{B(r)} = \min \{x^0_{B(i)} / -y_{B(j)} \mid y_{B(j)} < 0\}$.
  - If $y_B \geq 0$, then y (defined as before) is greater or equal 0, and thus: $x^\lambda = x + \lambda y$ is feasible for all $\lambda \geq 0$.
- Definition
  - A solution x^1 obtained by an exchange as discussed above with $0 < \lambda < \infty$ is called a neighboring corner of x^0.

Basis Changes

- Remark
  - If
    $$\boxed{\min\left\{\frac{x^0_{B(j)}}{-y_{B(j)}} \mid y_{B(j)} < 0\right\} = 0,}$$
    then $x^1 = x^0$ is degenerated.
  - If r in the previous theorem is not unique, then x^1 is degenerated.

Objective Function Change

- Assume that $\lambda < \infty$ and we found a neighboring corner x^1. Then:

$$z(x^0) := c^T x^0 = c_B^T x_B^0 + c_N^T x_N^0 = c_B^T x_B^0 =$$

$$= c_B^T A_B^{-1} b =: \pi^T b$$

$\Rightarrow$ For all $x \in P$, by setting $z^T := \pi^T A = c_B^T A_B^{-1} A$, we have

$$z^T x = \pi^T Ax = \pi^T b = z(x^0)$$

$$\Rightarrow z(x) = c^T x = z(x^0) - z^T x + c^T x = z(x^0) + (c^T - z^T)x =$$

$$= z(x^0) + (c_B^T - z_B^T)x_B + (c_N^T - z_N^T)x_N =$$

$$= z(x^0) + (c_N^T - z_N^T)x_N$$

$$\Rightarrow z(x^1) = z(x^0) + (c_t - z_t)\lambda$$

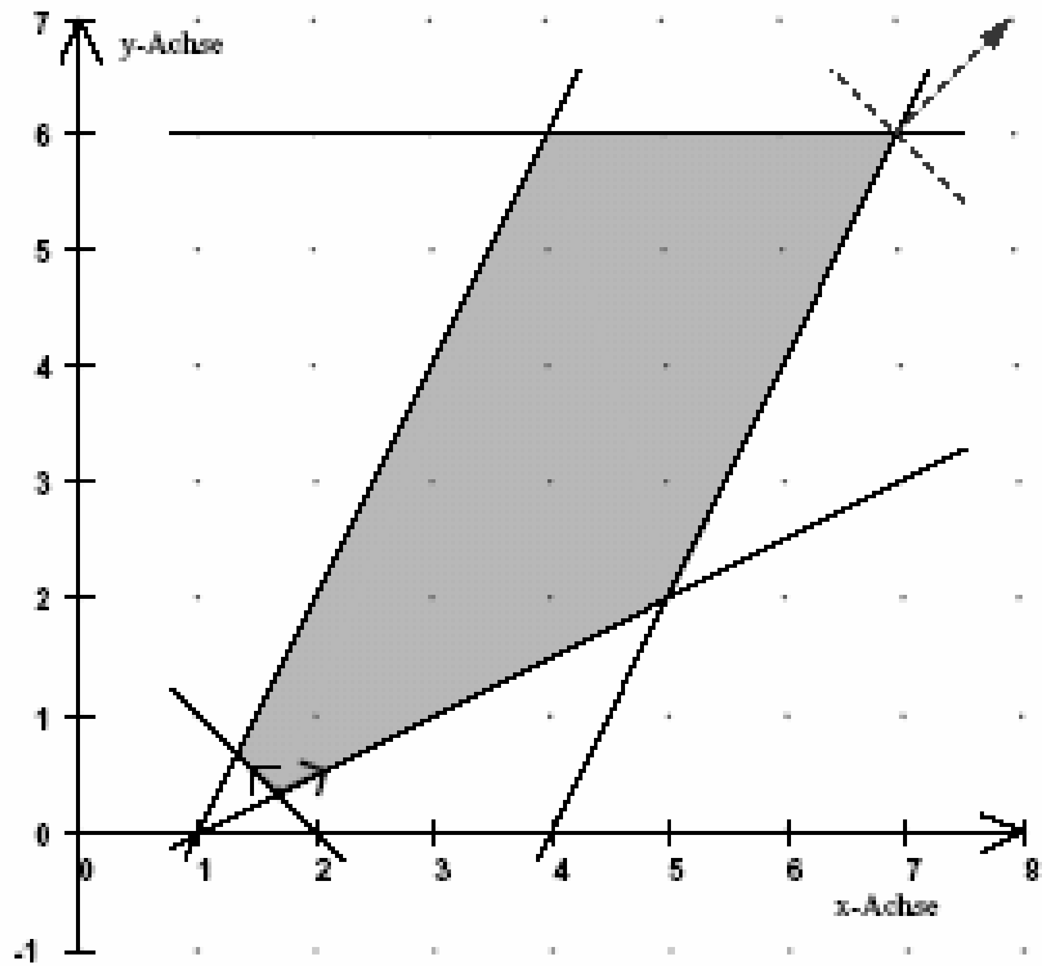
Objective Function Change

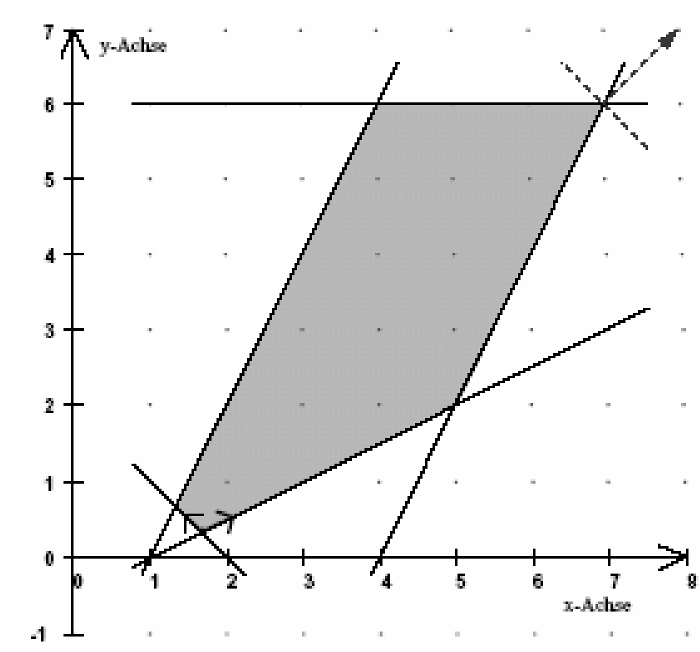
- Definition
  - We set $\boxed{\bar{c}^T := c^T - z^T = c^T - \pi^T A}$ and call the s'th component of this vector the relative costs of column a^s (with respect to B).

- Theorem
  - When changing the solution from x^0 to x^1, the costs change by $\boxed{\frac{x^0_{B(r)}}{-y_{B(r)}} \bar{c}_t}$.
  - If $c-z \geq 0$, then x^0 is optimal!

An Example

- Min - x - y
- 2x - 4y ≤ 2
- 2x - y ≤ 8
- -2x + y ≤ -2
- -2x - 2y ≤ -4
- y ≤ 6
- x,y ≥ 0



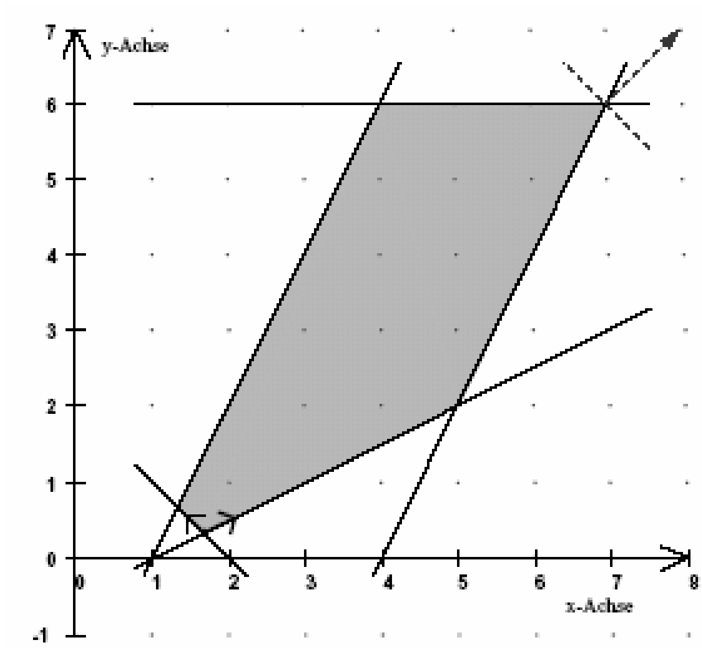


$$\begin{pmatrix} 1 & & & & & 2 & -4 \\ & 1 & & & & 2 & -1 \\ & & 1 & & & -2 & 1 \\ & & & 1 & & -2 & -2 \\ & & & & 1 & & 1 \end{pmatrix} \bullet \begin{pmatrix} s1 \\ s2 \\ s3 \\ s4 \\ s5 \\ a \\ b \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ -2 \\ -4 \\ 6 \end{pmatrix}$$

B = (6,2,3,7,5)   N = (1,4)

$$A_B = \begin{pmatrix} 2 & & & -4 & \\ 2 & 1 & & -1 & \\ -2 & & 1 & 1 & \\ -2 & & & -2 & \\ & & & 1 & 1 \end{pmatrix}$$

$$A_B^{-1} = \begin{pmatrix} \frac{1}{6} & & & -\frac{1}{3} & \\ -\frac{1}{2} & 1 & & \frac{1}{2} & \\ \frac{1}{2} & & 1 & -\frac{1}{2} & \\ -\frac{1}{6} & & & -\frac{1}{6} & \\ \frac{1}{6} & & & \frac{1}{6} & 1 \end{pmatrix}$$



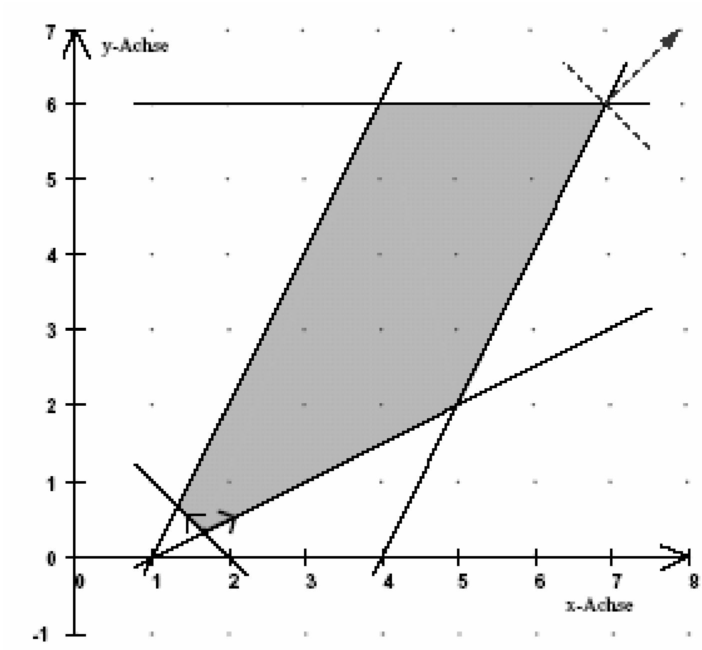
$$\begin{pmatrix} 1 & & & & & 2 & -4 \\ & 1 & & & & 2 & -1 \\ & & 1 & & & -2 & 1 \\ & & & 1 & & -2 & -2 \\ & & & & 1 & & 1 \end{pmatrix} \bullet \begin{pmatrix} s1 \\ s2 \\ s3 \\ s4 \\ s5 \\ a \\ b \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ -2 \\ -4 \\ 6 \end{pmatrix}$$

B = (6,2,3,7,5)   N = (1,4)

$$\boxed{x^{0^T} = (0,5,1,0,\tfrac{17}{3},\tfrac{5}{3},\tfrac{1}{3}) \Rightarrow Ax^0 = b}$$

$$\boxed{x^T = (0,0,6,10,4,5,2) \Rightarrow Ax = b}$$

$$\boxed{y^T = x^T - x^{0^T} = (0,-1,5,10,-\tfrac{5}{3},\tfrac{10}{3},\tfrac{5}{3}) \Rightarrow Ay = 0}$$



$$A_B^{-1} = \begin{pmatrix} \frac{1}{6} & & & -\frac{1}{3} & \\ -\frac{1}{2} & 1 & & \frac{1}{2} & \\ \frac{1}{2} & & 1 & -\frac{1}{2} & \\ -\frac{1}{6} & & & -\frac{1}{6} & \\ \frac{1}{6} & & & \frac{1}{6} & 1 \end{pmatrix}$$

B = (6,2,3,7,5)   N = (1,4)

$$x^{0\,T} = (0,5,1,0,\tfrac{17}{3},\tfrac{5}{3},\tfrac{1}{3})$$

$$x_N = (0,1) \quad \Rightarrow \quad (A_B^{-1} A_N x_N)^T = (-\tfrac{10}{3},5,-5,-\tfrac{5}{3},\tfrac{5}{3})$$

$$\Rightarrow \quad y = (0,-\tfrac{1}{2},\tfrac{1}{2},1,-\tfrac{1}{6},\tfrac{1}{3},\tfrac{1}{6})$$

$$\Rightarrow \quad \lambda = \min\{-\tfrac{5}{-\frac{1}{2}},-\tfrac{\frac{17}{3}}{-\frac{1}{6}}\} = \min\{10,34\} = 10$$

What have we achieved?

- The solution space of a linear programming problem is a convex polyeder. Corners of such a polyeder correspond to extreme points which correspond to basic feasible solutions.
- An optimal solution, if it exists, can always be found in a corner!
- Given a basic feasible solution, we can find a neighboring corner by exchanging exactly one basis column – which corresponds to following the direction given by a solution to the homogenous equation system!
- An improved basic feasible solution is found iff the corner is not degenerated and if the relative costs of the column that is introduced are negative. If no such column exists, the current solution is optimal!

Thank you!