Combinatorial Optimization

Winter course 2018/2019

Prof. Dr. Stefan Bock

University of Wuppertal
Business Computing and Operations Research

WINFOR

Information concerning the course

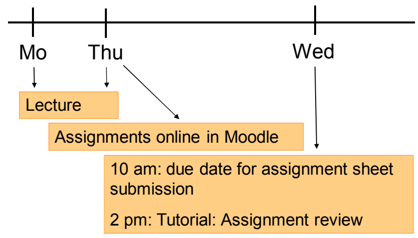
- Lecture:
  - Monday, 2:15 pm - 3:45 pm in M12.25
  - Thursday, 2:15 pm - 3:45 pm in M12.25
  - Start: October 11th, 2018
- Lecturer: Prof. Dr. Stefan Bock
  - Office: M12.02
  - Office hour: Monday, 4:00 pm - 6:00 pm (registration is mandatory (email to iwuester@winfor.de))
  - Secretary office: M12.01
  - E-mail: sbock@winfor.de

Information concerning the course

- Tutorial:
  - Wednesday, 2:00 pm - 4:00 pm in M.15.09
  - Start: October 17th, 2018
  - First assignment sheet is already available
- Supervisor: Anna Katharina Janiszczak
  - Office: M12.33
  - Office hour: Wednesday, 4:00 pm-6:00 pm (after agreement (per email))
  - E-mail: kjaniszczak@winfor.de
  - Coordinates the Tutorial

Information concerning the course

- Weekly assignment sheet submission
  - Submit in **Moodle** as PDF or **postbox** in room M11.25
  - Write down course name and make sure you can identify your submission on return



Preliminary Agenda

1. Linear programming
   1. Applications
   2. The Simplex Algorithm
   3. Geometry of the solution space
   4. How fast is the Simplex Method?
   5. Working with tableaus
2. Duality
   1. Motivation and the dual problem
   2. The Dual Simplex Algorithm
   3. The possible cases
   4. Interpreting the dual solution
   5. Farkas' Lemma

Preliminary Agenda

3. Computational considerations
   1. The Revised Simplex Algorithm
   2. Analyzing the complexity of the Revised Simplex Algorithm
   3. Solving the Max Flow Problem by the Revised Simplex Algorithm
   4. The Dantzig-Wolfe Decomposition
4. The Hitchcock Transportation Problem
   1. Using the standardized Simplex Algorithm
   2. The MODI Algorithm
5. The Primal-Dual Simplex Algorithm

Preliminary Agenda

6. Optimally solving the Shortest Path Problem
   1. Deriving the Dijkstra algorithm
   2. Bellman-Ford algorithm
   3. The Floyd-Warshall algorithm
7. Max-Flow and Min Cut Problems
   1. Max-Flow Problems
   2. Min-Cut Problems
   3. A Primal-Dual algorithm
   4. The Ford-Fulkerson algorithm
   5. Analyzing the Ford-Fulkerson algorithm
   6. An efficient algorithm for the Max-Flow Problem

Preliminary Agenda

8. Applying the Primal-Dual Simplex to the Transportation Problem – The alpha-beta algorithm
   1. Problem definition and analysis
   2. Analyzing the reduced primal (RP)
   3. Solving the DRP
   4. Complexity of the algorithm
9. Integer programming
   1. Basics
   2. Cutting Plane Methods – algorithm of Gomory
   3. A Branch&Bound Algorithm

Preliminary Agenda

10. Matrix Games
    1. Introducing examples
    2. Basic definitions
    3. Games and Linear Programming

Selected basic Literature

- Brucker, P.; Knust, S. (2012): Complex Scheduling. 2. ed., Springer, Berlin, Heidelberg, New York.
- Domschke, W.; Drexl, A.; Klein, R.; Scholl, A.; Voß, S. (2015): Übungen und Fallbeispiele zum Operations Research. 8. Aufl., Springer Gabler, 2015.
- Domschke, W.; Drexl, A. (2015): Einführung in Operations Research. 9. Aufl., Springer Gabler.
- Myerson, R.B. (1997): Game Theory. Analysis of Conflict. Havard University Press.
- Nemhauser,G.L.,&Wolsey,L.A.(1988).*Integer and combinatorial optimization*. JohnWiley&Sons, New York.
- Papadimitriou, C.H.; Steiglitz, K. (1982, 1988): Combinatorial Optimization. Algorithms and Complexity. Prentice-Hall, 1982 and Dover unabridged edition 1998.

Selected basic Literature

- Suhl, L.; Mellouli, T. (2013): Optimierungssysteme. 3. Aufl., Springer Gabler.
- Taha, H.A. (2010): Operations Research. An Introduction. 9th ed, Pearson Education.
- Chvátal, V. (2002): Linear Programming. 16th print. W.H. Freeman and Company, New York.
- Wolsey, L.A. (1998): Integer Programming. John Wiley&Sons.

And thousands of other good books dealing with Optimization, Linear Programming, or Combinatorial Optimization (references to further papers will be given in the respective sections)

Simplex calculators

- Excel Solver
  - Can be activated under Extras→Add-Ins (2003 Version), File→Options→Add-Ins (2010 Version)
  - Subsequently, you may use the Solver by Extras→Solver (2003 Version), Data→Solver (2010 Version)
  - It is not powerful but nice to play around with our simple examples
- Online Simplex calculator:

  http://www.zweigmedia.com/RealWorld/simplex.html

1 Linear Programming

- We deal with a large class of problems in this first section
- These problems can be mapped as Linear Programs, i.e.,
  - We define continuous variables
  - We define linear constraints to be fulfilled by the values of the variables
  - We define an objective function that provides an evaluation of each solution found
  - We want to find optimal solutions, i.e., solutions that fulfill all restrictions (then we denote them as feasible) and maximize or minimize the objective function value

Linear Program – Main attributes

- continuous decision variables
- linear constraints that must be fulfilled by the values of the decision variables
- objective function that provides an evaluation of each solution found

→ Finding an optimal solution

Solution: vector of decision variables
Feasible solution: solution that fulfills all constraints
Optimal solution: feasible solution with maximal or minimal objective function value

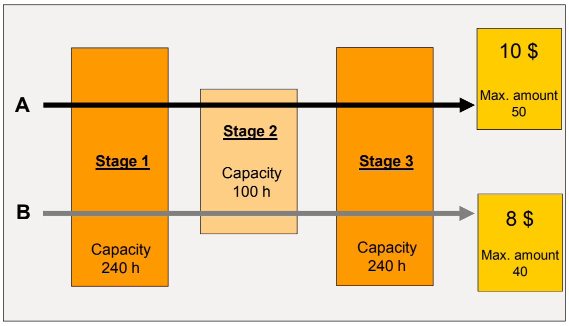
1.1 Linear Programming Applications

- We commence our work with representative applications of Linear Programming
  - Production Program Planning
  - Hitchcock problem, i.e., standardized balanced transportation problem
  - Diet Problem

Application 1 – Production Program Planning

- The production management of a plant of an orange juice producer plans the production program
- There are two types of orange juices that are pressed and mixed in this plant
- For simplicity reasons, let us denote them as A and B
- Both are produced on 3 stages in a predetermined sequence, i.e., 1 – 2 – 3 is the production sequence for both product types
- This is illustrated by the following figure

Production Program Planning - Illustration



… and just the values

- All types are produced on all stages
- Capacity on stages 1 and 3 are 240 h,
- Capacity on stage 2 is 100 h

- **Product A**
  - Price per gallon: 10$
  - Variable costs per gallon: 5$
  - Thus, we obtain a **marginal profit of 5$ per gallon**
  - Max. sales volume: 50 gallons
  - In order to produce one gallon of A
    - on stage 1, we need 2 hours,
    - on stage 2, 1 hour,
    - and on stage 3, 4 hours

… and product B

- **Product B**
  - Price per gallon: 8$
  - Variable costs per gallon: 2$
  - Thus, we obtain a **marginal profit of 6$ per gallon**
  - Max. sales volume: 40 gallons
  - In order to produce one gallon of B
    - on stage 1, A gallon B requires 4 hours
    - on stage 2, A gallon B requires 2 hours, and
    - on stage 3, A gallon B requires 2 hours

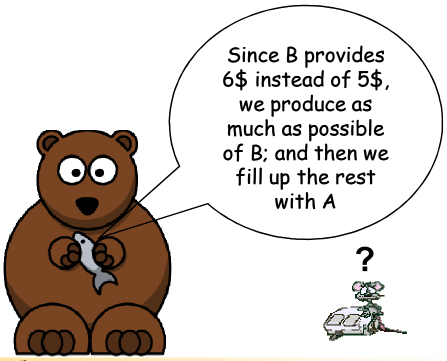
Optimal production program

- Clearly, we want to maximize our profit, i.e., the maximally obtainable total marginal profit
- Thus, we analyze what we are able to sell maximally
  - Both types of orange juice are worth to produce
  - Each item of A brings us a marginal profit of 5$ per gallon
  - Each item of B even 6$
- Consequently, we want to produce as much as possible of both products
- If the maximum volumes of sale can be produced, we have found the optimal production program

We calculate the maximum demand

- **We have the following demand levels**
  - Stage 1: $50 \cdot 2+40 \cdot 4=260>240$
    Thus, since demand is larger than capacity, we have a bottleneck !
  - Stage 2: $50 \cdot 1+40 \cdot 2=130>100$
    Thus, since demand is larger than capacity, we have a bottleneck !
  - Stage 3: $50 \cdot 4+40 \cdot 2=280>240$
    Thus, since demand is larger than capacity, we have a bottleneck !

What is to do?



If we do so, it turns out that …

- B needs
  - on stage 1 4h. Maximally produce min{40, 240/4} = 40
  - on stage 2 2h. Maximally produce min{40, 100/2} = 40
  - on stage 3 2h. Maximally produce min{40, 240/2} = 40

  Thus, B can be produced in its maximum volume of sales

- A needs
  - on stage 1 2h. Maximally produce min{50, (240-160)/2} = 40
  - on stage 2 1h. Maximally produce min{50, (100-80)/1=20/1} = 20
  - on stage 3 4h. Maximally produce min{50, (240-80)/4=160/4} = 40

  Thus, 20 items of A can be additionally produced and sold

Results in

a total profit of

20·5$ + 40·6$ = 100$ + 240$ = **340$**

➢ However, in order to analyze the problem more in detail, we want to formalize it

Linear Program of the production

Decision variables: x_A : Produced gallons of juice A
x_B : Produced gallons of juice B

Objective function: Maximize $z = 5 \cdot x_A + 6 \cdot x_B$ — Maximize the total revenue

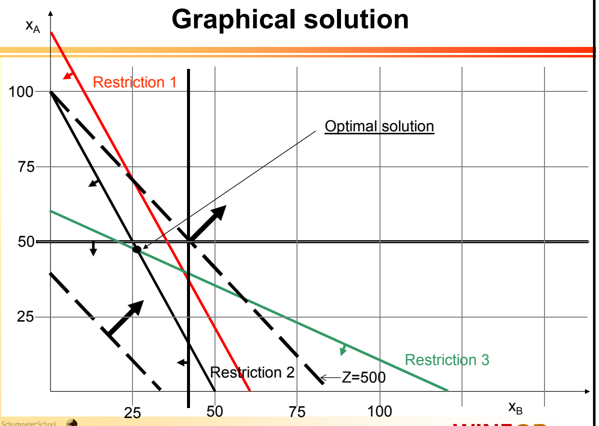
Constraints: subject to

$x_A \leq 50$ — Maximium volume of sales
$x_B \leq 40$

$2 \cdot x_A + 4 \cdot x_B \leq 240$ — Production capacities
$1 \cdot x_A + 2 \cdot x_B \leq 100$
$4 \cdot x_A + 2 \cdot x_B \leq 240$

$x_A, x_B \geq 0$ — Non-negativity constraints

Graphical solution



Optimal solution

- Obviously, the optimal solution is located at the point of intersection of restriction 2 and 3
- Thus, we have to solve

$$\begin{vmatrix} 1 \cdot x_A + 2 \cdot x_B = 100 \\ 4 \cdot x_A + 2 \cdot x_B = 240 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 100 - 2 \cdot x_B \\ 2 \cdot x_A + x_B = 120 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 100 - 2 \cdot x_B \\ 2 \cdot (100 - 2 \cdot x_B) + x_B = 120 \end{vmatrix}$$

$$\Leftrightarrow \begin{vmatrix} x_A = 100 - 2 \cdot x_B \\ 200 - 4 \cdot x_B + x_B = 120 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 100 - 2 \cdot x_B \\ 200 - 3 \cdot x_B = 120 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 100 - 2 \cdot x_B \\ x_B = 80/3 \end{vmatrix}$$

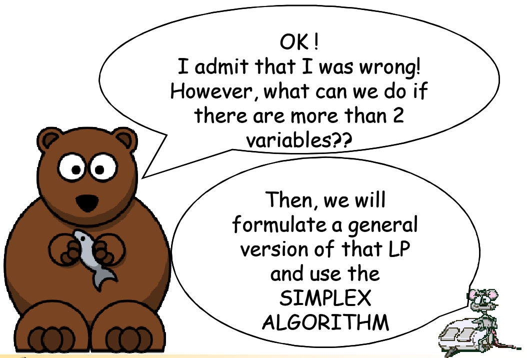
$$\Leftrightarrow \begin{vmatrix} x_A = 300/3 - 160/3 \\ x_B = 80/3 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 140/3 \\ x_B = 80/3 \end{vmatrix} \Leftrightarrow \begin{vmatrix} x_A = 46,\overline{6} \\ x_B = 26,\overline{6} \end{vmatrix}$$

Results in

a total (optimal) profit of

$26.\overline{6} \cdot 6\$ + 46.\overline{6} \cdot 5\$ = 233.\overline{3}\$ + 159.\overline{9}\$ =$ **$393.\overline{3}\$$**

Consequence



Production Program Planning

p_j : Marginal profit per item of product type $j = 1,...,n$ [MU/PU]

$c_{i,j}$: Production coefficient of product type $j = 1,...,n$ on machine $i = 1,...,m$ [CU/PU]

C_i : Maximum capacity of machine $i = 1,...,m$ [CU]

X_j : Maximum number of items of product type $j = 1,...,n$ salable in the planning horizon [PU]

x_j : Number of items of product type $j = 1,...,n$ to be produced in the planning horizon [PU]

Maximize $\sum_{j=1}^{n} p_j \cdot x_j$

subject to $\forall i \in \{1,...,m\} : \sum_{j=1}^{n} c_{i,j} \cdot x_j \le C_i$

$\forall j \in \{1,...,n\} : 0 \le x_j \le X_j$

Using Excel

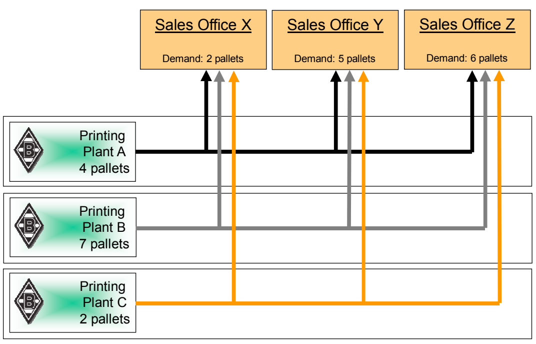
- The program Excel comprises a standard solver for Linear Programs
- It is neither really high-performance nor convenient to use but available and sufficient for our exemplary problem constellations
- Activate the Solver by Extras→Add-Ins (2003 Version), File→Options→Add-Ins (2010 Version)
- Subsequently, you may use the Solver by Extras→Solver (2003 Version), Data→Solver (2010 Version)
- Let us try it out…

Microsoft Excel-Arbeitsblatt

Application 2 – The Hitchcock Problem

- A service agent has three sales offices (C_X, C_Y, and C_Z) in Wuppertal
- These offices are supplied by three local printing plants (P_A, P_B, and P_C)
- In order to satisfy the numerous soccer fans in Wuppertal, the service agent has an exclusive license of sale for the famous football/soccer club Borussia Mönchengladbach
- While the tickets are printed in the printing plants at equal costs, the transport to the offices causes individual costs
- Additional information
  - Each printing plant has an individual inventory for the next day. Note that this inventory is extremely perishable
  - Each sales office has an individual maximum amount of sales

The Hitchcock Problem – Illustration



The Matrix

Producers			
	A	B	C
Supply	4	7	2
Consumers			
	X	Y	Z
Demand	2	5	6

Transportation Distances

Distance	X	Y	Z
A	2	3	4
B	4	6	8
C	5	2	7

What is the objective?

- Obviously, we have to decide about the quantities to be transported along each relation between a printing plant and a ticket office
- Specifically, we determine the precise number of product units that are transported along each relation
- Since quality is assumed to be negligible, a transportation cost minimization is appropriate to compare generated assignments

➢ **Thus, a solution is solely rated by the incurred transportation costs**

Let's solve the problem

- When you try something out, you usually provide a heuristic solution
- Heuristic solutions do not always guarantee a certain quality
- Usually, their performance is empirically validated or roughly anticipated for worst case scenarios
- In the following, we want to obtain some insights into the problem structure by applying some well-known simple heuristics

Minimum Method

- **Basic idea**:
  "Find the least possible combination of costs and use it exhaustively. Afterwards, proceed with the second lowest one,…"

- Let us do so…

Minimum Method

Distance	X D:2	Y D:5	Z D:6
A S:4	(2) 0	(3) 0	(4) 0
B S:7	(4) 0	(6) 0	(8) 0
C S:2	(5) 0	(2) 0	(7) 0

Minimum Method

Distance	X D:0	Y D:5	Z D:6
A S:2	(2) **2**	(3) 0	(4) 0
B S:7	(4) 0	(6) 0	(8) 0
C S:2	(5) 0	(2) 0	(7) 0

Minimum Method

Distance	X D:0	Y D:3	Z D:6
A S:2	(2) **2**	(3) 0	(4) 0
B S:7	(4) 0	(6) 0	(8) 0
C S:0	(5) 0	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:1	Z D:6
A S:0	(2) **2**	(3) **2**	(4) 0
B S:7	(4) 0	(6) 0	(8) 0
C S:0	(5) 0	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:1	Z D:6
A S:0	(2) **2**	(3) **2**	(4) 0
B S:7	(4) **0**	(6) 0	(8) 0
C S:0	(5) 0	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:1	Z D:6
A S:0	(2) **2**	(3) **2**	(4) **0**
B S:7	(4) **0**	(6) 0	(8) 0
C S:0	(5) 0	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:1	Z D:6
A S:0	(2) **2**	(3) **2**	(4) **0**
B S:7	(4) **0**	(6) 0	(8) 0
C S:0	(5) **0**	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:0	Z D:6
A S:0	(2) **2**	(3) **2**	(4) **0**
B S:6	(4) **0**	(6) **1**	(8) 0
C S:0	(5) **0**	(2) **2**	(7) 0

Minimum Method

Distance	X D:0	Y D:0	Z D:6
A S:0	(2) 2	(3) 2	(4) 0
B S:6	(4) **0**	(6) **1**	(8) 0
C S:0	(5) **0**	(2) **2**	(7) **0**

Minimum Method

Total costs = 68

Distance	X D:0	Y D:0	Z D:0
A S:0	(2) **2**	(3) **2**	(4) **0**
B S:0	(4) **0**	(6) **1**	(8) **6**
C S:0	(5) **0**	(2) **2**	(7) **0**

Vogel's Approximation Method

- **Basic Idea:**
  "Avoid larger deteriorations by identifying critical relations"
- Specifically, calculate the differences between the best and the second best relation for all producers and all consumers
- Select the best relation for the one with the largest difference
- Proceed until a complete solution is found

Vogel's Approximation Method

Distance	X D:2 Diff: 2	Y D:5 Diff: 1	Z D:6 **Diff: 3**
A S:4 Diff: 1	(2) 0	(3) 0	(4) 0
B S:7 Diff: 2	(4) 0	(6) 0	(8) 0
C S:2 Diff: 3	(5) 0	(2) 0	(7) 0

Vogel's Approximation Method

Distance	X D:2 Diff: 1	Y D:5 **Diff: 4**	Z D:2 Diff: 1
A S:0 Diff: 0	(2) 0	(3) 0	(4) **4**
B S:7 Diff: 2	(4) 0	(6) 0	(8) 0
C S:2 Diff: 3	(5) 0	(2) 0	(7) 0

Vogel's Approximation Method

Distance	X D:2 Diff: 0	Y D:3 Diff: 0	Z D:2 Diff: 0
A S:0 Diff: 0	(2) 0	(3) 0	(4) **4**
B S:7 **Diff: 2**	(4) 0	(6) 0	(8) 0
C S:0 Diff: 0	(5) 0	(2) **2**	(7) 0

Vogel's Approximation Method

Distance	X D:0 Diff: 0	Y D:3 Diff: 0	Z D:2 Diff: 0
A S:0 Diff: 0	(2) 0	(3) 0	(4) **4**
B S:5 **Diff: 2**	(4) **2**	(6) 0	(8) 0
C S:0 Diff: 0	(5) 0	(2) **2**	(7) 0

Vogel's Approximation Method

Distance	X D:0 Diff: 0	Y D:0 Diff: 0	Z D:2 Diff: 0
A S:0 Diff: 0	(2) 0	(3) 0	(4) **4**
B S:2 **Diff: 2**	(4) **2**	(6) **3**	(8) 0
C S:0 Diff: 0	(5) 0	(2) **2**	(7) 0

Vogel's Approximation Method

Total costs = 62

Distance	X D:0 Diff: 0	Y D:0 Diff: 0	Z D:0 Diff: 0
A S:0 Diff: 0	(2) 0	(3) 0	(4) **4**
B S:0 Diff: 2	(4) **2**	(6) **3**	(8) **2**
C S:0 Diff: 0	(5) 0	(2) **2**	(7) 0

Local Improvement Operations

- We may improve an existing solution by applying specific transformation moves, i.e., we slightly modify a current solution in a way that
  - feasibility is maintained, and
  - solution quality is improved
- A simple example is the **pairwise shift**
- Specifically, we select two consumer-producer relations (P_1&C_1, P_2&C_2) and ask for the change in costs by…
  - transporting one unit from P_1 to C_2 rather than from P_1 to C_1
  - For P_2, we simultaneously consider the same
  - Note that feasibility is ensured by the simultaneous consideration of both constellations

Pairwise shift

Total Costs = 68

Distance	X D:0	Y D:0	Z D:0
A S:0	**(2) 2**	(3) 2	**(4) 0**
B S:0	**(4) 0**	(6) 1	**(8) 6**
C S:0	(5) 0	(2) 2	(7) 0

Pairwise shift

Total Costs are 68-4=64

Distance	X D:0	Y D:0	Z D:0
A S:0	**(2) 0**	(3) 2	**(4) 2**
B S:0	**(4) 2**	(6) 1	**(8) 4**
C S:0	(5) 0	(2) 2	(7) 0

Pairwise shift

Total Costs = 64

Distance	X D:0	Y D:0	Z D:0
A S:0	(2) 0	(3) 2	(4) 2
B S:0	(4) 2	(6) 1	(8) 4
C S:0	(5) 0	(2) 2	(7) 0

Pairwise shift

Total Costs are 64-2=62

Distance	X D:0	Y D:0	Z D:0
A S:0	(2) 0	**(3) 0**	**(4) 4**
B S:0	(4) 2	**(6) 3**	**(8) 2**
C S:0	(5) 0	(2) 2	(7) 0

The (balanced) Transportation Problem

$c_{i,j}$: Delivery costs for each product unit that is transported from supplier $i=1,\ldots,m$ to customer $j=1,\ldots,n$ [CU/PU]

a_i : Total supply of $i=1,\ldots,m$ [PU]

b_j : Total demand of $j=1,\ldots,n$ [PU]

$x_{i,j}$: Quantity that supplier $i=1,\ldots,m$ delivers to the customer $j=1,\ldots,n$ [PU]

$$\text{Minimize } \sum_{i=1}^{m}\sum_{j=1}^{n} c_{i,j}\cdot x_{i,j}$$

$$\text{subject to} \quad \forall i\in\{1,\ldots,m\}: \sum_{j=1}^{n} x_{i,j} = a_i$$

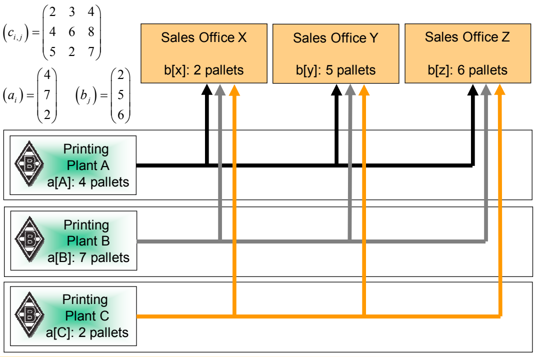
$$\forall j\in\{1,\ldots,n\}: \sum_{i=1}^{m} x_{i,j} = b_j$$

$$\forall i\in\{1,\ldots,m\}: \forall j\in\{1,\ldots,n\}: x_{i,j}\geq 0$$

The balanced Transportation Problem

$$(c_{i,j})=\begin{pmatrix}2&3&4\\4&6&8\\5&2&7\end{pmatrix}$$

$$(a_i)=\begin{pmatrix}4\\7\\2\end{pmatrix} \quad (b_j)=\begin{pmatrix}2\\5\\6\end{pmatrix}$$



The (balanced) Transportation Problem

With the previously defined parameters our problem is as follows:

$$(c_{i,j})=\begin{pmatrix}2&3&4\\4&6&8\\5&2&7\end{pmatrix};(a_i)=\begin{pmatrix}4\\7\\2\end{pmatrix};(b_j)=\begin{pmatrix}2\\5\\6\end{pmatrix}$$

$$\text{Minimize } 2\cdot x_{1,1}+3\cdot x_{1,2}+4\cdot x_{1,3}+4\cdot x_{2,1}+6\cdot x_{2,2}+8\cdot x_{2,3}+5\cdot x_{3,1}+2\cdot x_{3,2}+7\cdot x_{3,3}$$

subject to

$$\begin{array}{lllllllllr}
1\cdot x_{1,1}+ & 1\cdot x_{1,2}+ & 1\cdot x_{1,3}+ & 0\cdot x_{2,1}+ & 0\cdot x_{2,2}+ & 0\cdot x_{2,3}+ & 0\cdot x_{3,1}+ & 0\cdot x_{3,2}+ & 0\cdot x_{3,3} & =4\\
0\cdot x_{1,1}+ & 0\cdot x_{1,2}+ & 0\cdot x_{1,3}+ & 1\cdot x_{2,1}+ & 1\cdot x_{2,2}+ & 1\cdot x_{2,3}+ & 0\cdot x_{3,1}+ & 0\cdot x_{3,2}+ & 0\cdot x_{3,3} & =7\\
0\cdot x_{1,1}+ & 0\cdot x_{1,2}+ & 0\cdot x_{1,3}+ & 0\cdot x_{2,1}+ & 0\cdot x_{2,2}+ & 0\cdot x_{2,3}+ & 1\cdot x_{3,1}+ & 1\cdot x_{3,2}+ & 1\cdot x_{3,3} & =2\\
1\cdot x_{1,1}+ & 0\cdot x_{1,2}+ & 0\cdot x_{1,3}+ & 1\cdot x_{2,1}+ & 0\cdot x_{2,2}+ & 0\cdot x_{2,3}+ & 1\cdot x_{3,1}+ & 0\cdot x_{3,2}+ & 0\cdot x_{3,3} & =2\\
0\cdot x_{1,1}+ & 1\cdot x_{1,2}+ & 0\cdot x_{1,3}+ & 0\cdot x_{2,1}+ & 1\cdot x_{2,2}+ & 0\cdot x_{2,3}+ & 0\cdot x_{3,1}+ & 1\cdot x_{3,2}+ & 0\cdot x_{3,3} & =5\\
0\cdot x_{1,1}+ & 0\cdot x_{1,2}+ & 1\cdot x_{1,3}+ & 0\cdot x_{2,1}+ & 0\cdot x_{2,2}+ & 1\cdot x_{2,3}+ & 0\cdot x_{3,1}+ & 0\cdot x_{3,2}+ & 1\cdot x_{3,3} & =6
\end{array}$$

$$x_{1,1},x_{1,2},x_{1,3},x_{2,1},x_{2,2},x_{2,3},x_{3,1},x_{3,2},x_{3,3}\geq 0$$

Alternative depiction

$$\left(c_{i,j}\right)=\begin{pmatrix}2 & 3 & 4\\ 4 & 6 & 8\\ 5 & 2 & 7\end{pmatrix};\left(a_i\right)=\begin{pmatrix}4\\ 7\\ 2\end{pmatrix};\left(b_j\right)=\begin{pmatrix}2\\ 5\\ 6\end{pmatrix}$$

$$\begin{array}{l}
\text{Minimize } 2\cdot x_{1,1}+3\cdot x_{1,2}+4\cdot x_{1,3}+4\cdot x_{2,1}+6\cdot x_{2,2}+8\cdot x_{2,3}+5\cdot x_{3,1}+2\cdot x_{3,2}+7\cdot x_{3,3}\\
\text{subject to}\\
1\cdot x_{1,1}+1\cdot x_{1,2}+1\cdot x_{1,3} = 4\\
1\cdot x_{2,1}+1\cdot x_{2,2}+1\cdot x_{2,3} = 7\\
1\cdot x_{3,1}+1\cdot x_{3,2}+1\cdot x_{3,3} = 2\\
1\cdot x_{1,1}+1\cdot x_{2,1}+1\cdot x_{3,1} = 2\\
1\cdot x_{1,2}+1\cdot x_{2,2}+1\cdot x_{3,2} = 5\\
1\cdot x_{1,3}+1\cdot x_{2,3}+1\cdot x_{3,3} = 6\\
x_{1,1},x_{1,2},x_{1,3},x_{2,1},x_{2,2},x_{2,3},x_{3,1},x_{3,2},x_{3,3}\geq 0
\end{array}$$

The (balanced) Transportation Problem

We identify the characteristic structure of the problem.

$$(P) \text{ Minimize } c^T\cdot x$$

subject to

$$\begin{pmatrix}1_n^T & & & & \\ & 1_n^T & & & \\ & & \dots & \dots & \\ & & & & 1_n^T\\ E_n & E_n & E_n & E_n & E_n\end{pmatrix}\cdot x=\begin{pmatrix}a_1\\ \dots\\ \dots\\ a_m\\ b\end{pmatrix}$$

$$x=\left(x_{1,1},\dots,x_{1,n},x_{2,1},\dots,x_{2,n},\dots,x_{m,1},\dots,x_{m,n}\right)^T\geq 0$$

Application 3 – The Diet Problem

- Susan wonders how much money she has to spend on food in order to get the energy that brings her through the day
- Now, she thinks it is time to analyze…
- Altogether, she chooses six foods that seem to be cheap sources of the nutrients her body needs

Food	Size per serving	Energy (kcal)	Protein (g)	Calcium (mg)	Price ($ Cents)
Oatmeal	28 g	110	4	2	3
Chicken	100 g	205	32	12	24
Eggs	2 large	160	13	54	13
Whole milk	237 cc	160	8	285	9
Cherry pie	170 g	420	4	22	20
Pork with beans	260 g	260	14	80	19

Additional information

- Susan needs per day
  - 2,000 kcal
  - 55 g protein
  - 800 mg calcium
  - Iron and vitamins are satisfied by pills
- Consequently, 10 servings of pork and beans are sufficient per day…
  - Imagine, 10 times pork and beans per day…
  - This is disgusting
- Ok…We need to impose servings-per-day limits
  - Oatmeal: at most 4 servings per day
  - Chicken: at most 3 servings per day
  - Eggs: at most 2 servings per day
  - Milk: at most 8 servings per day
  - Cherry pie: at most 2 servings per day
  - Pork with beans: at most 2 servings per day

The Diet Problem

$x_1, x_2, x_3, x_4, x_5, x_6$ servings per day of the respective food

Minimize $3 \cdot x_1 + 24 \cdot x_2 + 13 \cdot x_3 + 9 \cdot x_4 + 20 \cdot x_5 + 19 \cdot x_6$

subject to

$0 \le x_1 \le 4 \wedge 0 \le x_2 \le 3 \wedge 0 \le x_3 \le 2 \wedge 0 \le x_4 \le 8 \wedge 0 \le x_5 \le 2 \wedge 0 \le x_6 \le 2$

$110 \cdot x_1 + 205 \cdot x_2 + 160 \cdot x_3 + 160 \cdot x_4 + 420 \cdot x_5 + 260 \cdot x_6 \ge 2000$

$4 \cdot x_1 + 32 \cdot x_2 + 13 \cdot x_3 + 8 \cdot x_4 + 4 \cdot x_5 + 14 \cdot x_6 \ge 55$

$2 \cdot x_1 + 12 \cdot x_2 + 54 \cdot x_3 + 285 \cdot x_4 + 22 \cdot x_5 + 80 \cdot x_6 \ge 800$

…and in general

$a_{i,j}$: Amount of ith nutrient in a serving of the jth food, $i = 1,...,m$, $j = 1,...,n$.

r_i: Daily requirement for the ith nutrient, $i = 1,...,m$.

c_j: Cost per serving for the jth food, $j = 1,...,n$.

X_j: Maximum number on servings of the jth food, $j = 1,...,n$.

Decision variables:

x_j: Daily consumption of the jth food, $j = 1,...,n$. A diet is denoted by a choice of a vector $x \ge 0, x \in IR^n$.

Objective function: Minimize $\sum_{j=1}^{n} c_j \cdot x_j$

subject to $\forall i \in \{1,...,m\}: \sum_{j=1}^{n} a_{i,j} \cdot x_j \ge r_i$

$\forall j \in \{1,...,n\}: 0 \le x_j \le X_j$

Consequence

- All applications are completely different, but their mathematical definitions are somehow strongly related
- All LPs have in common that…
  - …the variables are continuous
  - …the objective function is linear
  - …the restrictions are linear
  - …the objective function is either a maximization or minimization
  - …restrictions require the fulfillment of a minimum or maximum bound

LP in general

- In what follows, we introduce general forms in order to define **what a Linear Program (LP) is**
- In Literature, different forms of LPs are distinguished. Specifically, it can be found for instance
  - LP in general form
  - LP in canonical form
  - LP in standard form
- The Reader should be warned that this classification is far away from being unambiguous
- Moreover, what we will denote as a Linear Program in standard form is frequently introduced as the LP in canonical form

General Form

Let $A \in IR^{m \times n}$ with $A = \begin{pmatrix} a_1'^T \\ \dots \\ a_m'^T \end{pmatrix} \wedge a_i' \in IR^n, i \in \{1,\dots,m\}$, and $b \in IR^m$.

Furthermore, let M be the set of row indices corresponding to equality constraints, and let $\bar{M}$ be the set of row indices corresponding to inequality constraints. Additionally, let N be the set of column indices corresponding to constrained variables, and let $\bar{N}$ be the set of column indices corresponding to unrestricted variables. Then, the feasible solution space P is

$P = \left\{ x \in IR^n \mid \forall j \in N : x_j \geq 0 \wedge \forall i \in M : a_i'^T \cdot x = b_i \wedge \forall i \in \bar{M} : a_i'^T \cdot x \leq b_i \right\}$.

Furthermore, for $c \in IR^n$, we pursue the maximization of $z(x) = c^T \cdot x$.

Note that M and $\bar{M}$ form a partition of $\{1,\dots,m\}$. Moreover, N and $\bar{N}$ are a partition of $\{1,\dots,n\}$.

Canonical Form

Let $A \in IR^{m \times n}$ and $b \in IR^m$:
Then, the set of feasible solutions is defined as follows:
$P = \{x \in IR^n \mid x \geq 0 \text{ and } A \cdot x \leq b\}$
Solutions that belong to P are denoted as feasible.

In order to evaluate a solution x that is found, we introduce an additional vector. Hence, let $c \in IR^n : z(x) = c^T \cdot x$.

In the following, we pursue the maximization of z under the constraints $x \geq 0$ and $A \cdot x \leq b$.

Standard Form

Let $A \in IR^{m \times n}$ and $b \in IR^m$:
Then, the set of feasible solutions is defined as follows:
$P = \{x \in IR^n \mid x \geq 0 \text{ and } A \cdot x = b\}$
Solutions that belong to P are denoted as feasible.

In order to evaluate solution that is found, we introduce an additional vector. Hence, let $c \in IR^n$: $z(x) = c^T \cdot x$.

In the following, we pursue the minimization or maximization of z under the constraints $x \geq 0$ and $A \cdot x = b$.

Problem transformations

- In order to prove that it is **sufficient to consider LPs in standard form only**, we have to think about problem transformations
- Obviously, in particular, the diet problem does not correspond to our class
- In addition, what about equalities?
- And what about unrestricted variables, i.e., variables that may become negative?
- This is briefly considered in the following

Transformation – Equality I

Replace $\sum_{j=1}^{n} a_{i,j} \cdot x_j = b_i$ by two inequalities

$$\sum_{j=1}^{n} a_{i,j} \cdot x_j \le b_i$$

$$\wedge$$

$$(-1) \cdot \left(\sum_{j=1}^{n} a_{i,j} \cdot x_j \right) \le (-1) \cdot (b_i)$$

Transformation – Equality II

Replace $\sum_{j=1}^{n} a_{i,j} \cdot x_j \le b_i$ by introducing a new (positive) variable that is called a **slack variable**:

$$\sum_{j=1}^{n} a_{i,j} \cdot x_j + y_i = b_i \wedge y_i \ge 0$$

We call all variables x that belong to the original problem **structure variables** to distinguish them from the slack variables.

Transformation – Objective function

Just replace the original objective function

Minimize $\sum_{j=1}^{n} c_j \cdot x_j$

by the modified equivalent objective function

Maximize $\sum_{j=1}^{n} (-1) \cdot c_j \cdot x_j$

Transformation – Free variables

Create two additional variables $x_j^+ \geq 0, x_j^- \geq 0$ for each unrestricted variable $x_j \in IR$, and substitute $x_j = x_j^+ - x_j^-$.

This leads to a doubling of the corresponding jth column in c^T and in A while the added column is multiplied with -1:

$$(c_1,\dots,c_j,\dots,c_n)\cdot\begin{pmatrix}x_1\\ \dots\\ x_j\\ \dots\\ x_n\end{pmatrix} \rightarrow (c_1,\dots,c_j,-c_j,\dots,c_n)\cdot\begin{pmatrix}x_1\\ \dots\\ x_j^+\\ x_j^-\\ \dots\\ x_n\end{pmatrix} \text{ and } A\cdot x = (a_1,\dots,a_j,\dots,a_n)\cdot\begin{pmatrix}x_1\\ \dots\\ x_j\\ \dots\\ x_n\end{pmatrix}$$

$$\rightarrow (a_1,\dots,a_j,-a_j,\dots,a_n)\cdot\begin{pmatrix}x_1\\ \dots\\ x_j^+\\ x_j^-\\ \dots\\ x_n\end{pmatrix}, \text{with:} \forall i \in \{1,\dots,n\}: c_i \in IR, \forall i \in \{1,\dots,n\}: a_i \in IR^m$$

Conclusion

- Since all forms of LPs are equivalent, we switch between them arbitrarily
- I.e., we always use the form that seems to be most useful

1.2 The Simplex Method

- Coming back to the Production Program Planning, we consider now the following problem constellation
- This time, a producer has to decide about 3 product types (1,2, and 3) to be produced on 3 stages
  - Product 1
    - Marginal Profit: 5
    - Production Coefficients: 2, 4, and 3
  - Product 2
    - Marginal Profit: 4
    - Production Coefficients: 3, 1, and 4
  - Product 3
    - Marginal Profit: 3
    - Production Coefficients: 1, 2, and 2

Altogether, we get

$$
\begin{aligned}
&\text{Max } 5\cdot x_1 + 4\cdot x_2 + 3\cdot x_3 = z \\
&\text{s.t.} \quad 2\cdot x_1 + 3\cdot x_2 + 1\cdot x_3 \le 5 \\
&\qquad 4\cdot x_1 + 1\cdot x_2 + 2\cdot x_3 \le 11 \\
&\qquad 3\cdot x_1 + 4\cdot x_2 + 2\cdot x_3 \le 8 \\
&\qquad x_1, x_2, x_3 \ge 0
\end{aligned}
$$

Getting equalities by slack variables

$$
\begin{aligned}
&\text{Max } z = 5\cdot x_1 + 4\cdot x_2 + 3\cdot x_3 \\
&\text{s.t.} \quad 2\cdot x_1 + 3\cdot x_2 + 1\cdot x_3 \le 5 \\
&\qquad 4\cdot x_1 + 1\cdot x_2 + 2\cdot x_3 \le 11 \\
&\qquad 3\cdot x_1 + 4\cdot x_2 + 2\cdot x_3 \le 8 \\
&\qquad x_1, x_2, x_3 \ge 0
\end{aligned}
$$

$$
\begin{aligned}
&\text{Max } z = 5\cdot x_1 + 4\cdot x_2 + 3\cdot x_3 \\
&\text{s.t.} \quad 2\cdot x_1 + 3\cdot x_2 + 1\cdot x_3 + x_4 = 5 \\
&\qquad 4\cdot x_1 + 1\cdot x_2 + 2\cdot x_3 + x_5 = 11 \\
&\qquad 3\cdot x_1 + 4\cdot x_2 + 2\cdot x_3 + x_6 = 8 \\
&\qquad x_1, x_2, x_3, x_4, x_5, x_6 \ge 0
\end{aligned}
$$

- The variables in the original LP are denoted as **structure variables**
- In contrast to this, the variables that are additionally introduced in the second LP (with equalities) are denoted as **slack variables**

Getting a first solution

$$
\begin{aligned}
&\text{Max } z = 5\cdot x_1 + 4\cdot x_2 + 3\cdot x_3 \\
&\text{s.t.} \quad x_4 = 5 - 2\cdot x_1 - 3\cdot x_2 - 1\cdot x_3 \\
&\qquad x_5 = 11 - 4\cdot x_1 - 1\cdot x_2 - 2\cdot x_3 \\
&\qquad x_6 = 8 - 3\cdot x_1 - 4\cdot x_2 - 2\cdot x_3 \\
&\qquad x_1, x_2, x_3, x_4, x_5, x_6 \ge 0
\end{aligned}
$$

This is a **dictionary**.

We call the $n-m$ variables that appear on the right-hand side **non-basic variables** and set their value to zero.

The m variables on the left-hand side (i.e., **the basic variables**) form a **basic solution**: $x_4 = 5 \wedge x_5 = 11 \wedge x_6 = 8 \Rightarrow z = 0$.

We call the objective function now **reduced costs**.

Quality of the solution found…

- …is obviously not really convincing ($z=0$)
- How can we improve it?
  - The value is not surprising
    - Only slack variables are unequal to zero
    - Slack variables which are unequal to zero do not provide any benefit according to the objective function value
  - Let us consider the set of variables
    - For this purpose, we consider the objective function coefficients belonging to the current solution
    - Owing to positive coefficients, an increase of x_1, x_2, or x_3 will raise z. Since x_1 has the largest positive coefficient, we first try it out. We call this pivoting strategy the **largest coefficient rule**
    - How much can we increase x_1?

$$\text{Max } 5 \cdot x_1 + 4 \cdot x_2 + 3 \cdot x_3 = z$$

How much can we increase x_1?

$$\text{Max } 5 \cdot x_1 + 4 \cdot x_2 + 3 \cdot x_3 = z$$

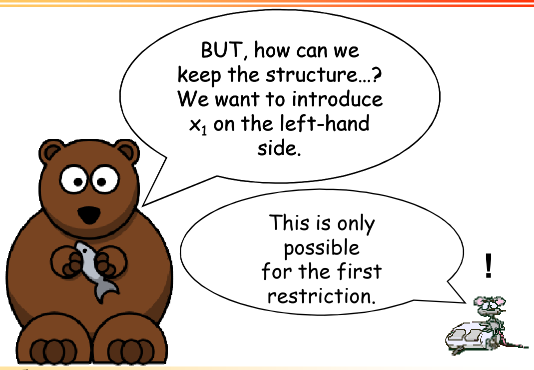
$$\text{s.t.} \quad x_4 = 5 - 2 \cdot x_1 - 3 \cdot x_2 - 1 \cdot x_3 \geq 0 \Rightarrow 5 \geq 2 \cdot x_1 \Rightarrow 5/2 \geq x_1$$

$$x_5 = 11 - 4 \cdot x_1 - 1 \cdot x_2 - 2 \cdot x_3 \geq 0 \Rightarrow 11 - 4 \cdot x_1 \geq 0 \Rightarrow 11/4 \geq x_1$$

$$x_6 = 8 - 3 \cdot x_1 - 4 \cdot x_2 - 2 \cdot x_3 \geq 0 \Rightarrow 8 - 3 \cdot x_1 \geq 0 \Rightarrow 8/3 \geq x_1$$

$$\Rightarrow x_1 = \min\left\{5/2, 11/4, 8/3\right\} = 5/2$$

And now?



Transforming the dictionary

$$\text{Max } 5\cdot\left(5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4\right) + 4\cdot x_2 + 3\cdot x_3 = z$$

$$\begin{aligned}
\text{s.t. } \; & x_1 = 5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4 \\
& x_5 = 11 - 4\cdot\left(5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4\right) - 1\cdot x_2 - 2\cdot x_3 \\
& x_6 = 8 - 3\cdot\left(5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4\right) - 4\cdot x_2 - 2\cdot x_3 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
\end{aligned}$$

$$\text{Max } 25/2 - 7/2\cdot x_2 + 1/2\cdot x_3 - 5/2\cdot x_4 = z$$

$$\begin{aligned}
\text{s.t. } \; & x_1 = 5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4 \\
& x_5 = 1 + 5\cdot x_2 + 2\cdot x_4 \\
& x_6 = 1/2 + 1/2\cdot x_2 - 1/2\cdot x_3 + 3/2\cdot x_4 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \Rightarrow x_1 = 5/2 \wedge x_5 = 1 \wedge x_6 = 1/2 \Rightarrow z = 25/2
\end{aligned}$$

Consider the solution

$$\boxed{\text{Max } 25/2 - 7/2\cdot x_2 + 1/2\cdot x_3 - 5/2\cdot x_4 = z}$$

- The objective function reveals us that we can improve the solution further.
- This is possible by increasing x_3.
- Again, we ask for bounds limiting the increase of this variable.
- For this purpose, we have to consider the dictionary.

How much can we increase x_3?

$$\text{Max } 25/2 - 7/2\cdot x_2 + 1/2\cdot x_3 - 5/2\cdot x_4 = z$$

$$\begin{aligned}
\text{s.t. } \; & x_1 = 5/2 - 3/2\cdot x_2 - 1/2\cdot x_3 - 1/2\cdot x_4 \geq 0 \Rightarrow x_3 \leq 5 \\
& x_5 = 1 + 5\cdot x_2 + 2\cdot x_4 \geq 0 \\
& x_6 = 1/2 + 1/2\cdot x_2 - 1/2\cdot x_3 + 3/2\cdot x_4 \geq 0 \Rightarrow x_3 \leq 1 \\
& \Rightarrow x_3 = \min\{5,1\} = 1
\end{aligned}$$

Transforming the dictionary

$$\text{Max } {}^{25}\!/_{2} - {}^{7}\!/_{2} \cdot x_2 + {}^{1}\!/_{2} \cdot (1 + x_2 + 3 \cdot x_4 - 2 \cdot x_6) - {}^{5}\!/_{2} \cdot x_4 = z$$

$$\text{s.t.} \quad x_1 = {}^{5}\!/_{2} - {}^{3}\!/_{2} \cdot x_2 - {}^{1}\!/_{2} \cdot (1 + x_2 + 3 \cdot x_4 - 2 \cdot x_6) - {}^{1}\!/_{2} \cdot x_4$$

$$x_5 = 1 + 5 \cdot x_2 + 2 \cdot x_4$$

$$2 \cdot x_6 = 1 + 1 \cdot x_2 - 1 \cdot x_3 + 3 \cdot x_4 \Rightarrow x_3 = 1 + x_2 + 3 \cdot x_4 - 2 \cdot x_6$$

$$\text{Max } 13 - 3 \cdot x_2 - x_4 - x_6 = z$$

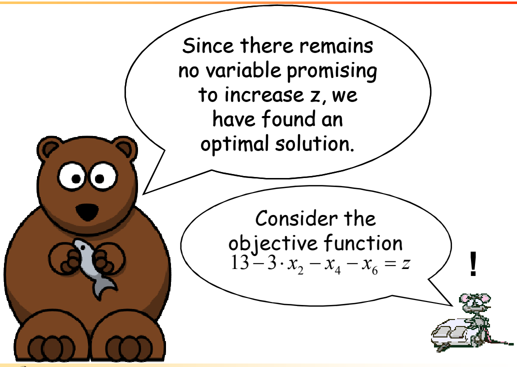
$$\text{s.t.} \quad x_1 = 2 - 2 \cdot x_2 - 2 \cdot x_4 + x_6$$

$$x_5 = 1 + 5 \cdot x_2 + 2 \cdot x_4$$

$$x_3 = 1 + x_2 + 3 \cdot x_4 - 2 \cdot x_6$$

$$x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \Rightarrow x_1 = 2 \wedge x_3 = 1 \wedge x_5 = 1 \Rightarrow z = 13$$

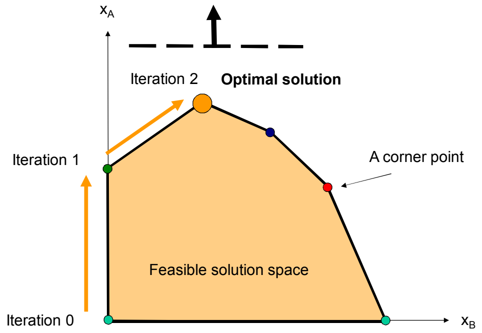
And now?



Calculation with dictionaries

- Left-hand side
  - Variables that are allowed to be unequal to zero
  - Here, we have altogether *m* variables
  - We call these variables **basic variables**
- Right-hand side
  - Variables that are equal to zero
  - Here, we have altogether at least *n-m* variables
  - We call these variables **non-basic variables**
- Objective function
  - Positive coefficients increase the objective function value and vice versa.
  - Later on, coefficients that belong to structure variables are denoted as **reduced costs**
- We execute a swap of a basic and a non-basic variable in each step.
- By doing so, we try to improve the solution quality.
- Moreover, we jump along the edge of the solution space. More specifically, **from corner point to corner point**

The solution space and the Simplex Algorithm



The Primal Simplex Algorithm with Dictionaries

1. Transform the problem into a canonical form and generate equations.
2. Initialization with a feasible basic solution.
3. Are there *strictly* positive reduced cost coefficients in the current solution?

"Yes": Iteration:
- Largest coefficient rule: Choose a variable x_B that has the largest positive reduced cost coefficient.
- Determine a positive upper bound on x_B.
- If there exists an upper bound on feasible values for x_B, set x_B to the minimal upper bound that is given by equation i ; otherwise terminate since the solution space is unbounded and no optimal solution exists.
- Transform equation i such that x_B appears on the left-hand side.
- Substitute x_B in all other equations as well as in the objective function with the obtained equation i.
- Rearrange the equation system.
- Go to step 3.

"No": Termination. An optimal basic solution is found.

Pitfalls and how to avoid them

- The presented calculation went pretty smoothly
- The danger that may occur was not pointed out
- Three kinds of pitfalls have to be considered
  - Initialization
    - Obviously, we need an initial solution
    - Are there constellations thinkable where this is not possible?
  - Iteration
    - Is there a danger of getting stuck throughout the calculation?
    - Is it always possible to swap from one basic solution to the next one?
  - Termination
    - Is the calculation always finite?
    - Are cyclical computations possible?

Initialization

- For what follows, we need at first a feasible solution to the LP. Fortunately, this is quite simple to provide
- If *b* is positive, we may just make use of the introduction of slack variables; i.e., all structure variables are set to zero and slack variables equal the right-hand side *b*
- Otherwise, we apply the simple procedure that is depicted on the following slides

Pitfall: Initialization with a feasible solution

- If ***b*≥0**, take the **trivial solution**: all structure variables are set to zero and all slack variables equal the right-hand side *b*
- If there is an *i* with **b_i<0**, apply the **Two-Phase Method**

1. Auxiliary LP: Maximize $z = -x_0$
   subject to $A \cdot x - 1_m \cdot x_0 \le b$
   $x, x_0 \ge 0$
2. Initial solution to the auxiliary LP: $x^{ini} = \left(x_1^{ini}, \ldots, x_n^{ini}, x_0^{ini}\right)^T$
   with $x_1^{ini} = \ldots = x_n^{ini} = 0 \wedge x_0^{ini} = -\min\left\{b_i \mid b_i < 0 \wedge i \in \{1, \ldots, m\}\right\}$
   Since $i \in \{1, \ldots, m\}$ exists with $b_i < 0$, x^{ini} is feasible
3. Solve the auxiliary LP and get its optimal solution x^*
4. If z>0, terminate (this procedure) because the original LP is not solvable
5. Initial feasible solution to the original LP: $x^* = \left(x_1^*, \ldots, x_n^*\right)^T$

Two-Phase Method – Conclusions I

1.2.1 Observation: Since the objective function value is lower bounded by zero, the auxiliary LP is solvable

1.2.2 Lemma: If and only if the optimal solution to the auxiliary problem has the objective function value zero, the original LP is solvable

Proof: "⇒": Since *z=0* holds x_0*=0* follows. The optimal auxiliary LP solution yields a feasible solution to the original LP
"⇐": If the original problem is solvable, we have x_0*=0* and, therefore, *z=0*

Two-Phase Method – Conclusions II

Optimal auxiliary LP solution:

- $x_0>0$ is basic:
  The original LP is **not solvable** because at least one constraint is violated
- x_0 is non-basic or $x_0=0$ is basic:
  Erase x_0 and switch to the original LP with the feasible solution just generated

  Special case here:
  $x_0=0$ is basic: Consider the next-to-last step where the objective function becomes zero. Here, x_0 was decreased to zero. Consequently, in this step x_0 was a candidate for being erased. Hence, we can adjust this step accordingly

Initialization – Example I

$$\begin{aligned} &\text{Max } x_1 - x_2 + x_3 = z \\ &\text{s.t.} \quad 2\cdot x_1 - x_2 + 2\cdot x_3 \le 4 \\ &\qquad 2\cdot x_1 - 3\cdot x_2 + x_3 \le -5 \\ &\qquad -x_1 + x_2 - 2\cdot x_3 \le -1 \\ &\qquad x_1, x_2, x_3 \ge 0 \end{aligned}$$

$$\begin{aligned} &\text{Max } -x_0 = z \\ &\text{s.t.} \quad 2\cdot x_1 - x_2 + 2\cdot x_3 - x_0 + x_4 = 4 \\ &\qquad 2\cdot x_1 - 3\cdot x_2 + x_3 - x_0 + x_5 = -5 \leftarrow \min \\ &\qquad -x_1 + x_2 - 2\cdot x_3 - x_0 + x_6 = -1 \\ &\qquad x_0, x_1, x_2, x_3, x_4, x_5, x_6 \ge 0 \end{aligned}$$

Initialization – Example II

$$\begin{aligned} &\text{Max } -(5 + 2\cdot x_1 - 3\cdot x_2 + x_3 + x_5) = z \\ &\text{s.t.} \quad 2\cdot x_1 - x_2 + 2\cdot x_3 - (5 + 2\cdot x_1 - 3\cdot x_2 + x_3 + x_5) + x_4 = 4 \\ &\qquad x_0 = 5 + 2\cdot x_1 - 3\cdot x_2 + x_3 + x_5 \\ &\qquad -x_1 + x_2 - 2\cdot x_3 - (5 + 2\cdot x_1 - 3\cdot x_2 + x_3 + x_5) + x_6 = -1 \\ &\qquad x_0, x_1, x_2, x_3, x_4, x_5, x_6 \ge 0 \end{aligned}$$

$$\begin{aligned} &\text{Max } -5 - 2\cdot x_1 + 3\cdot x_2 - x_3 - x_5 = z \\ &\text{s.t.} \quad 2\cdot x_2 + x_3 + x_4 - x_5 = 9 \\ &\qquad x_0 = 5 + 2\cdot x_1 - 3\cdot x_2 + x_3 + x_5 \\ &\qquad -3\cdot x_1 + 4\cdot x_2 - 3\cdot x_3 - x_5 + x_6 = 4 \\ &\qquad x_0, x_1, x_2, x_3, x_4, x_5, x_6 \ge 0 \end{aligned}$$

Initialization – Example III

$$
\begin{aligned}
&\text{Max } -5-2\cdot x_1+3\cdot x_2-x_3-x_5=z\\
&\text{s.t.}\quad x_4=9-2\cdot x_2-x_3+x_5\\
&\qquad x_0=5+2\cdot x_1-3\cdot x_2+x_3+x_5\\
&\qquad x_6=4+3\cdot x_1-4\cdot x_2+3\cdot x_3+x_5\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0 \Rightarrow \text{Introduce } x_2
\end{aligned}
$$

$$
\begin{aligned}
&\text{Max } -5-2\cdot x_1+3\cdot x_2-x_3-x_5=z\\
&\text{s.t.}\quad x_4=9-2\cdot x_2-x_3+x_5 \Rightarrow x_4=9-2\cdot x_2\ge 0\Rightarrow x_2\le 9/2\\
&\qquad x_0=5+2\cdot x_1-3\cdot x_2+x_3+x_5 \Rightarrow x_0=5-3\cdot x_2\ge 0\Rightarrow x_2\le 5/3\\
&\qquad x_6=4+3\cdot x_1-4\cdot x_2+3\cdot x_3+x_5 \Rightarrow x_6=4-4\cdot x_2\ge 0\Rightarrow x_2\le 1 \hookleftarrow\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0
\end{aligned}
$$

Initialization – Example IV

$$
\begin{aligned}
&\text{Max } -5-2\cdot x_1+3\cdot\left(1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\right)-x_3-x_5=z\\
&\text{s.t.}\quad x_4=9-2\cdot\left(1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\right)-x_3+x_5\ge 0\\
&\qquad x_0=5+2\cdot x_1-3\cdot\left(1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\right)+x_3+x_5\\
&\qquad x_2=1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0
\end{aligned}
$$

$$
\begin{aligned}
&\text{Max } -2+1/4\cdot x_1+5/4\cdot x_3-1/4\cdot x_5-3/4\cdot x_6=z\\
&\text{s.t.}\quad x_4=7-3/2\cdot x_1-5/2\cdot x_3+1/2\cdot x_5+1/2\cdot x_6\\
&\qquad x_0=2-1/4\cdot x_1-5/4\cdot x_3+1/4\cdot x_5+3/4\cdot x_6\\
&\qquad x_2=1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0 \Rightarrow \text{Introduce } x_3
\end{aligned}
$$

Initialization – Example V

$$
\begin{aligned}
&\text{Max } -2+1/4\cdot x_1+5/4\cdot x_3-1/4\cdot x_5-3/4\cdot x_6=z\\
&\text{s.t.}\quad x_4=7-3/2\cdot x_1-5/2\cdot x_3+1/2\cdot x_5+1/2\cdot x_6\ge 0\Rightarrow 5/2\cdot x_3\le 7\Rightarrow x_3\le 14/5\\
&\qquad x_0=2-1/4\cdot x_1-5/4\cdot x_3+1/4\cdot x_5+3/4\cdot x_6\ge 0\Rightarrow x_3\le 8/5 \hookleftarrow\\
&\qquad x_2=1+3/4\cdot x_1+3/4\cdot x_3+1/4\cdot x_5-1/4\cdot x_6\ge 0\Rightarrow x_3\ge -4/3\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0
\end{aligned}
$$

$$
\begin{aligned}
&\text{Max } 0-x_0=z\\
&\text{s.t.}\quad x_4=3-x_1+2\cdot x_0-x_6\\
&\qquad x_3=8/5-1/5\cdot x_1-4/5\cdot x_0+1/5\cdot x_5+3/5\cdot x_6\\
&\qquad x_2=11/5+3/5\cdot x_1-3/5\cdot x_0+2/5\cdot x_5+1/5\cdot x_6\\
&\qquad x_0,x_1,x_2,x_3,x_4,x_5,x_6\ge 0 \;\Rightarrow x_1=0, x_2=11/5, x_3=8/5 \text{ is a feasible solution}
\end{aligned}
$$

Initialization – Example VI

Consequently, we resume with the Simplex applied to the following dictionary in order to solve the original problem

$$\begin{aligned}
\text{Max } x_1 - x_2 + x_3 &= \text{Max } x_1 - \left(11/5 + 3/5 \cdot x_1 + 2/5 \cdot x_5 + 1/5 \cdot x_6\right) + \left(8/5 - 1/5 \cdot x_1 + 1/5 \cdot x_5 + 3/5 \cdot x_6\right) \\
&= \text{Max } -3/5 + 1/5 \cdot x_1 - 1/5 \cdot x_5 + 2/5 \cdot x_6 = z
\end{aligned}$$

$$\begin{aligned}
\text{s.t.} \quad & x_4 = 3 - x_1 - x_6 \\
& x_3 = 8/5 - 1/5 \cdot x_1 + 1/5 \cdot x_5 + 3/5 \cdot x_6 \\
& x_2 = 11/5 + 3/5 \cdot x_1 + 2/5 \cdot x_5 + 1/5 \cdot x_6 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \geq 0
\end{aligned}$$

Cases to be distinguished

- Altogether, we have to deal with the following cases after solving the auxiliary problem
  1. x_0 is non-basic, i.e., we have the simple case where we can directly switch to the original problem with the feasible solution justly generated. x_0 is erased
  2. $x_0>0$ is basic, i.e., the original problem is not solvable at all because at least one constraint is violated
  3. $x_0=0$ is basic, i.e., this variable can be erased from the basis without affecting the solution quality. In order to make this obvious, consider the next-to-last step where the objective function becomes zero. Here, x_0 was decreased to zero. Consequently, in this step x_0 was a candidate for being erased. Hence, we can adjust this step accordingly

Pitfall: Iteration and Termination

- In each iteration, we erase one variable from the basis and replace it by another variable with a positive contribution to the objective function value
- However, this choice is ambiguous
  - There may be more than one non-basic candidate for entering the basis
  - Thus, we may choose the one with the largest improvement factor
  - If there is no candidate at all, the current solution is optimal (this point will be addressed thoroughly in Section 1.3)
- In addition, the choice of the leaving variable is ambiguous as well
  - If there is no candidate, the solution is unbounded, i.e., we can improve the solution arbitrarily
  - Otherwise, if there are several equal bounds, we have alternative choices. But, here we obtain a degenerate solution

Primal Degeneration – Example I

Consider the following dictionary

$$\text{Max } 2\cdot x_1 - x_2 + 8\cdot x_3 = z$$
$$\text{s.t.} \quad x_4 = 1 - 2\cdot x_3$$
$$x_5 = 3 - 2\cdot x_1 + 4\cdot x_2 - 6\cdot x_3$$
$$x_6 = 2 + x_1 + 3\cdot x_2 - 4\cdot x_3, \text{ with } x_1, x_2, x_3, x_5, x_6 \geq 0$$

We introduce x_3 into the basis

$$\text{Max } 2\cdot x_1 - x_2 + 8\cdot x_3 = z$$
$$\text{s.t.} \quad x_4 = 1 - 2\cdot x_3 \geq 0 \Rightarrow x_3 \leq \tfrac{1}{2}$$
$$x_5 = 3 - 2\cdot x_1 + 4\cdot x_2 - 6\cdot x_3 \geq 0 \Rightarrow x_3 \leq \tfrac{1}{2}$$
$$x_6 = 2 + x_1 + 3\cdot x_2 - 4\cdot x_3 \geq 0 \Rightarrow x_3 \leq \tfrac{1}{2}, \text{ with } x_1, x_2, x_3, x_5, x_6 \geq 0$$

Primal Degeneration – Example II

$$\text{Max } 4 + 2\cdot x_1 - x_2 - 4\cdot x_4 = z$$
$$\text{s.t.} \quad x_3 = \tfrac{1}{2} - \tfrac{1}{2}\cdot x_4$$
$$x_5 = 0 - 2\cdot x_1 + x_2$$
$$x_6 = 0 + x_1 - 3\cdot x_2 + 2\cdot x_3, \text{ with } x_1, x_2, x_3, x_5, x_6 \geq 0$$

- We observe that we have two basic variables (x_5, x_6) with value zero
- Although this is not harmful in its own right, it may have annoying side effects
- Specifically, primal degeneracy may cause cyclical calculations, i.e., in this case it prevents termination

Termination

- Termination may be prevented by cyclical calculations
- Note that cycling is only a rare phenomenon. Specifically, such kind of instances are hard to generate
- But, how does cycling become possible?
  - Primal degeneration may cause non-improving moves
  - Specifically, a basic variable with value zero leaves the basis and is replaced by a non-basic one
  - Note that a calculation that only comprises improving moves cannot cycle

Smallest subscript rule (rule of Bland)

- The rule proposed by Bland (Bland (1976)) is a relatively late development in the history of linear programming.
  - It is a very simple rule that allows for proving the termination of the simplex calculation
  - It bases on the so-called smallest subscript rule
- Pivoting strategy (**smallest subscript rule**): Choose the non-basic variable with the smallest index that has positive reduced costs to become a **basic variable**
- Choose the basic variable with the smallest index to become a **non-basic variable** from all equations that provides the minimal upper bound on the new basic variable

Termination of the Simplex algorithm

1.2.3 Theorem:

The Simplex Method terminates as long as the entering and leaving variables are selected by the smallest subscript rule

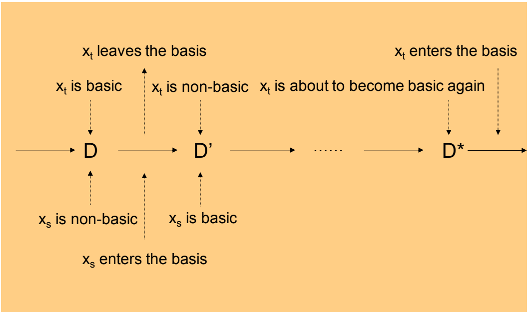
Proof by contradiction:

Assume that the opposite holds (i.e., there is a cycle with the smallest subscript rule applied) and show that this leads to a logical contradiction

Proof of Theorem 1.2.3 – Basics

- Let us assume that we have a cycle of dictionaries D_0-D_1-…-D_k, with D_0=D_k
- A variable is denoted as volatile if this variable is basic as well as non-basic throughout these dictionaries
  - Let x_t be the volatile variable with the largest subscript
  - D is the dictionary where x_t is basic and becomes non-basic in the next dictionary
  - x_s is non-basic in D and becomes basic in the next dictionary
  - Further along in the sequence, there is a dictionary D* where x_t becomes basic again

Proof of Theorem 1.2.3 – Illustration



Proof of Theorem 1.2.3 – Dictionaries D, D*

$$x_i = b_i - \sum_{j \notin B} a_{i,j} \cdot x_j \; (\forall i \in B)$$

––––––––––––––

$$z = v + \sum_{j \notin B} c_j \cdot x_j$$

Consider the calculation from D to D^*. Since we have a cycle, all these dictionaries are degenerate and the objective function value is kept unchanged. Hence, we obtain for the dictionary D^* the objective function $z = v + \sum_{j \notin B^*} c_j^* \cdot x_j$, with B^* as the basis of D^*.

Dictionaries D and D*

...

––––––––––––––––––––

$$z = v + \sum_{j} c_j^* \cdot x_j, \text{ with } c_j^* = 0, \forall j \in B^*$$

This dictionary D^* is generated by algebraic manipulations out of D. Therefore, each feasible solution of D is feasible for D^* and thus for each feasible solution of D it holds:

$$z = v + \sum_{j} c_j^* \cdot x_j$$

Proof of Theorem 1.2.3 – A new solution

We know that x_s enters basis B in dictionary D' (next to D)

By introducing y as the new value for x_s, we generate the following solution $\tilde{x}$ for each chosen value of y which is obviously feasible for D

$x_i = b_i - a_{i,s} \cdot y \quad (\forall i \in B)$

$x_s = y$

$x_i = 0, \forall i \notin B \wedge i \neq s$

We substitute all new basic variables of this solution in the objective function of the dictionaries D and D^* and obtain:

$$\Rightarrow z(\tilde{x}) = v + c_s \cdot y + \underbrace{\sum_{j \neq s} c_j \cdot x_j}_{\forall j \notin B \wedge j \neq s : x_j = 0} = v + c_s \cdot y + \underbrace{\sum_{j \in B} c_j \cdot x_j}_{\forall j \in B : c_j = 0} = v + c_s \cdot y$$

$$= v + c_s^* \cdot y + \underbrace{\sum_{j \neq s} c_j^* \cdot x_j}_{\forall j \notin B \wedge j \neq s : x_j = 0} = v + c_s^* \cdot y + \sum_{j \in B} c_j^* \cdot x_j$$

Proof of Theorem 1.2.3 – Transforming

Therefore, we obtain:

$$v + c_s \cdot y = v + c_s^* \cdot y + \sum_{j \in B} c_j^* \cdot x_j, \text{ with } c_j^* = 0, \forall j \in B^*$$

$$\Leftrightarrow c_s \cdot y = c_s^* \cdot y + \sum_{j \in B} c_j^* \cdot x_j$$

$$\Leftrightarrow c_s \cdot y = c_s^* \cdot y + \sum_{j \in B} c_j^* \cdot \left(b_j - a_{j,s} \cdot y\right)$$

$$\Leftrightarrow \left(c_s - c_s^*\right) \cdot y = \sum_{j \in B} c_j^* \cdot b_j - \sum_{j \in B} c_j^* \cdot a_{j,s} \cdot y$$

$$\Leftrightarrow \left(c_s - c_s^* + \sum_{j \in B} c_j^* \cdot a_{j,s}\right) \cdot y = \underbrace{\sum_{j \in B} c_j^* \cdot b_j}_{\text{Is a constant independent of } y}$$

Proof of Theorem 1.2.3 – Conclusion

$$\left(c_s - c_s^* + \sum_{j \in B} c_j^* \cdot a_{j,s}\right) \cdot y = \underbrace{\sum_{j \in B} c_j^* \cdot b_j}_{\text{Is independent of } y \Rightarrow \text{constant for all chosen values of } y}$$

$\Rightarrow$ Consequently, both sides have the value zero.

Thus, $c_s - c_s^* + \sum_{j \in B} c_j^* \cdot a_{j,s} = 0$

Since x_s enters the basis in D, it holds $c_s > 0$. Thus, x_s is volatile and therefore $s < t$. Since x_s not enters the basis in D^*, it holds $c_s^* \leq 0$.

Proof of Theorem 1.2.3 – Conclusion

Since $c_s - c_s^* + \sum_{j \in B} c_j^* \cdot a_{j,s} = 0 \wedge c_s > 0 \wedge c_s^* \leq 0$

$\Rightarrow c_s - c_s^* > 0 \Rightarrow \exists r \in B : c_r^* \cdot a_{r,s} < 0 \Rightarrow c_r^* \neq 0$

Since $r \in B$, x_r is basic in D.

Since $c_r^* \neq 0$, we know that $r \notin B^* \Rightarrow x_r$ is volatile.

Note that $r \neq t$. Since t enters in D^*, we have $c_t^* > 0$. In addition, t is leaving in D and thus, we conduct the following transformation

$$x_t = b_t - a_{t,s} \cdot x_s + \sum_{j \notin B \wedge j \neq s} a_{t,j} \cdot x_j \Leftrightarrow x_s = \frac{b_t}{\underbrace{a_{t,s}}_{\geq 0}} - \frac{x_t}{a_{t,s}} + \sum_{j \notin B \wedge j \neq s} \frac{a_{t,j}}{a_{t,s}} \cdot x_j$$

$\Rightarrow a_{t,s} > 0 \Rightarrow c_r^* \cdot a_{r,s} < 0 \Rightarrow c_r^* \cdot a_{r,s} \neq c_t^* \cdot a_{t,s} \geq 0 \Rightarrow t \neq r$

Proof of Theorem 1.2.3 – Conclusion

Consequently, $r < t$, but x_r has not entered in D^*.

Although x_r is not basic in D^*, x_t has entered in D^*.

$\Rightarrow c_r^* \leq 0$, actually, $c_r^* < 0$ since $c_r^* \cdot a_{r,s} < 0 \Rightarrow a_{r,s} > 0$

Since all solutions between D and D^* are degenerate, and x_r and x_t are volatile, we have in all solutions $x_r = x_t = 0$.

$\Rightarrow b_r = 0 \wedge b_t = 0$ in dictionary D and both $(x_t$ and $x_r)$ were candidates for leaving the basis B.

But, we choose x_t although $t > r$. This violates the Smallest Subscript rule and is therefore a contradiction.

This completes the proof.

What we know so far…

- We have learned how to solve general LPs by applying the Simplex procedure that explores a sequence of basic solutions
- We have seen that under certain circumstances (i.e., if we make use of a specific subscript rule) this algorithm always terminates
- We have learned to deal with problems where an initial solution is not directly available
  - In order to do this, we have generated the Two-Phase Method
  - It terminates either with an initial solution or with the cognition that the problem is not solvable at all
- In what follows, we will show why it is sufficient to concentrate the search process to basic solutions

1.3 The Geometry of the solution space

- In what follows, we have to do a little bit mathematics
- By doing so, we get (hopefully!) some insights into the problem structure
  - First of all, we focus on convexity
  - Then, we learn about the solution space that it is sufficient to focus our search on the corner points
- Therefore, let the solution space *P* be given as defined above in the standard form
- Convexity is a very convenient attribute of solution spaces. Note that it causes – among other advantages – that each local optimum is also a global optimum

Minimization problem

- In what follows, we consider minimization problems
- I.e., unless it is indicated differently, we consider minimization problems of the following structure

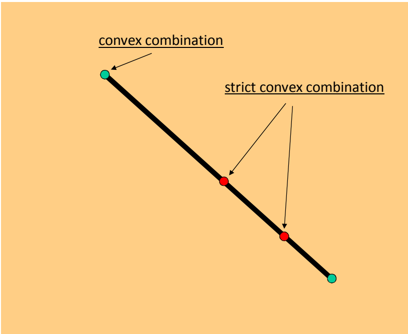
$$\text{Minimize } z(x) = c^T \cdot x, \text{ with } c \in IR^n, x \in P,$$
$$\text{and } P = \{x \in IR^n \mid x \geq 0 \text{ and } A \cdot x = b\}$$

Convex combinations

1.3.1 Definition
(Convex combination)

Let $a^1,...,a^k \in IR^n$ and $\alpha_1,...,\alpha_k \in IR, \alpha_i \geq 0$. Then, $\sum_{i=1}^{k} \alpha_i \cdot a^i$ is denoted as a non-negative linear combination and as a convex combination if additionally $\sum_{i=1}^{k} \alpha_i = 1$. If $\forall i \in \{1,...,n\} : \alpha_i > 0$, then $\sum_{i=1}^{k} \alpha_i \cdot a^i$ is a strict convex combination.

Illustration – Convex combinations



Convex sets

1.3.2 Definition

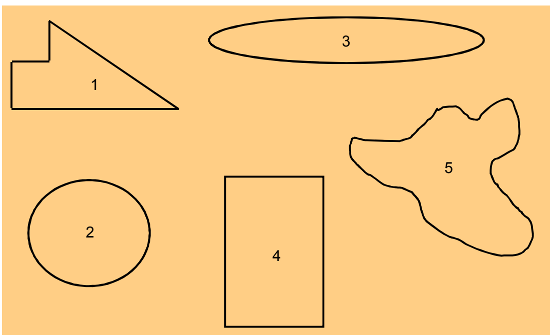
$C(a^1,\dots,a^k)$ denotes the set of all convex combinations of $a^1,\dots,a^k$

if $k=2$, $C(a^1,a^2)$ denotes the direct connection between a^1 and a^2.

1.3.3 Definition

A set $S \subseteq IR^n$ is **convex** if it contains all convex combinations of pairs of points $x,y \in S$, i.e., $C(x,y) \subseteq S,\ \forall x,y \in S$.

Examples – Who is convex, who not?



Solution space of LPs is convex

Consider two elements $y, z \in P = \{x \mid x \in IR^n \wedge x \geq 0 \wedge A \cdot x \leq b\}$.

Then, consider λ with $0 \leq \lambda \leq 1$ and $\lambda \cdot y + (1-\lambda) \cdot z = \tilde{x}$.

Obviously, $\tilde{x} \geq 0$.

$A \cdot \tilde{x} = A \cdot (\lambda \cdot y + (1-\lambda) \cdot z) = \lambda \cdot (A \cdot y) + (1-\lambda) \cdot (A \cdot z)$

$\leq \lambda \cdot b + (1-\lambda) \cdot b = b$

$\Rightarrow \tilde{x} \in P = \{x \mid x \in IR^n \wedge x \geq 0 \wedge A \cdot x \leq b\}$.

Additionally, since $0 \leq \lambda \leq 1$ and $y, z \geq 0$,

it holds $\lambda \cdot y + (1-\lambda) \cdot z \geq 0$.

Intersection

1.3.4 Lemma

The intersection of any number of convex sets S_i is convex.

Proof:

Let us consider two elements of the set $\cap S_i$. Then, each convex combination belongs to every set S_i. Consequently, it also belongs to the intersection of all sets $\cap S_i$. This completes the proof.

Linear combinations in convex sets

1.3.5 Lemma

Let S be a convex set and $a^1, \ldots, a^k \in S$, then $C(a^1, \ldots, a^k) \subseteq S$.

We provide this proof by induction

- Beginning of induction:
  - Show that the lemma holds for k=2. This is obviously trivial.
- Induction step: k>2
  - The proposition is held for all values up to k-1
  - Let us now consider k

Proof of Lemma 1.3.5

Let us now consider: $\sum_{i=1}^{k}\alpha_i\cdot a^i$, $\sum_{i=1}^{k}\alpha_i=1\Leftrightarrow 1-\alpha_k=\sum_{i=1}^{k-1}\alpha_i$

Thus, by assumption of induction:

$$\frac{1}{1-\alpha_k}\cdot\sum_{i=1}^{k-1}\alpha_i\cdot a^i=\sum_{i=1}^{k-1}\frac{\alpha_i}{1-\alpha_k}\cdot a^i=\tilde{a}\in S$$

Consequently, it holds: $(1-\alpha_k)\cdot\tilde{a}+\alpha_k\cdot a^k\in S$

We get: $(1-\alpha_k)\cdot\tilde{a}+\alpha_k\cdot a^k=(1-\alpha_k)\cdot\left(\sum_{i=1}^{k-1}\frac{\alpha_i}{1-\alpha_k}\cdot a^i\right)+\alpha_k\cdot a^k=$

$\sum_{i=1}^{k-1}\alpha_i\cdot a^i+\alpha_k\cdot a^k=\sum_{i=1}^{k}\alpha_i\cdot a^i$

Hyperplanes

Let $a\in IR^n\setminus\{0\}$ and $\alpha\in IR$. Then, $H=\{x\in IR^n\mid a^T\cdot x=\alpha\}$ is denoted as a hyperplane.

Hyperplanes are obviously convex. This can be easily shown:

Let $x^1,x^2\in H, 0\le\lambda\le 1$.
Let us now consider:
$a^T\cdot(\lambda\cdot x^1+(1-\lambda)\cdot x^2)=\lambda\cdot a^T\cdot x^1+(1-\lambda)\cdot a^T\cdot x^2$
$=\lambda\cdot\alpha+(1-\lambda)\cdot\alpha=\alpha.$

Half spaces

Let $a\in IR^n\setminus\{0\}$ and $\alpha\in IR$. Then, $H^{\ge}=\{x\in IR^n\mid a^T\cdot x\ge\alpha\}$ is denoted as a half space.

Half spaces are obviously convex. This can be easily shown as follows:

Let $x^1,x^2\in H^{\ge}, 0\le\lambda\le 1$.
Let us consider:
$a^T\cdot(\lambda\cdot x^1+(1-\lambda)\cdot x^2)=\lambda\cdot a^T\cdot x^1+(1-\lambda)\cdot a^T\cdot x^2$
$\ge\lambda\cdot\alpha+(1-\lambda)\cdot\alpha=\alpha.$

Observation

- A hyperplane in an n-dimensional space has the dimension n-1
- A hyperplane defines two separated half spaces, i.e., it divides the space into two parts

In the IR^n, the hyperplane $H = \{x \in IR^n | a^T \cdot x = \alpha\}$ determines the two half spaces $H_1^{\geq} = \{x \in IR^n | a^T \cdot x \geq \alpha\}$ and $H_2^{\geq} = \{x \in IR^n | -a^T \cdot x \geq -\alpha\}$

Convex hull

1.3.6 Definition

$$CH(M) = \bigcup \{C(a^1,...,a^k) \mid a^1,...,a^k \in M, k \in IN\}$$

is denoted as the convex hull to $M \subseteq IR^n$.

The set *CH(M)* is convex since a convex combination of two convex combinations of elements of set *M* is again a convex combination of elements of set *M*

1.3.7 Observation

It holds: $CH(M) = \bigcap \{K \mid M \subseteq K \wedge K \text{ convex}\}$, i.e., $CH(M)$ is the smallest convex set that contains M.

Proof of Observation 1.3.7

Let $SCH(M)$ be the smallest convex set that contains M

1. $SCH(M) \subseteq CH(M)$:

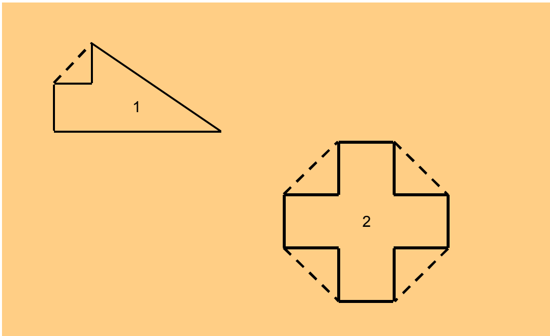
This is correct since $SCH(M)$ is the smallest convex set that contains M, $CH(M)$ is convex, and it holds that $M \subseteq CH(M)$.

2. $CH(M) \subseteq SCH(M)$:

Consider $x \in CH(M)$. Then, we know $x = \sum_{i=1}^{k} \alpha_i \cdot a^i$, with $a^1,...,a^k \in M$ and $M \subseteq SCH(M)$.

By applying Lemma 1.3.5 and the convexity of set $SCH(M)$, we obtain $x \in SCH(M)$.

Illustration – convex hull

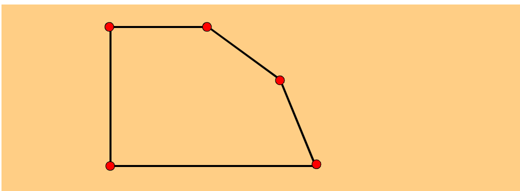


Extreme points

1.3.8 Definition

A point $x \in K, K \subseteq IR^n$ convex is denoted as an extreme point of K if it is not defineable by a strict convex combination.

Let $\varepsilon(K)$ be the set of all extreme points of K.



Important attributes of extreme points

1.3.9 Observation

The following propositions are equivalent

1. x^0 is an extreme point
2. $\forall a,b \in K, x^0 \in C(a,b) = \overline{ab} \Rightarrow x^0 = a \vee x^0 = b$
3. $\forall y \in IR^n \setminus \{0\} : x^0 + y \notin K \vee x^0 - y \notin K$
4. $K \setminus \{x^0\}$ is convex

Proof of Observation 1.3.9

$1 \Rightarrow 2$:

trivial, since if $x^0 \in C(a,b)$, i.e., if $x^0 = \bar{a} \cdot a + \bar{b} \cdot b, \bar{a}, \bar{b} \in IR$,
we can conclude that this convex combination is not
strict. Thus, $\bar{a} = 0 \vee \bar{b} = 0$.

$2 \Rightarrow 3$:

Let us assume (2) holds and $x^0 - y \in K \wedge x^0 + y \in K, y \in IR^n \setminus \{0\}$.

Consider $\lambda \cdot (x^0 - y) + (1-\lambda) \cdot (x^0 + y), 0 \le \lambda \le 1$

$\lambda \cdot x^0 - \lambda \cdot y + x^0 + y - \lambda \cdot x^0 - \lambda \cdot y = x^0 + y - 2 \cdot \lambda \cdot y$

$= x^0 + (1 - 2 \cdot \lambda) \cdot y$.

Let $\lambda = 0.5 \wedge a = x^0 - y \wedge b = x^0 + y \Rightarrow x^0 = 0.5 \cdot a + 0.5 \cdot b$

This contradicts (2)

Proof of Observation 1.3.9

$3 \Rightarrow 4$:

Consider $a, b \in K \setminus \{x^0\}$ and $\lambda \cdot a + (1-\lambda) \cdot b$, with $0 \le \lambda \le 1$. We have to distinguish two cases:

1. $\lambda \cdot a + (1-\lambda) \cdot b \ne x^0$, then we know $\lambda \cdot a + (1-\lambda) \cdot b \in K \setminus \{x^0\}$ due to the convexity of K
2. $\lambda \cdot a + (1-\lambda) \cdot b = x^0$, then we set $y = \begin{cases} a - x^0 & \text{if } \lambda \ge 0{,}5 \\ b - x^0 & \text{otherwise} \end{cases}$. In both cases we have $y \ne 0$ if $0 < \lambda < 1$.

We consider: $x^0 + y = \begin{cases} a & \text{if } \lambda \ge 0{,}5 \\ b & \text{otherwise} \end{cases} \wedge x^0 - y = \begin{cases} -a + 2x^0 = (2\lambda - 1) \cdot a + (2 - 2\lambda) \cdot b & \text{if } \lambda \ge 0{,}5 \\ -b + 2x^0 = 2\lambda \cdot a + (1 - 2\lambda) \cdot b & \text{otherwise} \end{cases}$

Since all weights are positive and $(2\lambda - 1) + (2 - 2\lambda) = 2\lambda + (1 - 2\lambda) = 1$, we conclude that
$x^0 + y \in K \wedge x^0 - y \in K$. This contradicts (3). Hence, we obtain $\lambda \cdot a + (1-\lambda) \cdot b \in K \setminus \{x^0\}$

$4 \Rightarrow 1$:

Let K and $K \setminus \{x^0\}$ be convex. We assume that x^0 is not an extreme point,
i.e., $\exists a, b \in K : \exists \bar{a}, \bar{b} \ge 0 \in IR : \bar{a} \cdot a + \bar{b} \cdot b = x^0 \wedge \bar{a} + \bar{b} = 1 \wedge a \ne x^0 \wedge b \ne x^0$

But then it holds:

$a, b \in K \setminus \{x^0\} \wedge \bar{a} \cdot a + \bar{b} \cdot b = x^0 \notin K \setminus \{x^0\} \wedge \bar{a} + \bar{b} = 1 \wedge \bar{a}, \bar{b} \ge 0$

This contradicts the convexity of $K \setminus \{x^0\}$

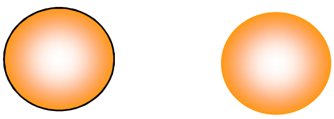
Examples of extreme points

$$1. \varepsilon(\{x \in IR^n \mid x \ge 0\}) = \{0\}$$

$$2. \varepsilon(H^{\ge}) = \varnothing$$

$$3. \varepsilon(\{x \in IR^n \mid \|x\|_2 \le 1\}) = \{x \in IR^n \mid \|x\|_2 = 1\}$$

$$4. \varepsilon(\{x \in IR^n \mid \|x\|_2 < 1\}) = \varnothing$$



Bounded sets

Let $|\cdot|: \mathbb{R}^n \to \mathbb{R}$, $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sqrt{(x_1^2 + \ldots + x_n^2)}$

This mapping is denoted as the **Euclidean norm** and satisfies the **norm properties**, i.e.,

- $|x| = 0 \Leftrightarrow x = 0$
- $|\lambda x| = |\lambda||x| \quad \forall\, \lambda \in \mathbb{R}$ and $\forall x \in \mathbb{R}^n$
- $|x + y| \leq |x| + |y| \quad \forall\, x \in \mathbb{R}^n$ and $\forall\, y \in \mathbb{R}^n$.

The mapping $|\cdot|: \mathbb{R} \to \mathbb{R}$, $\quad x \mapsto \begin{cases} x, & if\ x \geq 0 \\ -x, & else \end{cases}$
is denoted as the **absolute value function**.

A set $M \subseteq \mathbb{R}^n$ is called bounded, if there is an arbitrary, but fixed positive real number r, such that

$$|x| \leq r \quad \forall\, x \in \mathrm{M}.$$

Convex polyhedron

1.3.10 Definition

A convex polyhedron is an intersection of a finite number of half spaces, i.e.,

$$P = \bigcap_{i=1}^{m} H^{\geq} = \{x \mid A \cdot x \geq b\}$$

If $P \neq \varnothing \wedge |P| < \infty$ (bounded), then P is a convex polytope.

A hyperplane $H = \{x \mid a^T \cdot x = \alpha \wedge a \in IR^n\}$ is significant if

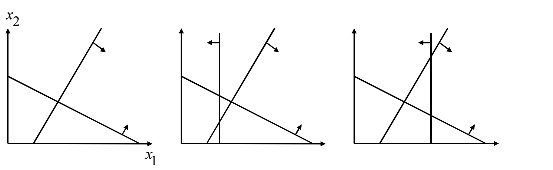
$H \cap P \neq \varnothing$ and $P \subseteq H^{\geq} = \{x \mid a^T \cdot x \geq \alpha \wedge a \in IR^n\}$ for one half space.

If $H \cap P = \{x_0\}$, then x_0 is denoted as a corner point.

If $H \cap P = \overline{ab}$, then $\overline{ab}$ is denoted as an edge.

Convex polyhedron - Examples

Consider a two-dimensional solution space P

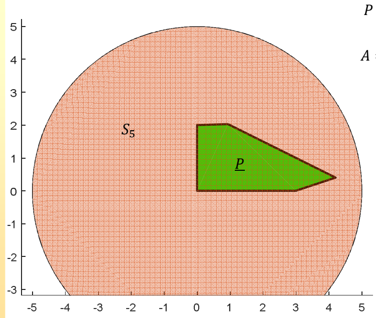


unbounded polyhedron

empty polyhedron

polytope, which is a bounded and non-empty polyhedron

Polytope – Example



$P = \{x \in \mathbb{R}^n | Ax \geq b \wedge x \geq 0\}$

$A = \begin{pmatrix} -1 & -2 \\ -2 & -6 \\ -1 & 3 \end{pmatrix} \quad b = \begin{pmatrix} -5 \\ -12 \\ -3 \end{pmatrix}$

$S_5 = \{x \in \mathbb{R}^n | \ |x| \leq 5\}$

Since $P \subseteq S_5$, the polyhedron P is bounded.

Conclusion

1.3.11 Theorem

Let P be a convex polyhedron. Then, all corner points are extreme points.

Proof of Theorem 1.3.11

Let x^0 be a corner point of P. In addition, let

$H=\{x\in IR^n \mid a^T\cdot x=\alpha\}$ a significant hyperplane with $P\cap H=\{x^0\}$.

We now make use of Observation 1.3.9 and consider $y\in IR^n$

with $x^0+y\in P \wedge x^0-y\in P \Rightarrow x^0+y\in H^{\geq}\wedge x^0-y\in H^{\geq}$

Thus, it holds:

$a^T\cdot(x^0+y)=a^T\cdot x^0+a^T\cdot y\geq\alpha \wedge a^T\cdot(x^0-y)=a^T\cdot x^0-a^T\cdot y\geq\alpha$

and since $P\cap H=\{x^0\}$

$\alpha=a^T\cdot x^0\Rightarrow a^T\cdot y=0$

Therefore, it holds: $a^T\cdot(x^0+y)=a^T\cdot x^0+a^T\cdot y=\alpha$

$\Rightarrow x^0+y\in H\wedge x^0+y\in P\Rightarrow x^0+y\in P\cap H=\{x^0\}\Rightarrow y=0$

Basic solutions

- In what follows, we are able to bridge the gap to basic solutions, i.e., it is now possible to provide a definition of basic solutions
  - Recall that these are just the solutions the Simplex Algorithm is focusing on
  - We feel that basic solutions are just corner points or extreme points of the solution space

Basic solutions and corner points

1.3.12 Definition

Let $B:\{1,2,...,m\}\to\{1,...,n\}, N:\{1,...,n-m\}\to\{1,...,n\}$ injective with $B(\{1,...,m\})\cup N(\{1,...,n-m\})=\{1,...,n\}$ and $A_B=\left(a^{B(1)},...,a^{B(m)}\right)\in IR^{m\times m},\ a^{B(1)},...,a^{B(m)}\in IR^m$ an invertible matrix.

Let $x_B=A_B^{-1}\cdot b$ and $x_N=0$. Then, $x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$ is denoted as a basic feasible solution (bfs) of $A\cdot x=b$, if $x\geq 0$.

Observations

Obviously, it holds:

1. $A\cdot x=(A_B,A_N)\cdot\begin{pmatrix}x_B\\x_N\end{pmatrix}=A_B\cdot x_B+A_N\cdot x_N$
$=A_B\cdot\left(A_B^{-1}\cdot b\right)+A_N\cdot x_N$
$=\left(A_B\cdot A_B^{-1}\right)\cdot b+A_N\cdot x_N=b+0=b$

2. A_B is invertible $\Leftrightarrow\left\{a^{B(1)},...,a^{B(m)}\right\}$ is a base of IR^m

Conclusions

1.3.13 Theorem

Let $A \in IR^{m \times n}$ with $rank(A) = m \le n$ and let $b \in IR^m$.

Furthermore, let $P = \{x \in IR^n \mid x \ge 0 \text{ and } A \cdot x = b\}$,

for $x^0 \in P$. The following propositions are equivalent:

1. x^0 is an extreme point of P
2. $\{a^j \mid x_j^0 > 0\}$ are linearly independent
3. x^0 is a basic feasible solution (bfs)
4. x^0 is a corner point of P

Proof of Theorem 1.3.13

$(1) \Rightarrow (2)$:

Let us consider $x^0 = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$. We sort the entries so that it holds:

$x_1, \ldots, x_r > 0$ and $x_{r+1}, \ldots, x_n = 0$

Case 1 : $r = 0$

$\Rightarrow x^0 = 0$. Then, $\{a^j \mid x_j^0 > 0\} = \varnothing$. By definition, this set is a set of linearly independent vectors.

Proof of Theorem 1.3.13

Case 2 : $r > 0$

We introduce $\alpha_1, \ldots, \alpha_r \in IR : \sum_{i=1}^{r} \alpha_i \cdot a^i = 0$

Now, consider $y = (\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)^T \in IR^n$

We compute $x^0 + y, x^0 - y$, i.e., $x^0 \pm y$

Consider $A \cdot (x^0 \pm y) = A \cdot x^0 \pm A \cdot y = b + 0 = b$

Thus, it holds: $x^0 + y, x^0 - y \in P$

Proof of Theorem 1.3.13

Thus, it holds: $x^0 + y, x^0 - y \in P$

Since x^0 is an extreme point and we are making use of Observation 1.3.9, we can conclude that

$y = 0 \Rightarrow a^1,...,a^r$ are linearly independent.

Proof of Theorem 1.3.13

$(2) \Rightarrow (3)$

We assume that $a^1,...,a^r$ are linearly independent.

If $r < m$ since $rank(A) = m$, we altogether have m linearly independent columns in A. Let $li(A)$ be this set. Then, w.l.o.g., we can assume $a^1,...,a^r \in li(A)$. We define B and N accordingly. If $r = m$, we define $li(A) = \{a^1,...,a^r\}$. Then, it holds:

$A_B \in IR^{m \times m}$ is invertible and it holds:

$A_B^{-1} \cdot b = A_B^{-1} \cdot A \cdot x^0 = A_B^{-1} \cdot A_B \cdot x_B^0 + A_N \cdot x_N^0 = x_B^0$

Since $x^0 \in P$, we know $x^0 \geq 0$ and, therefore, x^0 is the basic feasible solution.

Proof of Theorem 1.3.13

$(3) \Rightarrow (4)$

We assume that $x^0 = \begin{pmatrix} x_B^0 \\ x_N^0 \end{pmatrix}$ is the basic feasible solution.

Let $a = \left(\underbrace{0,...,0}_{m \text{ elements}}, \underbrace{1,...,1}_{n-m \text{ elements}} \right)$. Furthermore, let

$H = \{x \in IR^n \mid a^T \cdot x = 0\}$.

We can conclude the following:

1. Let $x \in P : a^T \cdot x = a_B^T \cdot x_B + a_N^T \cdot x_N = a_N^T \cdot x_N \geq 0$ since $x \geq 0$

$\Rightarrow x \in H^{\geq} \Rightarrow P \subseteq H^{\geq}$

Proof of Theorem 1.3.13

2. Consider $a^T \cdot x^0 = a_B^T \cdot x_B^0 + a_N^T \cdot x_N^0 = 0 \Rightarrow x^0 \in H$

Thus, $x^0 \in H \cap P$

3. We consider $y = \begin{pmatrix} y_B \\ y_N \end{pmatrix} \in H \cap P \Rightarrow a_N^T \cdot y_N = 0$. Since $y \geq 0$, it holds $y_N = 0$

Additionally, it holds:

$b = A \cdot y = A_B \cdot y_B + A_N \cdot y_N = A_B \cdot y_B \Leftrightarrow y_B = A_B^{-1} \cdot b = x_B^0$

$\Rightarrow y = x^0 \Rightarrow H \cap P = \{x^0\}$

Consequently, x^0 is a corner point

Proof of Theorem 1.3.13

$(4) \Rightarrow (1)$

Assuming x^0 is a corner point.

$$P = \{x \in IR^n \mid x \geq 0 \wedge A \cdot x = b\} = \left\{ x \in IR^n \mid \begin{pmatrix} A \\ -A \\ E \end{pmatrix} \cdot x \geq \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix} \right\}$$

Obviously, P is a convex polyhedron (Definition 1.3.10). Since $x^0 \in P$ is the corner point and all corner points in P are extremal points, x^0 is an extremal point.

Conclusions

1.3.14 Observation

1. $0 \in P \Rightarrow 0 \in \varepsilon(P)$

This is true due to the fact that if $0 \in P \Rightarrow 0$ fulfills Restriction (2) of Theorem 1.3.13 $\Rightarrow 0$ is a corner point

2. $x \in \varepsilon(P) \Rightarrow \left|\{x_i \mid x_i > 0\}\right| \leq m$

In order to conceive this proposition, we again make use of Theorem 1.3.13. Owing to the fact that $\{a^j \mid x_j^0 > 0\}$ are linearly independent and $rank(A) = m$, the proposition immediately follows

3. $|\varepsilon(P)| \leq \binom{n}{m} = \dfrac{n!}{m!(m-n)!}$ Binomial coefficient

Degeneration

1.3.15 Definition

A basic feasible solution $x \in \varepsilon(P)$ is denoted as degenerated if $\left|\{x_i \mid x_i > 0\}\right| < m$.

1.3.16 Observation

The respective set of base vectors is unambiguously defined for each non-degenerate basic feasible solution $x \in \varepsilon(P)$.

The structure of the solution space

- In what follows, we are able to provide a very compact and fundamental definition for the solution space *P*
- For this purpose, however, we have to distinguish if the solution space is
  - bounded or
  - unbounded
- Consequently, if the latter case applies, an infinite number of new elements of *P* can be generated iteratively

Preliminary Definitions

In what follows, we consider an LP with a solution space $P = \left\{x \in IR^n \mid x \geq 0 \wedge A \cdot x = b\right\}$

1.3.17 Definition

Let $D(P) = \left\{y \in IR^n \mid \forall x \in P : \forall \lambda > 0 : x + \lambda \cdot y \in P\right\}$, for $P \neq \varnothing$.

1.3.18 Lemma

$D(P) = \left\{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\right\}$

Proof of Lemma 1.3.18

1. $D(P) \subseteq \{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\}$

Let $\tilde{y} \in D(P) = \{y \in IR^n \mid \forall x \in P : \forall \lambda > 0 : x + \lambda \cdot y \in P\}$,
for $P \neq \emptyset$.
Then, it holds: $\forall x \in P : \forall \lambda > 0 : x + \lambda \cdot \tilde{y} \in P$
$\Rightarrow \forall x \in P : \forall \lambda > 0 :$
$A \cdot (x + \lambda \cdot \tilde{y}) = A \cdot x + \lambda \cdot A \cdot \tilde{y} = b + \lambda \cdot A \cdot \tilde{y} = b \Leftrightarrow \lambda \cdot A \cdot \tilde{y} = 0$
$\Leftrightarrow A \cdot \tilde{y} = 0.$
In addition, we know that $\forall x \in P : \forall \lambda > 0 : x + \lambda \cdot \tilde{y} \geq 0$
$\Rightarrow \tilde{y} \geq 0 \Rightarrow \tilde{y} \in \{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\}$
$\Rightarrow D(P) \subseteq \{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\}$

Proof of Lemma 1.3.18

2. $\{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\} \subseteq D(P)$

Let $\tilde{y} \in \{y \in IR^n \mid y \geq 0 \wedge A \cdot y = 0\}$.

Consider $x + \lambda \cdot \tilde{y}, x \in P \wedge \lambda > 0$. Then, it holds:
$A \cdot (x + \lambda \cdot \tilde{y}) = A \cdot x + \lambda \cdot A \cdot \tilde{y} = b + \lambda \cdot 0 = b$
$\wedge x + \lambda \cdot y \geq 0$ since $x \geq 0 \wedge \lambda \cdot y \geq 0 \Rightarrow \tilde{y} \in D(P)$.

Final result – The solution space

1.3.19 Theorem

$$P = \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$$

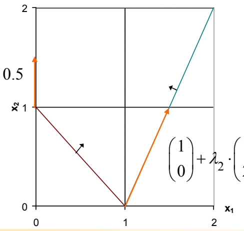
A non-empty polyhedron P is represented by a convex combination y of its **extreme (corner) points** $\varepsilon(P)$ and by $z \in D(P)$. If $z \neq 0$, then z is called a **ray**.

Theorem 1.3.19 – Example I

Let $P=\left\{\begin{pmatrix}x_1\\x_2\end{pmatrix}\ge 0 \,\middle|\, \begin{pmatrix}-1&-1\\2&-1\end{pmatrix}\cdot\begin{pmatrix}x_1\\x_2\end{pmatrix}\le\begin{pmatrix}-1\\2\end{pmatrix}\right\}$. It is $\varepsilon(P)=\left\{\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\0\end{pmatrix}\right\}$ and

$D(P)=\left\{\lambda_1\cdot\begin{pmatrix}0\\1\end{pmatrix}+\lambda_2\cdot\begin{pmatrix}1\\2\end{pmatrix}\,\middle|\,\lambda_1,\lambda_2\ge 0\right\}\cup\left\{\begin{pmatrix}0\\0\end{pmatrix}\right\}$ that is an infinite set.

$\begin{pmatrix}0\\1\end{pmatrix}+\lambda_1\cdot\begin{pmatrix}0\\1\end{pmatrix}$ with $\lambda_1=0.5$



$\begin{pmatrix}1\\0\end{pmatrix}+\lambda_2\cdot\begin{pmatrix}1\\2\end{pmatrix}$ with $\lambda_2=0.5$

Theorem 1.3.19 – Example II

Since $A\cdot x\le b$ in P, Lemma 1.3.18 becomes $D(P)=\left\{y\in IR^n \mid y\ge 0\wedge A\cdot y\le 0\right\}$.

Note that this is satisfied by all $y\in D(P)$. Theorem 1.3.19 states that P is equivalently written as

$$P=\left\{x\in IR^n \mid x=\alpha\cdot\begin{pmatrix}0\\1\end{pmatrix}+(1-\alpha)\cdot\begin{pmatrix}1\\0\end{pmatrix}+\lambda_1\cdot\begin{pmatrix}0\\1\end{pmatrix}+\lambda_2\cdot\begin{pmatrix}1\\2\end{pmatrix}\wedge\lambda_1,\lambda_2\ge 0\right\}$$

Note that $\begin{pmatrix}0\\0\end{pmatrix}\in D(P)$ is omitted in the representation of P because it is neutral to x.

Proof of Theorem 1.3.19

1. $\left\{x\in IR^n \mid x=y+z\wedge y\in C(\varepsilon(P))\wedge z\in D(P)\right\}\subseteq P$

Let $x\in\left\{x\in IR^n \mid x=y+\lambda\cdot z\wedge y\in C(\varepsilon(P))\wedge z\in D(P)\right\}$

$\Rightarrow x=y+z\wedge y\in C(\varepsilon(P))\wedge z\in D(P)$

$\Rightarrow A\cdot(y+z)=A\cdot y+A\cdot z=A\cdot\left(\sum_{k=1}^{K}\zeta_k\cdot y_k\right)+A\cdot z,$

with: $y_1,\ldots,y_K\in\varepsilon(P)\wedge\sum_{k=1}^{K}\zeta_k=1$

$=\sum_{k=1}^{K}\zeta_k\cdot\left(\underbrace{A\cdot y_k}_{=b}\right)+A\cdot z=\underbrace{\sum_{k=1}^{K}\zeta_k}_{=1}\cdot b+\underbrace{A\cdot z}_{=0}=b+0=b$

Additionally, it holds: $x=y+z\ge 0$

Proof of Theorem 1.3.19

2. $P \subseteq \{x \in IR^n \mid x = y + \lambda \cdot z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

The proof is conducted by induction by $n_x = \left|\{j \mid x_j > 0\}\right|$

We show : $\forall l \in IN: \forall x \in P: l = n_x: \exists \lambda_1,\ldots,\lambda_k (\geq 0) \in IR:$

$$x = \sum_{i=1}^{k} \lambda_i \cdot x^i + y \wedge \sum_{i=1}^{k} \lambda_i = 1 \wedge x^1,\ldots,x^k \in \varepsilon(P) \wedge y \in D(P)$$

Proof of Theorem 1.3.19

We commence with $n_x = 0 \Rightarrow x = 0 \Rightarrow x \in \varepsilon(P) \Rightarrow$

$x \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

Now, we assume that the proposition holds for all $x \in P$ with $n_x < l$. Consider $x \in P$ with $n_x = l$. Obviously, if $x \in \varepsilon(P)$, the proposition immediately follows.

Proof of Theorem 1.3.19

Consequently, we assume $x \notin \varepsilon(P)$.

Since $x \notin \varepsilon(P)$, it holds: $\exists y(\neq 0) \in IR^n : x + y \in P$

$\wedge x - y \in P \Rightarrow$ Therefore, we can assume $|y_j| \leq x_j, \forall j$

We compute

$A \cdot y = A \cdot y + A \cdot x - A \cdot x = A \cdot (x + y) - A \cdot x$

$= b - b = 0$

$A \cdot (x + \lambda \cdot y) = A \cdot x + \lambda \cdot A \cdot y = b + 0 = b, \forall \lambda \in IR$

Thus, it holds: $x + \lambda \cdot y \in P \Leftrightarrow x + \lambda \cdot y \geq 0$

Proof of Theorem 1.3.19

Case 1 : $x \geq y \geq 0$

Calculate $\lambda^{-} = \max\left\{-\frac{x_j}{y_j} \mid y_j > 0\right\} = -\frac{x_{\bar{j}}}{y_{\bar{j}}} \in [-1, -\infty]$

Furthermore, let: $x^{-} = x + \lambda^{-} \cdot y = x - \frac{x_{\bar{j}}}{y_{\bar{j}}} \cdot y = \begin{pmatrix} \hat{x}_1 \\ \dots \\ \underset{\bar{j}}{0} \\ \dots \\ \hat{x}_n \end{pmatrix} \geq 0$

Obviously, it holds: $n_{x^-} < n_x \underset{Induction}{\Rightarrow} x^{-} = v + w, v \in C(\varepsilon(P)) \wedge w \in D(P)$

Proof of Theorem 1.3.19

Obviously, it holds: $n_{x^-} < n_x \underset{Induction}{\Rightarrow} x^{-} = v + w, v \in C(\varepsilon(P))$
$\wedge w \in D(P)$

$x^{-} = x + \lambda^{-} \cdot y \Rightarrow v + w = x + \lambda^{-} \cdot y$
$\Rightarrow x = v + w - \lambda^{-} \cdot y, v \in C(\varepsilon(P))$

Note that $\lambda^{-} < 0 \Rightarrow \lambda^{-} \cdot y < 0 \Rightarrow w - \lambda^{-} \cdot y \geq 0$ and
$A \cdot (w - \lambda^{-} \cdot y) = A \cdot w - \lambda^{-} \cdot A \cdot y = 0 + 0 = 0 \Rightarrow w - \lambda^{-} \cdot y \in D(P)$
$\Rightarrow x \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

Proof of Theorem 1.3.19

Case 2 : $y \leq 0$

Calculate $\lambda^{+} = \min\left\{-\frac{x_j}{y_j} \mid y_j < 0\right\} = -\frac{x_{\bar{j}}}{y_{\bar{j}}} \in [1, \infty]$

Furthermore, let: $x^{+} = x + \lambda^{+} \cdot y = x - \frac{x_{\bar{j}}}{y_{\bar{j}}} \cdot y = \begin{pmatrix} \hat{x}_1 \\ \dots \\ \underset{\bar{j}}{0} \\ \dots \\ \hat{x}_n \end{pmatrix} \geq 0$

Obviously, it holds: $n_{x^+} < n_x \underset{Induction}{\Rightarrow} x^{+} = v + w, v \in C(\varepsilon(P)) \wedge w \in D(P)$

Proof of Theorem 1.3.19

Obviously, it holds: $n_{x^+} < n_x$

$\underset{Induction}{\Rightarrow} x^+ = v + w, v \in C(\varepsilon(P)) \wedge w \in D(P)$

$x^+ = x + \lambda^+ \cdot y \Rightarrow v + w = x + \lambda^+ \cdot y \Rightarrow x = v + w - \lambda^+ \cdot y, v \in C(\varepsilon(P))$

Note that $\lambda^+ > 0 \Rightarrow w - \lambda^+ \cdot y \geq 0$ (since $y \leq 0$) and

$A \cdot (w + \lambda^+ \cdot y) = A \cdot w + \lambda^+ \cdot A \cdot y = 0 + 0 = 0 \Rightarrow w + \lambda^+ \cdot y \in D(P)$

$\Rightarrow x \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

Proof of Theorem 1.3.19

Case 3 : $\exists j \neq i : y_j < 0 \wedge y_i > 0$

Calculate $\lambda^+ = \min\left\{-\frac{x_j}{y_j} \mid y_j < 0\right\} = -\frac{x_{\bar{j}}}{y_{\bar{j}}} \in [1, \infty]$

Calculate $\lambda^- = \max\left\{-\frac{x_j}{y_j} \mid y_j > 0\right\} = -\frac{x_{\tilde{j}}}{y_{\tilde{j}}} \in [-\infty, -1]$

Furthermore, let: $x^- = x + \lambda^- \cdot y = x - \frac{x_{\tilde{j}}}{y_{\tilde{j}}} \cdot y = \begin{pmatrix} \hat{x}_1 \\ \ldots \\ \underset{\tilde{j}}{0} \\ \ldots \\ \hat{x}_n \end{pmatrix} \geq 0$

Proof of Theorem 1.3.19

Obviously, it holds: $n_{x^-} < n_x \underset{Induction}{\Rightarrow} x^- = a_- + b_-, a_- \in C(\varepsilon(P)) \wedge b_- \in D(P)$

Furthermore, let: $x^+ = x + \lambda^+ \cdot y = x + \frac{x_{\bar{j}}}{y_{\bar{j}}} \cdot y = \begin{pmatrix} \hat{x}_1 \\ \ldots \\ \underset{\bar{j}}{0} \\ \ldots \\ \hat{x}_n \end{pmatrix} \geq 0$

(since the considered entries of y are < 0)

Obviously, it holds: $n_{x^+} < n_x \underset{Induction}{\Rightarrow} x^+ = a_+ + b_+, a_+ \in C(\varepsilon(P)) \wedge b_+ \in D(P)$

Proof of Theorem 1.3.19

Consider the following convex combination

(note that $\lambda^- < 0 \Rightarrow \lambda^+ - \lambda^- > 0$)

$$\frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot x^- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot x^+ = \frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot \left(x + \lambda^- \cdot y\right)$$

$$- \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot \left(x + \lambda^+ \cdot y\right)$$

$$= \frac{\lambda^+ \cdot x + \lambda^+ \cdot \lambda^- \cdot y}{\lambda^+ - \lambda^-} - \frac{\lambda^- \cdot x + \lambda^- \cdot \lambda^+ \cdot y}{\lambda^+ - \lambda^-}$$

$$= \frac{\left(\lambda^+ - \lambda^-\right) \cdot x + \lambda^+ \cdot \lambda^- \cdot y - \lambda^- \cdot \lambda^+ \cdot y}{\lambda^+ - \lambda^-} = \frac{\left(\lambda^+ - \lambda^-\right) \cdot x}{\lambda^+ - \lambda^-} = x.$$

Proof of Theorem 1.3.19

Thus, we know:

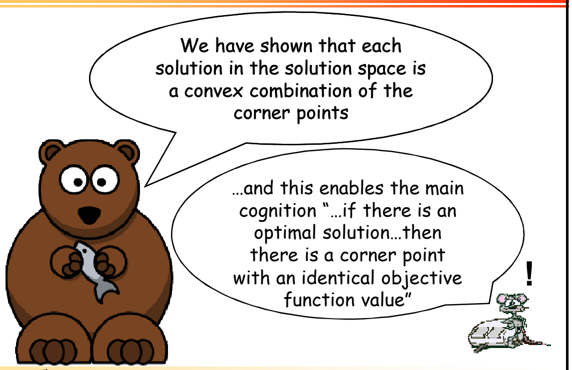
$$x = \frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot x^- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot x^+ = \frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot \left(a_- + b_-\right) - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot \left(a_+ + b_+\right)$$

$$= \underbrace{\left(\frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot a_- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot a_+\right)}_{\in C(\varepsilon(P))} + \underbrace{\frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot b_- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot b_+}_{\substack{\geq 0 \wedge A \cdot \left(\frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot b_- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot b_+\right) = \frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot A \cdot b_- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot A \cdot b_+ = 0 \\ \Rightarrow \frac{\lambda^+}{\lambda^+ - \lambda^-} \cdot b_- - \frac{\lambda^-}{\lambda^+ - \lambda^-} \cdot b_+ \in D(P)}}$$

$$x \in \left\{x \in IR^n \mid x = y + z \wedge y \in C\left(\varepsilon\left(P\right)\right) \wedge z \in D\left(P\right)\right\}$$

$$\Rightarrow P = \left\{x \in IR^n \mid x = y + z \wedge y \in C\left(\varepsilon\left(P\right)\right) \wedge z \in D\left(P\right)\right\}$$

Now, we are almost done with it



Important consequence

1.3.20 Observation

1. $P \neq \emptyset \Rightarrow \varepsilon(P) \neq \emptyset$
2. If there is an optimal solution in P, there exists a corner point with identical optimal costs,i.e., there is also an optimal corner point
3. If $P \neq \emptyset$ and if there is no optimal solution in P
   $\Rightarrow \exists y \in D(P) : c^T \cdot y > 0$
4. $P \neq \emptyset \wedge P$ bounded $\Rightarrow P = C(\varepsilon(P))$

Proof of Observation 1.3.20

1. $P \neq \emptyset \Rightarrow \exists x \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

$\Rightarrow x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)$

Case 1: $0 \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

$\Rightarrow 0 \in \varepsilon(P)$

Case 2: $0 \notin \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

$\Rightarrow \exists x(\neq 0) \in \{x \in IR^n \mid x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)\}$

$\Rightarrow x = y + z > 0 \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)$

Since it holds that $A \cdot z = 0$, we conclude $y \in P \wedge y = \sum_{i=1}^{k} \alpha_i \cdot a^i,\ a^i \in \varepsilon(P) \Rightarrow \varepsilon(P) \neq \emptyset$

Proof of Observation 1.3.20

$2+3$. Let $\{x^1, \ldots, x^k\} = \varepsilon(P)$. We introduce $x^{\hat{j}}$ as the corner point that possesses maximal objective function value, i.e., $c^T \cdot x^{\hat{j}} = \max\{c^T \cdot x^i \mid x^i \in \varepsilon(P)\}$

Consider now $x \in P \Rightarrow x = y + z \wedge y \in C(\varepsilon(P)) \wedge z \in D(P)$

Calculate

$$c^T \cdot x = c^T \cdot (y + z) = c^T \cdot y + c^T \cdot z = c^T \cdot \left(\sum_{i=1}^{k} \alpha_i \cdot x^i\right) + c^T \cdot z,$$

with $\sum_{i=1}^{k} \alpha_i = 1 \Rightarrow c^T \cdot \left(\sum_{i=1}^{k} \alpha_i \cdot x^i\right) + c^T \cdot z \leq c^T \cdot x^{\hat{j}} + c^T \cdot z$

Proof of Observation 1.3.20

We know $A\cdot z=0 \Rightarrow x+\xi\cdot z\in P$

Thus, we have to distinguish

Case 1: $c^T\cdot z\le 0\Rightarrow$

There are optimal solutions in P. Specifically,

$x^{\tilde{j}}\in\varepsilon(P)$ is one of them.

Case 2: $c^T\cdot z>0\Rightarrow$

There is no optimal solution in P.

Proof of Observation 1.3.20

4. $P\neq\varnothing\wedge P$ bounded

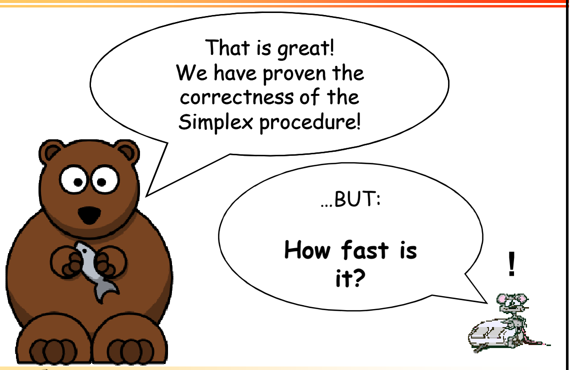
$\Rightarrow$

$\exists x\in P=\{x\in IR^n \mid x=y+z\wedge y\in C(\varepsilon(P))\wedge z\in D(P)\}$

Since $A\cdot z=0$, we know $\lambda\cdot z\in D(P), \lambda>0\Rightarrow z=0$

$\Rightarrow x\in C(\varepsilon(P))$

Correctness has been proven



1.4 How fast is the Simplex Method?

- We consider average practical problems of the general form

$$\text{Maximize} \sum_{j=1}^{n} c_j \cdot x_j$$
$$\text{s.t. } \forall i \in \{1,\ldots,m\}: \sum_{j=1}^{n} a_{i,j} \cdot x_j \le b_i \wedge \forall j \in \{1,\ldots,n\}: x_j \ge 0$$

- Dantzig (1963) reported that the number of iterations that the Simplex procedure conducts is usually less than *3m/2* and only rarely going to *3m* (*m<50* and *m+n<200*)
- In fact, recent empirical findings underline that the average running time of the Simplex procedure is linear
- Specifically, it increases proportionally to m

1.4.1 Analyzing the average case

- Some investigations reveal that for fixed *m* the average total number of iterations to be conducted is upper bounded by *log(n)*
- Thus, if each iteration is executed efficiently, modern computers are able to solve problems with about 100 constraints and variables in a few seconds
- Even cases with n and m of size 1,000 can be solved efficiently
- However, as a prerequisite, this requires an efficient implementation of each iteration, i.e., each basis changes
- For this purpose, two attributes are decisive…
  - an appropriate pivot strategy
  - an efficient update handling

1.4.2 Analyzing the worst case

- An open question for a long time was whether the solving of an Linear Program is in P
- Based on the findings of Klee and Minty (1972), we construct a worst case polytope
- We introduce the following parameters

$$\text{For some } 0<\varepsilon<\frac{1}{2}: \ 1 \ge x_1 \ge \varepsilon \ \text{ and } \ 1-\varepsilon \cdot x_{j-1} \ge x_j \ge \varepsilon \cdot x_{j-1}, \ \forall j = 2,3,\ldots,d$$

- Moreover, we introduce the following LP (LP 1.4.2.1)

$$\begin{aligned}
&\min \ -x_d \\
&x_1 - r_1 = \varepsilon \\
&x_1 + s_1 = 1 \\
&x_j - \varepsilon \cdot x_{j-1} - r_j = 0, \ \forall j = 2,3,\ldots,d \\
&x_j + \varepsilon \cdot x_{j-1} + s_j = 1, \ \forall j = 2,3,\ldots,d \ \text{ with } x_j, r_j, s_j \ge 0, \ \forall j = 1,\ldots,d
\end{aligned}$$

The basic feasible solutions (bfs)

1.4.2.1 Lemma

The set of feasible bases of the LP $(1.4.2.1)$ is the set of subsets of $\{x_1,...,x_d,r_1,...,r_d,s_1,...,s_d\}$ containing all x-variables and exactly one of s_j, r_j for each $j = 1,...,d$. Furthermore, all these bases are nondegenerate.

Proof of Lemma 1.4.2.1:

Because $x_1 \geq \varepsilon$ and $x_{j+1} \geq \varepsilon \cdot x_j, \ \forall j = 1,...,d-1,$ we conclude that in each feasible solution we have $x_j \geq \varepsilon^j > 0$. Hence, all feasible bases must contain all d columns corresponding to the x-variables.
Moreover, assume that $\exists j \in \{1,...,d\} : r_j = s_j = 0$.
Case $1: j = 1 \Rightarrow$ Since $x_1 - r_1 = \varepsilon,$ it holds that $x_1 = \varepsilon$ and through $x_1 - s_1 = 1,$ it holds that $x_1 = 1$.
However, this implies $x_1 = \varepsilon = 1$ and contradicts the assumed parameter setting of ε.

Proof of Lemma 1.4.2.1

Case $2: j > 1 \Rightarrow$ Since $x_j - \varepsilon \cdot x_{j-1} - r_j = 0 \wedge x_j + \varepsilon \cdot x_{j-1} + s_j = 1,$ it holds that $x_j = \varepsilon \cdot x_{j-1}$ and $x_j = 1 - \varepsilon \cdot x_{j-1} \Rightarrow \varepsilon \cdot x_{j-1} = 1 - \varepsilon \cdot x_{j-1} \Rightarrow 2 \cdot \varepsilon \cdot x_{j-1} = 1$.
Since $x_1 + s_1 = 1$ and $x_j + \varepsilon \cdot x_{j-1} + s_j = x_j + \varepsilon \cdot x_{j-1} = 1, \ \forall j = 2,3,...,d,$ we have $x_j \leq 1, \ \forall j = 1,2,3,...,d$. Due to $\varepsilon < \frac{1}{2},$ we have a contradiction to $2 \cdot \varepsilon \cdot x_{j-1} = 1$.
This results from $2 \cdot \varepsilon \cdot x_{j-1} \leq 2 \cdot \varepsilon < 1$.
Therefore, each feasible basis must contain one of the columns corresponding to s_j and r_j for every $j \in \{1,...,d\}$. However, there are already $2d = m$ elements in the basis. Moreover, since all these variables are non-zero, these solutions are nondegenerate.

The set of basic feasible solutions

- In what follows, we write a bfs of the LP (1.4.2.1) as *x(S)* with *S* giving a subset of set *{1,…,d}* that indicates the nonzero *r*'s in *x(S)*
- The value of the x_j-variable in *x(S)* is abbreviated by $x_j(S)$
- Based on these abbreviations, we formulate the following Lemma

Comparing the objective values

1.4.2.2 Lemma

Suppose that $d \in S$ but $d \notin S'$; then $x_d(S) > x_d(S')$. Moreover, if additionally $S' = S - \{d\}$, we have $x_d(S') = 1 - x_d(S)$.

Proof:

Since $d \in S$, we have $s_d = 0 \Rightarrow$ Due to $x_d(S) + \varepsilon \cdot x_{d-1}(S) + s_d = x_d(S) + \varepsilon \cdot x_{d-1}(S) = 1$, it holds that $x_d(S) = 1 - \varepsilon \cdot x_{d-1}(S)$. Since $x_j, r_j, s_j \geq 0$ and $\varepsilon > 0$, we have $x_{d-1}(S) \leq 1$.

By $\varepsilon < \frac{1}{2}$ and $x_d(S) = 1 - \varepsilon \cdot x_{d-1}(S)$, we conclude that $x_d(S) = 1 - \varepsilon \cdot x_{d-1}(S) > 1 - \frac{1}{2} \cdot x_{d-1}(S) \geq \frac{1}{2}$

Moreover, since $d \notin S'$, we have $r_d = 0$. And by $x_d(S') - \varepsilon \cdot x_{d-1}(S') - r_d = 0$, we have

$x_d(S') - \varepsilon \cdot x_{d-1}(S') = 0 \Rightarrow x_d(S') = \varepsilon \cdot x_{d-1}(S') < \frac{1}{2}$. Consequently, we have $x_d(S') < x_d(S)$.

If $S' = S - \{d\}$, we have $x_{d-1}(S) = x_{d-1}(S')$ and it holds that

$x_d(S') = \varepsilon \cdot x_{d-1}(S') = 1 - (1 - \varepsilon \cdot x_{d-1}(S')) = 1 - (1 - \varepsilon \cdot x_{d-1}(S)) = 1 - x_d(S)$, with $x_d(S) > \frac{1}{2}$

$\Rightarrow x_d(S) - x_d(S') = 2 \cdot x_d(S) - 1$

Comparing the objective values

1.4.2.3 Lemma

We assume that the subsets of set $\{1, \dots, d\}$ are enumerated in such a way that $x_d(S_1) \leq x_d(S_2) \leq \dots \leq x_d(S_{2^d})$. Then, the inequalities are strict, and the basic feasible solutions $x(S_j)$ and $x(S_{j+1})$ are adjacent for $j = 1, 2, \dots, 2^d - 1$.

Proof:

We give the proof by induction:

$d = 1$: In this case there are two basic feasible solutions, namely $(x_1, r_1, s_1) \in \{(\varepsilon, 0, 1-\varepsilon), (1, 1-\varepsilon, 0)\}$ since we have $x_1 - r_1 = \varepsilon$ and $x_1 + s_1 = 1$. Clearly, the solutions have unequal nonzero x_1-values and are adjacent since exactly two columns are exchanged.

Proof of Lemma 1.4.2.3

$d > 1$: We assume that the proposition holds for d. Therefore, there is the appropriate enumeration $S_1, \dots, S_{2^d}$ of subsets of set $\{1, \dots, d\}$.

Clearly, these are also subsets of set $\{1, \dots, d, d+1\}$. Due to $r_{d+1} = 0$ in the corresponding solutions and $x_{d+1}(S_j) - \varepsilon \cdot x_d(S_j) - r_d = 0$, we have $x_{d+1}(S_j) = \varepsilon \cdot x_d(S_j)$.

$\Rightarrow$ By applying the induction hypothesis, we conclude that

$x_{d+1}(S_1) < x_{d+1}(S_2) < \dots. < x_{d+1}(S_{2^d})$.

Furthermore, we consider the remaining subsets of set $\{1, \dots, d, d+1\}$, namely, $S'_j = S_j \cup \{d+1\}$, i.e., with $s_{d+1} = 0$. By applying $x_{d+1}(S'_j) + \varepsilon \cdot x_d(S'_j) + s_{d+1} = 1$, we obtain $x_{d+1}(S'_j) = 1 - \varepsilon \cdot x_d(S'_j)$. By Lemma 1.4.2.3, we have $x_{d+1}(S'_j) > x_{d+1}(S_{2^d})$ and $x_{d+1}(S'_j) = 1 - x_{d+1}(S_j)$ with $x_{d+1}(S_j) < \frac{1}{2}$. Thus, we have

$x_{d+1}(S_1) < x_{d+1}(S_2) < \dots. < x_{d+1}(S_{2^d}) < x_{d+1}(S'_{2^d}) < x_{d+1}(S'_{2^d-1}) < \dots. < x_{d+1}(S'_2) < x_{d+1}(S'_1)$.

Proof of Lemma 1.4.2.3

By induction hypothesis, we know that $x(S_j)$ and $x(S_{j+1})$ are adjacent for $j=1,....,2^d-1$.
Also $x(S'_j)$ and $x(S'_{j+1})$ are adjacent for $j=1,....,2^d-1$ since, again by induction hypothesis, $x(S_j)$ and $x(S_{j+1})$ are adjacent.
Moreover, $x(S_{2^d})$ and $x(S'_{2^d})$ are adjacent since r_{d+1} is added to the basis while s_{d+1} leaves the basis.
This completes the proof.

Main result

1.4.2.4 Theorem

For every $d>1$ there is an LP with 2d equations, 3d variables, and integer coefficients with absolute value bounded by 4, such that the simplex algorithm may take 2^d-1 iterations to find the optimal solution.

Proof:

We set $\varepsilon=\frac{1}{4}$ and multiply all equations of the LP (1.4.2.1) by 4.
Therefore, all coefficients become integer. Since the objective is to maximize x_d $(=\min -x_d)$, the exponentially long chain of 2^d adjacent bfs's whose existance is established by Lemma 1.4.2.4 has decreasing costs.
This proves the Theorem.

Consequences

- Results similar to Theorem 1.4.2.4 are known for all variations of simplex algorithms, including several heuristic pivoting rules

Another worst case example

$$\text{Maximize} \sum_{j=1}^{n} 10^{n-j} \cdot x_j$$

$$\text{s.t. } \forall i \in \{1,\ldots,n\}: \left(2 \cdot \sum_{j=1}^{i-1} 10^{i-j} \cdot x_j\right) + x_i \le 100^{i-1} \wedge \forall j \in \{1,\ldots,n\}: x_j \ge 0$$

$$\Rightarrow n = 3:$$
$$\text{Maximize } 100 \cdot x_1 + 10 \cdot x_2 + x_3$$
$$\text{s.t. } x_1 \le 1 \wedge 20 \cdot x_1 + x_2 \le 100 \wedge 200 \cdot x_1 + 20 \cdot x_2 + x_3 \le 10{,}000$$
$$\wedge\, x_1, x_2, x_3 \ge 0$$

Using the largest coefficient rule

$$x_1 + x_4 = 1$$
$$20 \cdot x_1 + x_2 + x_5 = 100$$
$$200 \cdot x_1 + 20 \cdot x_2 + x_3 + x_6 = 10{,}000$$
$$-----------------$$
$$100 \cdot x_1 + 10 \cdot x_2 + x_3 = z$$

Using the largest coefficient rule

$$x_4 = 1 - x_1$$
$$x_5 = 100 - 20 \cdot x_1 - x_2$$
$$x_6 = 10{,}000 - 200 \cdot x_1 - 20 \cdot x_2 - x_3$$
$$-----------------$$
$$\boxed{100 \cdot x_1} + 10 \cdot x_2 + x_3 = z$$

Using the largest coefficient rule

Iteration 1

$$x_1 = 1 - x_4$$
$$x_5 = 80 + 20 \cdot x_4 - x_2$$
$$x_6 = 9{,}800 + 200 \cdot x_4 - 20 \cdot x_2 - x_3$$
$$100 - 100 \cdot x_4 + \boxed{10 \cdot x_2} + x_3 = z$$

Using the largest coefficient rule

Iteration 2

$$x_1 = 1 - x_4$$
$$x_2 = 80 + 20 \cdot x_4 - x_5$$
$$x_6 = 8{,}200 - 200 \cdot x_4 + 20 \cdot x_5 - x_3$$
$$900 + \boxed{100 \cdot x_4} - 10 \cdot x_5 + x_3 = z$$

Using the largest coefficient rule

Iteration 3

$$x_4 = 1 - x_1$$
$$x_2 = 100 - 20 \cdot x_1 - x_5$$
$$x_6 = 8{,}000 + 200 \cdot x_1 + 20 \cdot x_5 - x_3$$
$$1{,}000 - 100 \cdot x_1 - 10 \cdot x_5 + \boxed{x_3} = z$$

Using the largest coefficient rule

Iteration 4

$$x_4 = 1 - x_1$$
$$x_2 = 100 - 20 \cdot x_1 - x_5$$
$$x_3 = 8{,}000 + 200 \cdot x_1 + 20 \cdot x_5 - x_6$$
$$9{,}000 + \boxed{100 \cdot x_1} + 10 \cdot x_5 - x_6 = z$$

Using the largest coefficient rule

Iteration 5

$$x_1 = 1 - x_4$$
$$x_2 = 80 + 20 \cdot x_4 - x_5$$
$$x_3 = 8{,}200 - 200 \cdot x_4 + 20 \cdot x_5 - x_6$$
$$9{,}100 - 100 \cdot x_4 + \boxed{10 \cdot x_5} - x_6 = z$$

Using the largest coefficient rule

Iteration 6

$$x_1 = 1 - x_4$$
$$x_5 = 80 + 20 \cdot x_4 - x_2$$
$$x_3 = 9{,}800 + 200 \cdot x_4 - 20 \cdot x_2 - x_6$$
$$9{,}900 + \boxed{100 \cdot x_4} - 10 \cdot x_2 - x_6 = z$$

Using the largest coefficient rule

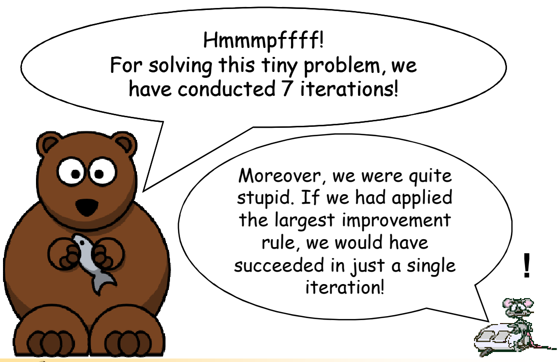
Iteration 7

$$x_4 = 1 - x_1$$
$$x_5 = 100 - 20 \cdot x_1 - x_2$$
$$x_3 = 10{,}000 - 200 \cdot x_1 - 20 \cdot x_2 - x_6$$

$$10{,}000 - 100 \cdot x_1 - 10 \cdot x_2 - x_6 = z$$

$\Rightarrow$ Optimal solution!

Assessing this calculation



Largest improvement rule

$$x_4 = 1 - x_1$$
$$x_5 = 100 - 20 \cdot x_1 - x_2$$
$$x_6 = 10{,}000 - 200 \cdot x_1 - 20 \cdot x_2 - x_3$$

$$100 \cdot x_1 + 10 \cdot x_2 + x_3 = z$$

$$\forall j = 1,\ldots,n : c_j > 0 : \text{Improvement}_j = c_j \cdot \min\left\{ \frac{b_i}{-a_{ij}} \middle| a_{ij} < 0, i = 1,\ldots,m \right\}$$

$$\Rightarrow x_1 \le \min\left\{1, \frac{100}{20}, \frac{10{,}000}{200}\right\} = \min\{1,5,50\} = 1 \Rightarrow \text{Improvement}_1 = 100$$

$$x_2 \le \min\left\{100, \frac{10{,}000}{20}\right\} = \min\{100, 500\} = 100 \Rightarrow \text{Improvement}_2 = 1{,}000$$

$$x_3 \le \min\{10{,}000\} = 10{,}000 \Rightarrow \text{Improvement}_3 = 10{,}000$$

Choose the variable with the largest improvement $\Rightarrow x_3$

Largest improvement rule

Iteration 1

$$x_4 = 1 - x_1$$
$$x_5 = 100 - 20 \cdot x_1 - x_2$$
$$x_3 = 10{,}000 - 200 \cdot x_1 - 20 \cdot x_2 - x_6$$
$$------------------$$
$$10{,}000 - 100 \cdot x_1 - 10 \cdot x_2 - x_6 = z$$

$\Rightarrow$ Optimal solution!

Alternative pivoting rules

- Two efficiency aspects to be considered for assessing pivoting rules
  - Number of iterations that are induced by the application of the rule
  - Effort of each iteration
- Generally, it can be stated that the number of iterations required by the largest improvement rule is usually smaller than the number of iterations caused by the largest coefficient rule
- This was underlined empirically
- However, the costs caused by each iteration are increased by the largest improvement rule
- Nevertheless, in a direct comparison the reduced number of iterations prevails and therefore the largest improvement rule outperforms the largest coefficient rule

Attention

- **However, as already mentioned, each rule has its own specific worst case scenario**
- Thus, according to worst case considerations, there is no real distinction between different pivoting rules
- In modern software packages, pivoting rules are chosen according to the handling of large sized problems on a computer
- However, it can be shown that LP is polynomially solvable. But this is done by using a different solution strategy

Complexity of Linear Programming

- Until 1979, the open question whether there can be any polynomial-time algorithm for LP was a most perplexing question (Papadimitriou and Steiglitz (1982,1988), p.170)
- Specifically, there was conflicting evidence about the possible answer
  - On the one hand, LP was certainly one of the problems (together with the TSP and many others) which seemed to defy all reasonable attempts at the development of a polynomial-time algorithm.
  - However, on the other hand, LP had two positive features that made it completely different from the other classical hard problems
    - First, LP has a strong duality theory, which is conspicuously lacking for all the other hard combinatorial problems
    - Secondly, LP has an algorithm, the simplex method, which – although exponential in its worst case – certainly works empirically on instances of seemingly unlimited size.

Complexity of Linear Programming

- In the spring of 1979 the Soviet mathematician L.G. Khachian proposed an exact polynomial-time solution algorithm for LP (see the paper of Khachian, L. G. (1979)), the so-called **ellipsoid algorithm**
- Therefore, it was proven that LP is well solvable (in the language of the Complexity Theory), i.e., $LP \in P$
- This important work was assessed, evaluated and further extended by the papers of Aspvall and Stone (1980), Dantzig, G.B. (1979), and Goldfarb, D., Todd, M.J. (1980)

Performance of the ellipsoid algorithm

- Despite the great theoretical value of the ellipsoid algorithm (for worst case scenarios), this algorithm seems to be not very useful in practice
  - The most obvious among many obstacles is the large precision that is required by the conducted computations
  - Hence, average running times are not competitive, i.e., it is outperformed by the Simplex algorithm for real-world problems

Interior point methods

- In 1984, however, Karmarkar, N. (1984) proposes a further polynomial exact solution algorithm for Linear Programming
- In contrast to the Simplex algorithm that moves from edge point to edge point, this procedure finds an optimal solution by iteratively moving through the interior of the solution space until optimality was proven
- Interior methods are also very efficient in practice and are competitive with the Simplex algorithm
  - This applies in particular to LP with sparsely populated matrices
  - However, the Simplex algorithm is superior if a series of problems has to be solved (e.g., applied as a subroutine within a Branch&Bound algorithm for integer problems)

1.5 How to work with tableaus

- In order to provide a direct understanding of the Simplex procedure, we have illustrated its calculations on the basis of dictionaries
- However, in what follows, we make use of tableaus
- Tableaus are directly derived from the use of matrices in order to solve Linear Programs
- By making use of them, we are able to illustrate several aspects of the matrix transformations executed during the conduction of the Simplex procedure
- Moreover, matrix operations play a crucial role for implementing the Simplex Method as efficiently as possible
- This is done by the so called Revised Simplex Method

1.5.1 Basis change

Consider a basic feasible solution $x^0 = \begin{pmatrix} x_B^0 \\ x_N^0 \end{pmatrix}$, with $x_B^0 = A_B^{-1} \cdot b \wedge x_N^0 = 0$

and additionally a feasible solution x with $A \cdot x = b$. Furthermore, let $y \in IR^n$ with $A \cdot y = 0$ and it holds: $x = x^0 + y$.

Then, since $x_N^0 = 0$, it obviously holds: $x_N = y_N$.

Additionally, we derive:

$$0 = A \cdot y = A_B \cdot y_B + A_N \cdot x_N \Rightarrow 0 = A_B^{-1} \cdot A_B \cdot y_B + A_B^{-1} \cdot A_N \cdot x_N$$

$$\Rightarrow 0 = y_B + A_B^{-1} \cdot A_N \cdot x_N \Rightarrow y_B = -A_B^{-1} \cdot A_N \cdot x_N \Rightarrow y = \begin{pmatrix} -A_B^{-1} \cdot A_N \cdot x_N \\ x_N \end{pmatrix}.$$

We define the basis vectors by B and the remaining vectors by N.

Basis change

$$y = \begin{pmatrix} y_B \\ y_N \end{pmatrix} = \begin{pmatrix} -A_B^{-1} \cdot A_N \cdot x_N \\ x_N \end{pmatrix}$$

In the following, we introduce the component t in the basis.
We introduce the shortcut $\overline{A} = A_B^{-1} \cdot A$,
with $A_B^{-1} \in IR^{m \times m}, \overline{A} = \left(\overline{a}^1, ..., \overline{a}^n\right) = A_B^{-1} \cdot \left(a^1, ..., a^n\right)$, with $a^j \in IR^m$.

Note that $A_B^{-1} \in IR^{m \times m} \wedge A \in IR^{m \times n} \Rightarrow \overline{A} \in IR^{m \times n}$.

Let us now set: $x_N = e^{k_0} = \left(0, ..., 0, \underbrace{1}_{\text{Position } k_0}, 0, ..., 0\right)^T \in IR^{n-m}$.

Basis change

Furthermore, we assume $t = N(k_0)$. I.e., the tth column in the original matrix represents the k_0th non-basic column.
Thus, it follows in this case:

$$y_B = -A_B^{-1} \cdot A_N \cdot x_N = -\overline{a}^t \Rightarrow y = \begin{pmatrix} y_B \\ y_N \end{pmatrix} = \begin{pmatrix} -\overline{a}^t_{(B)} \\ e^{k_0} \end{pmatrix}$$

Let $x^\lambda = x^0 + \lambda \cdot y \Rightarrow A \cdot \left(x^0 + \lambda \cdot y\right) = A \cdot x^0 + A \cdot \lambda \cdot y = b + 0 = b$

$$x^0 + \lambda \cdot y \geq 0 \Leftrightarrow \lambda \cdot y \geq -x^0 \Leftrightarrow \lambda \cdot \begin{pmatrix} -\overline{a}^t_{(B)} \\ e^{k_0} \end{pmatrix} \geq -x^0$$

Conducting a basis change

$$\lambda \cdot \begin{pmatrix} -\overline{a}^t_{(B)} \\ e^{k_0} \end{pmatrix} \geq -x^0 \Leftrightarrow \begin{pmatrix} x^0_{(B)} - \lambda \cdot \overline{a}^t_{(B)} \\ \lambda \cdot e^{k_0} \end{pmatrix} \geq 0 \Rightarrow$$

$$\lambda = \begin{cases} \min\left\{ \dfrac{x^0_{(B),j}}{\overline{a}^t_j} \mid \overline{a}^t_j > 0 \right\} \text{ if } \exists j : \overline{a}^t_j > 0 \\ \geq 0 \text{ otherwise} \end{cases}$$

Note that $x^\lambda \geq 0$ is feasible for these values of λ. Let us consider the

result of our calculations: $x^\lambda = \begin{pmatrix} x^0_{(B)} - \lambda \cdot \overline{a}^t_{(B)} \\ 0 \\ ... \\ \lambda \\ ... \\ 0 \end{pmatrix}$, i.e., $x^\lambda_j = \begin{cases} x^0_{(B),j} - \lambda \cdot \overline{a}^t_i & \text{if } B(i) = j \\ \lambda & \text{if } j = t \\ 0 & \text{otherwise} \end{cases}$

Observation

- By setting λ to a maximal feasible value, we erase the corresponding variable out of the basis and introduce the *t*th entry instead
- In the following, we examine a simple example

Basis change – An illustrative example I

$$\left|\begin{array}{l} 100\cdot a+250\cdot b\geq 500 \\ 150\cdot a+200\cdot b\geq 600 \\ 100\cdot a+50\cdot b\geq 250 \end{array}\right| \Leftrightarrow \left|\begin{array}{l} 2\cdot a+5\cdot b\geq 10 \\ 3\cdot a+4\cdot b\geq 12 \\ 2\cdot a+1\cdot b\geq 5 \end{array}\right|$$

We have $m=3 \wedge n=2+3=5$

Min $f(a,b)=20\cdot a+30\cdot b$

$$A=\begin{pmatrix} 2 & 5 & -1 & 0 & 0 \\ 3 & 4 & 0 & -1 & 0 \\ 2 & 1 & 0 & 0 & -1 \end{pmatrix} \wedge b=\begin{pmatrix} 10 \\ 12 \\ 5 \end{pmatrix}, \text{set } B(1)=2, B(2)=3, B(3)=4$$

Basis change – An illustrative example II

$$A=\begin{pmatrix} 2 & 5 & -1 & 0 & 0 \\ 3 & 4 & 0 & -1 & 0 \\ 2 & 1 & 0 & 0 & -1 \end{pmatrix} \wedge b=\begin{pmatrix} 10 \\ 12 \\ 5 \end{pmatrix}, \text{set } B(1)=2, B(2)=3, B(3)=4$$

$$A_B=\begin{pmatrix} 5 & -1 & 0 \\ 4 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \Rightarrow A_B^{-1}=\begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 5 \\ 0 & -1 & 4 \end{pmatrix}$$

$$\Rightarrow A_B^{-1}\cdot A=\begin{pmatrix} 2 & 1 & 0 & 0 & -1 \\ 8 & 0 & 1 & 0 & -5 \\ 5 & 0 & 0 & 1 & -4 \end{pmatrix} \wedge A_B^{-1}\cdot b=\begin{pmatrix} 5 \\ 15 \\ 8 \end{pmatrix}$$

Basis change – An illustrative example III

We introduce column 1, i.e., $t=1$

$$\Rightarrow \lambda = \min\left\{\frac{5}{2},\frac{15}{8},\frac{8}{5}\right\} = \frac{8}{5} \Rightarrow x^{\lambda} = x^0 + \lambda\cdot\begin{pmatrix} -\overline{a}^t_{(B)} \\ e^{k_0} \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 5 \\ 15 \\ 8 \\ 0 \end{pmatrix} + \frac{8}{5}\cdot\begin{pmatrix} 1 \\ -2 \\ -8 \\ -5 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{8}{5} \\ \frac{9}{5} \\ \frac{11}{5} \\ 0 \\ 0 \end{pmatrix}$$ and obtain $B(1)=2, B(2)=3, B(3)=1$

Basis change – An illustrative example IV

$$A = \begin{pmatrix} 2 & 5 & -1 & 0 & 0 \\ 3 & 4 & 0 & -1 & 0 \\ 2 & 1 & 0 & 0 & -1 \end{pmatrix} \wedge b = \begin{pmatrix} 10 \\ 12 \\ 5 \end{pmatrix}$$, set $B(1)=2, B(2)=3, B(3)=4$

$$A_B = \begin{pmatrix} 5 & -1 & 0 \\ 4 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} \Rightarrow A_B^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 5 \\ 0 & -1 & 4 \end{pmatrix}$$

$$\Rightarrow A_B^{-1}\cdot A = \begin{pmatrix} 2 & 1 & 0 & 0 & -1 \\ 8 & 0 & 1 & 0 & -5 \\ 5 & 0 & 0 & 1 & -4 \end{pmatrix} \wedge A_B^{-1}\cdot b = \begin{pmatrix} 5 \\ 15 \\ 8 \end{pmatrix}$$

Instead of introducing column 1, now, we try column 5, i.e., $t=5$

Basis change – An illustrative example V

Thus, we obtain:

$$\Rightarrow \overline{a}^5 = \begin{pmatrix} -1 \\ -5 \\ -4 \end{pmatrix} < 0$$

$$\Rightarrow \lambda \geq 0 \Rightarrow x^{\lambda} = x^0 + \lambda\cdot\begin{pmatrix} -\overline{a}^t_{(B)} \\ e^{k_0} \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 15 \\ 8 \\ 0 \end{pmatrix} + \lambda\cdot\begin{pmatrix} 0 \\ 1 \\ 5 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5+\lambda \\ 15+5\cdot\lambda \\ 8+4\cdot\lambda \\ \lambda \end{pmatrix}$$

and yield a ray.

Direct consequences

1.5.2 Theorem

1. If there exists an i with $\bar{a}_i^t > 0$, then $\lambda_0 = \min\left\{ \frac{x^0_{B(i)}}{\bar{a}_i^t} \mid \bar{a}_i^t > 0 \right\} = \frac{x^0_{B(\tilde{i})}}{\bar{a}_{\tilde{i}}^t}$,

and $\bar{x} = x^{\lambda_0} = x^0 + \lambda_0 \cdot y$ is a basic feasible solution (bfs) of P. Basis is $B/\{\tilde{i}\} \cup \{t\}$.

2. If $\bar{a}^t \leq 0$, then $y = \begin{pmatrix} y_B \\ y_N \end{pmatrix}$ with $y_B = -\bar{a}^t$, and $y_N = e^{k_0}$ greater equal 0

and, therefore, it holds: $x^\lambda = x^0 + \lambda \cdot y \in P, \forall \lambda \geq 0$

Proof of Theorem 1.5.2

1. Define $\bar{N}$ by $\bar{N}(i) = \begin{cases} N(i) & \text{if } i \neq \tilde{i} \\ B(\tilde{i}) & \text{otherwise} \end{cases}$.

$\Rightarrow \bar{x} = x^0 + \lambda_0 \cdot y \Rightarrow \forall j \in \bar{N}(\{1,\ldots,n-m\})$:

$\bar{x}_j = 0 \Rightarrow \bar{x}_{(\bar{N})} = 0$

Since $a_i^t > 0 \Rightarrow a^t = A_B \cdot \bar{a}^t, \bar{a}^t = A_B^{-1} \cdot a^t$.

We replace $B(\tilde{i})$ by t in the basis. Thus, $\bar{B}$ arises.

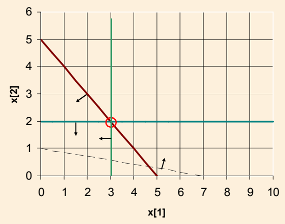
$A \cdot \bar{x} = b.$

$\Rightarrow \bar{x}$ is basic feasible solution for P.

2. trivial

Kinds of degeneration

- Primal degeneration of x^0:
  If λ_0=0, then one basic variable equals zero. The objective function value is kept unchanged



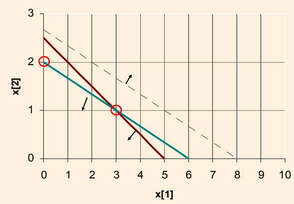
$$\begin{aligned} &\textit{Maximize} && x_1 + 7x_2 \\ &s.t. && x_1 + x_2 \leq 5 \\ &&& x_2 \leq 2 \\ &&& x_1 \leq 3 \\ &&& x_1, x_2 \geq 0 \end{aligned}$$

Observation

- Consider the example
  - We have five variables altogether (two structure variables and three slack variables)
  - Since m=3, we have always three basic variables
- Clearly, one slack variable becomes zero in the optimal solution
- Note that this is not restricted to optimal solutions

Kinds of degeneration

- Dual degeneration of x^0:
  If for one non-basic variable the relative costs are zero, we are facing a constellation of dual degeneration, i.e., all solutions integrating this variable into the basis yield the same objective function value



$$\begin{aligned} \textit{Maximize} \quad & x_1 + 3x_2 \\ s.t. \quad & x_1 + 2x_2 \le 5 \\ & 2x_1 + 6x_2 \le 12 \\ & x_1, x_2 \ge 0 \end{aligned}$$

Neighboring basic feasible solution

1.5.3 Definition
Two basic feasible solutions that can be mutually transformed in each other by changing a single basis vector are denoted as **neighboring basic feasible solutions.**

Basis change and objective function value

Let $x^0 = \begin{pmatrix} x_B^0 \\ x_N^0 \end{pmatrix}$, with $x_B^0 = A_B^{-1} \cdot b, x_N^0 = 0$ a basic feasible solution.

Assuming it holds: $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = x^0 + \begin{pmatrix} y_B \\ y_N \end{pmatrix}$, $y_B = -A_B^{-1} \cdot A_N \cdot x_N \wedge$ $y_N = x_N$ for x with $A \cdot x = b$.

We calculate: $c^T \cdot x = c^T \cdot x^0 + c_B^T \cdot y_B + c_N^T \cdot y_N$

$$= c^T \cdot x^0 - c_B^T \cdot A_B^{-1} \cdot A_N \cdot x_N + c_N^T \cdot y_N$$
$$= c^T \cdot x^0 - c_B^T \cdot A_B^{-1} \cdot A_N \cdot x_N + c_N^T \cdot x_N$$
$$= c^T \cdot x^0 + \left(c_N^T - c_B^T \cdot A_B^{-1} \cdot A_N\right) \cdot x_N$$

Substitution

We introduce $\pi^T = c_B^T \cdot A_B^{-1}$ and $z^T = \pi^T \cdot A$

$$\Rightarrow c^T \cdot x^0 + \left(c_N^T - c_B^T \cdot A_B^{-1} \cdot A_N\right) \cdot x_N$$
$$= c^T \cdot x^0 + \left(c_N^T - z_N^T\right) \cdot x_N$$

$$\Rightarrow z_B^T = \left(c_B^T \cdot A_B^{-1} \cdot A\right)_B = c_B^T \cdot A_B^{-1} \cdot A_B = c_B^T$$
$$\wedge\, z_N^T = c_B^T \cdot A_B^{-1} \cdot A_N$$

Relative cost change

Altogether, we get
$c^T \cdot x = c^T \cdot x^0 + \left(c_N^T - z_N^T\right) \cdot x_N = c^T \cdot x^0 + \left(c^T - z^T\right) \cdot x.$

Thus,
$\bar{c}^T = c^T - z^T$ defines the **relative costs** change
(or reduced costs/prices) when x^0 is transformed into x.

Specifically, it holds: $c^T \cdot x = c^T \cdot x^0 + \bar{c}^T \cdot x.$

Optimality criterion

1.5.4 Theorem

1. By moving between neighboring basic feasible solutions as introduced above, the objective function value is modified by $\overline{c}_t \cdot \lambda_0$.
2. If $\overline{c} \geq 0$, then x^0 is an optimal solution for a minimization problem.

Note that $\overline{c} \leq 0$ is the optimality criterion for maximization problems.

Proof of Theorem 1.5.4

1. Calculate

$c^T \cdot \overline{x} = c^T \cdot x^0 + \left(c_N^T - z_N^T\right) \cdot \overline{x}_N = c^T \cdot x^0 + \overline{c}_t \cdot \lambda_0$, since $\overline{x}_N = \lambda_0 \cdot e^k, N(k) = t$

2. Consider an arbitrary solution $x \in P$ and a minimization problem

$\Rightarrow A \cdot x = b \wedge x \geq 0 \Rightarrow c^T \cdot x = c^T \cdot x + z^T \cdot x - z^T \cdot x$

$\left(\text{We assume } \overline{c} \geq 0 \Rightarrow c^T \cdot x - z^T \cdot x \geq 0\right)$

$\geq z^T \cdot x = c_B^T \cdot A_B^{-1} \cdot A \cdot x = \pi^T \cdot A \cdot x = \pi^T \cdot b$

$= \pi^T \cdot A \cdot x^0 = z^T \cdot x^0 = z_B^T \cdot x_B^0 + z_N^T \cdot x_N^0$

$= z_B^T \cdot x_B^0 + 0 = c_B^T \cdot x_B^0 = c^T \cdot x^0$

$\Rightarrow \overline{c} \geq 0 \Rightarrow \forall x \in P : c^T \cdot x \geq c^T \cdot x^0 \Rightarrow x^0$ is optimal

Summary

Assuming an LP Problem is given in standard form, i.e., minimize $c^T \cdot x$ with respect to $x \geq 0 \wedge A \cdot x = b$.

Furthermore, we assume $rank(A) = m$ and that x^0 is a basic feasible solution (bfs).

We are transforming the problem by A_B^{-1}. Denote E_m as an $m \times m$ elementary matrix.

We introduce $\overline{A} = A_B^{-1} \cdot A = \left(\overline{A}_B, \overline{A}_N\right) = \left(E_m, \overline{A}_N\right)$,

$\overline{b} = A_B^{-1} \cdot b = x_B \geq 0$, and $\overline{c}^T = c^T - \pi^T \cdot A$.

Summary

By multiplying A_B^{-1}, we get the following equivalent problem:

Minimize $c^T \cdot x$, s.t. $A_B^{-1} \cdot A \cdot x = A_B^{-1} \cdot b \Leftrightarrow \overline{A} \cdot x = \overline{b}$

$$
\begin{aligned}
c^T \cdot x &= c^T \cdot x + z^T \cdot x - z^T \cdot x = \left(c^T - z^T\right) \cdot x + z^T \cdot x \\
&= \left(c^T - z^T\right) \cdot x + c_B^T \cdot A_B^{-1} \cdot A \cdot x = \pi^T \cdot b + \left(c^T - z^T\right) \cdot x \\
&= \pi^T \cdot A \cdot x^0 + \left(c^T - z^T\right) \cdot x = z^T \cdot x^0 + \left(c^T - z^T\right) \cdot x \\
&= z_B^T \cdot x_B^0 + z_N^T \cdot x_N^0 + \left(c^T - z^T\right) \cdot x = z_B^T \cdot x_B^0 + 0 + \left(c^T - z^T\right) \cdot x \\
&= c_B^T \cdot x_B^0 + \left(c^T - z^T\right) \cdot x = c^T \cdot x^0 + \left(c^T - z^T\right) \cdot x = c^T \cdot x^0 + \overline{c}^T \cdot x
\end{aligned}
$$

3 Cases may occur

1. $\overline{c} \geq 0 \Rightarrow x^0$ is an optimal solution to P
2. $\exists t : \overline{c}_t < 0 \wedge \overline{a}^t \leq 0 \Rightarrow$ The objective function is not bounded against ∞
3. $\exists t : \exists j : \overline{c}_t < 0 \wedge \overline{a}_j^t > 0 \Rightarrow \exists x^1 \in \varepsilon(P) : c^T \cdot x^1 \leq c^T \cdot x^0$.

If $\lambda_0 > 0$ it holds $c^T \cdot x^1 < c^T \cdot x^0$

Note that there are constellations possible where cases 2 and 3 apply, simultaneously. Furthermore, it is worth mentioning that all results are directly derived from Theorem 1.5.4.

1.5.5 The Tableau

- In order to obtain a basic feasible solution, the following transformation is conducted
- This is illustrated by the following tableaus
- Specifically, the modifications are stepwise done by ordinary row transformations, i.e., we produce the matrix E_m out of A and 0 out of c^T for the basis vectors of $A_B^{-1} \cdot A$

0	c^T
b	A

$\rightarrow$

$-c^T x^0$	$\overline{c}^T = c^T - c_B^T \cdot A_B^{-1} \cdot A = c^T - z^T$
$\overline{b} = A_B^{-1} \cdot b$	$\overline{A} = A_B^{-1} \cdot A$

The Primal Simplex Algorithm

1. Transform the primal problem into canonical form, generate equations and a minimization objective function.
2. Initialization with a feasible basic solution to the primal problem.
3. Are there *strictly* negative cost coefficients $\bar{c}$ in the current solution?

   "Yes": Iteration

   - Pivot column *t*: - Largest coefficient rule $\min\left\{\bar{c}_j < 0 \,\middle|\, \forall j = 1,\ldots,n\right\}$

     - Smallest subscript rule $\min\left\{j = 1,\ldots,n \,\middle|\, \bar{c}_j < 0\right\}$

     - Largest improvement rule $\max\left\{c_j \cdot \min\left\{\frac{\bar{b}_i}{\bar{a}_{ij}} \,\middle|\, \forall i : \bar{a}_{ij} > 0\right\} \,\middle|\, \forall j : c_j < 0\right\}$
   - If $\bar{a}_{it} \le 0 \; \forall i = 1,\ldots,m$, then terminate since the solution space is unbounded.
   - Pivot row s: $\min\left\{\bar{b}_i / \bar{a}_{it} \,\middle|\, \forall i = 1,\ldots,m : \bar{a}_{it} > 0\right\}$ that is an upper bound on x_t.
   - Basis change: x_t enters the basis and $x_{B(s)}$ becomes a non-basic variable.
     Apply a linear transformation of the constraint equalities by the Gauß-Jordan algorithm to yield a unit vector with $\bar{a}_{st}=1$ at the pivot element (i.e., e^s in column t). Go to step 3.

   "No": Termination. An optimal basic solution to the primal problem is found.

Calculation with tableaus – Example I.I

Consider the example from the calculation with dictionaries.

$$
\begin{aligned}
\text{Max } & 5\cdot x_1 + 4\cdot x_2 + 3\cdot x_3 = z \\
\text{s.t. } & 2\cdot x_1 + 3\cdot x_2 + 1\cdot x_3 \le 5 \\
& 4\cdot x_1 + 1\cdot x_2 + 2\cdot x_3 \le 11 \\
& 3\cdot x_1 + 4\cdot x_2 + 2\cdot x_3 \le 8 \\
& x_1, x_2, x_3 \ge 0
\end{aligned}
$$

Introducing the slack variables x_4, x_5, x_6, we transform the problem.

$$
\begin{aligned}
\text{Min } & -5\cdot x_1 - 4\cdot x_2 - 3\cdot x_3 + 0\cdot x_4 + 0\cdot x_5 + 0\cdot x_6 = -z \\
\text{s.t. } & 2\cdot x_1 + 3\cdot x_2 + 1\cdot x_3 + 1\cdot x_4 + 0\cdot x_5 + 0\cdot x_6 = 5 \\
& 4\cdot x_1 + 1\cdot x_2 + 2\cdot x_3 + 0\cdot x_4 + 1\cdot x_5 + 0\cdot x_6 = 11 \\
& 3\cdot x_1 + 4\cdot x_2 + 2\cdot x_3 + 0\cdot x_4 + 0\cdot x_5 + 1\cdot x_6 = 8 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \ge 0
\end{aligned}
$$

Calculation with tableaus – Example I.II

We commence with the basis $B(1) = 4 \wedge B(2) = 5 \wedge B(3) = 6$

$$\Rightarrow A_B = E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$ The trivial initial solution is feasible,

and the start tableau is as follows:

$$
\begin{array}{c|cccccc}
0 & -5 & -4 & -3 & 0 & 0 & 0 \\
\hline
5 & 2 & 3 & 1 & 1 & 0 & 0 \\
11 & 4 & 1 & 2 & 0 & 1 & 0 \\
8 & 3 & 4 & 2 & 0 & 0 & 1
\end{array}
\Rightarrow x = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 11 \\ 8 \end{pmatrix} \in P \quad \bar{c}^T = (-5, -4, -3, 0, 0, 0)
$$

Calculation with tableaus – Example I.III

Applying the largest coefficient rule, a further improvement is possible.

We try to introduce $t = 1$ into the basis

$\Rightarrow i_0 = 1$ since $a_1^t = \min\left\{5/2, 11/4, 8/3\right\} = 5/2$

$$\begin{array}{c|cccccc} 0 & -5 & -4 & -3 & 0 & 0 & 0 \\ \hline 5 & (2) & 3 & 1 & 1 & 0 & 0 \\ 11 & 4 & 1 & 2 & 0 & 1 & 0 \\ 8 & 3 & 4 & 2 & 0 & 0 & 1 \end{array} \Rightarrow \begin{array}{c|cccccc} 25/2 & 0 & 7/2 & -1/2 & 5/2 & 0 & 0 \\ \hline 5/2 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & -5 & 0 & -2 & 1 & 0 \\ 1/2 & 0 & -1/2 & 1/2 & -3/2 & 0 & 1 \end{array} \begin{array}{l} \cdot 5/2 \\ \cdot 1/2 \\ \cdot -4/2 \\ \cdot -3/2 \end{array}$$

$\Rightarrow x^T = \left(5/2, 0, 0, 0, 1, 1/2\right) \in P \quad \bar{c}^T = \left(0, 7/2, -1/2, 5/2, 0, 0\right) \quad z = 25/2$

Calculation with tableaus – Example I.IV

Further improvement is possible. We try to introduce $t = 3$ into the basis

$\Rightarrow i_0 = 3$ since $a_3^t = \min\{5, 1\} = 1$

$$\begin{array}{c|cccccc} 25/2 & 0 & 7/2 & -1/2 & 5/2 & 0 & 0 \\ \hline 5/2 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 \\ 1 & 0 & -5 & 0 & -2 & 1 & 0 \\ 1/2 & 0 & -1/2 & (1/2) & -3/2 & 0 & 1 \end{array} \Rightarrow \begin{array}{c|cccccc} 13 & 0 & 3 & 0 & 1 & 0 & 1 \\ \hline 2 & 1 & 2 & 0 & 2 & 0 & -1 \\ 1 & 0 & -5 & 0 & -2 & 1 & 0 \\ 1 & 0 & -1 & 1 & -3 & 0 & 2 \end{array} \begin{array}{l} \cdot 1 \\ \cdot -1 \\ \\ \cdot 2 \end{array}$$

$\Rightarrow x^T = (2, 0, 1, 0, 1, 0) \in P \quad \bar{c}^T = (0, 3, 0, 1, 0, 1) \quad z = 13$

$\Rightarrow$ Since $\bar{c} \geq 0$, the solution is optimal and the total costs are $z = c^T \cdot x = 13$.

Furthermore, $(\bar{c}_4, \bar{c}_5, \bar{c}_6) = (0, 0, 0) - \pi^T \cdot E_3 = \pi^T = (1, 0, 1)$.

Calculation with tableaus – Example II.I

We consider the following example:

Minimize $2 \cdot x_1 + 3 \cdot x_2 + x_3 + 0 \cdot x_4$ with $x_1, x_2, x_3, x_4 \geq 0$

with subject to the restrictions

$x_1 + x_2 + x_3 = 5$

$2 \cdot x_1 + x_2 + 3 \cdot x_3 - x_4 = 9$

$\Rightarrow m = 2$

We commence with the basis

$$B(1) = 1 \wedge B(2) = 2 \Rightarrow A_B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$$

Calculation with tableaus – Example II.II

$$
\begin{array}{c|cccc}
0 & 2 & 3 & 1 & 0 \\ \hline
5 & 1 & 1 & 1 & 0 \\
9 & 2 & 1 & 3 & -1
\end{array}
\Rightarrow
\begin{array}{c|cccc}
-10 & 0 & 1 & -1 & 0 \\ \hline
5 & 1 & 1 & 1 & 0 \\
-1 & 0 & -1 & 1 & -1
\end{array}
\Rightarrow
$$

$$
\begin{array}{c|cccc}
-10 & 0 & 1 & -1 & 0 \\ \hline
5 & 1 & 1 & 1 & 0 \\
1 & 0 & 1 & -1 & 1
\end{array}
\Rightarrow
\begin{array}{c|cccc}
-10 & 0 & 1 & -1 & 0 \\ \hline
4 & 1 & 0 & 2 & -1 \\
1 & 0 & 1 & -1 & 1
\end{array}
$$

$$
\Rightarrow
\begin{array}{c|cccc}
-11 & 0 & 0 & 0 & -1 \\ \hline
4 & 1 & 0 & 2 & -1 \\
1 & 0 & 1 & -1 & 1
\end{array}
\Rightarrow x = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in P
$$

Calculation with tableaus – Example II.III

$$
\begin{array}{c|cccc}
-11 & 0 & 0 & 0 & -1 \\ \hline
4 & 1 & 0 & 2 & -1 \\
1 & 0 & 1 & -1 & 1
\end{array}
\Rightarrow x = \begin{pmatrix} 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} \in P \quad \bar{c}^T = (0,0,0,-1)
$$

$\Rightarrow$ Further improvement is possible

We try to introduce $t = 4$ in the basis $\Rightarrow i_0 = 2$ since $a_1^t = -1 < 0 \wedge a_2^t = 1$

$$
\begin{array}{c|cccc}
-11 & 0 & 0 & 0 & -1 \\ \hline
4 & 1 & 0 & 2 & -1 \\
1 & 0 & 1 & -1 & (1)
\end{array}
\Rightarrow
\begin{array}{c|cccc}
-10 & 0 & 1 & -1 & 0 \\ \hline
5 & 1 & 1 & 1 & 0 \\
1 & 0 & 1 & -1 & 1
\end{array}
$$

$$
\Rightarrow x = \begin{pmatrix} 5 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in P \quad \bar{c}^T = (0,1,-1,0)
$$

Calculation with tableaus – Example II.IV

$\Rightarrow$ Further improvement is possible. We try to introduce $t = 3$ in the basis $\Rightarrow i_0 = 1$ since $a_2^t = -1 < 0 \wedge a_1^t = 1$

$$
\begin{array}{c|cccc}
-10 & 0 & 1 & -1 & 0 \\ \hline
5 & 1 & 1 & (1) & 0 \\
1 & 0 & 1 & -1 & 1
\end{array}
\Rightarrow
\begin{array}{c|cccc}
-5 & 1 & 2 & 0 & 0 \\ \hline
5 & 1 & 1 & 1 & 0 \\
6 & 1 & 2 & 0 & 1
\end{array}
$$

$$
\Rightarrow x = \begin{pmatrix} 0 \\ 0 \\ 5 \\ 6 \end{pmatrix} \in P \quad \bar{c}^T = (1,2,0,0)
$$

$\Rightarrow$ Optimal solution with total costs $Z = c^T \cdot x = 5$

Additional literature to Section 1

- Aspvall, B. and Stone, R.E. (1980): "Khachiyan's Linear Programming Algorithm,“ Journal of Algorithms 1, no. 1.

Anticycling rule:

- Bland, R. G. (1976): "New Finite Pivoting Rules," Discussion Paper 7612, Center for Operations Research and Econometrics (CORE), Universite Catholique de Louvain, Heverlee, Belgium, June 1976 (revised January 1977).

The first publication of the simplex algorithm was:

- Dantzig, G.B. (1949, 1951): "Programming of Interdependent Activities, H, Mathematical Model," pp. 19-32, in Koopmans, T.C. (ed.): *Activity Analysis of Production and Allocation.* John Wiley & Sons, Inc., New York. Also in *Econometrics* 17, nos. 3 and 4 (July-Oct. 1949), 200-211.
- Dantzig, G.B. (1951): Maximization of a linear function of variables subject to linear inequalities. In: Activity Analysis of Production and Allocation (T.C. Koopmans, ed.), Wiley, New York, pp. 359–373

Additional literature to Section 1

- Dantzig, G.B. (1963): Linear Programming and Extensions. Princeton, N.J.: Princeton University Press.
- Dantzig, G.B. (1979): Comments on Khachian's Algorithm for Linear Programming. Technical Report SOL 79-22, Dept. of Operations Research, Stanford University.
- Goldfarb, D., Todd, M.J. (1980): Modifications and Implementation of the Schor-Khachian Algorithm for Linear Programming. Department of Computer Science, Cornell University.
- Grünbaum B. (1967): *Convex Polytopes.* John Wiley & Sons, Inc., New York.
- Karmarkar, N. (1984): A new polynomial-time algorithm for Linear Programming. Combinatorica 4(4), pp. 373-395.
- Khachian, L. G. (1979): A Polynomial Algorithm for Linear Programming,“ Doklady Akad. Nauk USSR, 244, no. 5 (1979), 1093-96. Translated in Soviet Math. Doklady, 20, 191-94.

Additional literature to Section 1

- Klee, V., Minty, G.J. (1972): How Good is the Simplex Algorithm? Inequalities III, edited by O. Shisha. New York: Academic Press, Inc, pp.159-175.
- Kuhn, H. W. (1976): Class Notes, Princeton University, 1976.
- Rockafellar, R.T. (1970): *Convex Analysis.* University Press, Princeton, N.J.

Dantzig (1961) attributes the formulation of the diet problem to

- Stigler, G.J. (1945): "The Cost of Subsistence," *J. Farm Econ.,* 27, no. 2 (May 1945), pp. 303-314.