



## **Introduction to Bilevel Programming**



Vilnius University Institute of Mathematics and Informatics

Scientific youth school "High-performance computing, optimization and applications", 7-11 September, 2017, Lobachevsky State University of Nizhni Novgorod, Russia



## Outline

### Introduction Motivation Applications

### Single-level Optimization Example 1

### Multi-Objective Optimization (MOO)

Example 2 Problem formulation

### Bilevel Optimization (BO)

Example 3 Problem formulation

### Goal of this course

To be able to distinguish **bilevel problems** from **other optimization problems** with **multiple decision makers** and/or **multiple decision levels**.

- Standard Mathematical Programming (MP) models are often inadequate in the real-life because more than a single objective and one Decision Maker (DM) are involved.
- Multi-Objective Programming (MOP) deals with the extension of optimization techniques to account several objective functions.
- Game Theory (GT) deals with the mathematical models of conflict and cooperation between intelligent rational decision makers

#### Goal of this course

To be able to distinguish **bilevel problems** from **other optimization problems** with **multiple decision makers** and/or **multiple decision levels**.

- Standard Mathematical Programming (MP) models are often inadequate in the real-life because more than a single objective and one Decision Maker (DM) are involved.
- Multi-Objective Programming (MOP) deals with the extension of optimization techniques to account several objective functions.
- Game Theory (GT) deals with the mathematical models of conflict and cooperation between intelligent rational decision makers

#### Goal of this course

To be able to distinguish **bilevel problems** from **other optimization problems** with **multiple decision makers** and/or **multiple decision levels**.

- Standard Mathematical Programming (MP) models are often inadequate in the real-life because more than a single objective and one Decision Maker (DM) are involved.
- Multi-Objective Programming (MOP) deals with the extension of optimization techniques to account several objective functions.
- Game Theory (GT) deals with the mathematical models of conflict and cooperation between intelligent rational decision makers

### Goal of this course

To be able to distinguish **bilevel problems** from **other optimization problems** with **multiple decision makers** and/or **multiple decision levels**.

- Standard Mathematical Programming (MP) models are often inadequate in the real-life because more than a single objective and one Decision Maker (DM) are involved.
- Multi-Objective Programming (MOP) deals with the extension of optimization techniques to account several objective functions.
- Game Theory (GT) deals with the mathematical models of conflict and cooperation between intelligent rational decision makers



**Bilevel Programming (BP)** in a narrow sense is the combination of both.

## Bilevel Programming Applications

The world is multilevel!

Applications of Bilevel Programming are diverse and include:

- ▶ Parameter Estimation [Mitsos et al., 2008]
- Management of Multi-Divisional Firms [Ryu et al., 2004]
- Environmental Policies: Biofuel Production [Bard et al., 2000]
- ► Traffic Planning [Migdalas, 1995]
- Chemical Equilibria [Clark, 1990]
- Design of Transportation Networks [LeBlanc and Boyce, 1985]
- Agricultural Planning [Fortuny-Amat and McCarl, 1981]
- Optimisation of Strategic Defence [Bracken and McGill, 1974]
- Resource Allocation [Cassidy et al., 1971]
- Stackelberg Games: Market Economy [Stackelberg, 1934]



### Bilevel Programming Applications Seller-Buyer Strategies

### Seller-Buyer Strategies

- An owner of a company (leader) dictates the selling price and supply. He wants to maximize his profit.
- The buyers (followers) look at the product quality, pricing and various other options available to maximize their performance.



### Bilevel Programming Applications Seller-Buyer Strategies

#### Seller-Buyer Strategies

- An owner of a company (leader) dictates the selling price and supply. He wants to maximize his profit.
- The buyers (followers) look at the product quality, pricing and various other options available to maximize their performance.



### Bilevel Programming Applications Seller-Buyer Strategies

### Seller-Buyer Strategies

- An owner of a company (leader) dictates the selling price and supply. He wants to maximize his profit.
- The buyers (followers) look at the product quality, pricing and various other options available to maximize their performance.



#### Bilevel formulation

max profit (selling price, supply, demand, other variables)

s.t. upper level constraints

max performance (selling price, quality, demand)

s.t. lower level constraints

### Single-objective optimization problem Single objective function

















#### **Objective Function & Constraints**





Contour lines of the function  $\,f\,$ 



A **contour line** of a function of two variables is a curve along which the function has a **constant value** 

## Single-objective optimization problem Gradient -f & Contour lines of the function f



The gradient points in the direction of the greatest rate of increase of the function

Optimal solution:  $(x^*, y^*) = (3, 6)$ 



Optimal solution:  $(x^*, y^*) = (3, 6)$  & Optimal function value:  $f^* = -21$ 



## Multi-Objective Optimization problem Two objective functions & Constraints



SOOP

## Multi-Objective Optimization Problem (MOOP) Single vs Multi

### Single-Objective Optimization Problem:

 $\min_{\mathbf{x}\in X} f(\mathbf{x})$ 

where  $X = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{h}(\mathbf{x}) = 0\}$  - constraints

Multiple objectives are typical in real life!

- Normally objectives conflicting with each other:
  - Quality vs Cost
  - Efficiency vs Portability
- ► The scalar concept of "optimality" does not apply directly in MOO.

## Multi-Objective Optimization Problem (MOOP) Single vs Multi



Multiple objectives are typical in real life!

- Normally objectives conflicting with each other:
  - Quality vs Cost
  - Efficiency vs Portability
- ► The scalar concept of "optimality" does not apply directly in MOO.

## Multi-Objective Optimization Problem (MOOP) Single vs Multi



- Normally objectives conflicting with each other:
  - Quality vs Cost
  - Efficiency vs Portability
- ► The scalar concept of "optimality" does not apply directly in MOO.

## Multi-Objective Optimization Problem (MOOP) Single vs Multi



- Normally objectives conflicting with each other:
  - Quality vs Cost
  - Efficiency vs Portability
- ► The scalar concept of "optimality" does not apply directly in MOO.

No traditional "optimality"



No traditional "optimality"



Vector  $\mathbf{x}^2$  is **better** than  $\mathbf{x}^1$  with respect to  $f_1 : f_1(\mathbf{x}^2) < f_1(\mathbf{x}^1)$  but worse with respect to  $f_2 : f_2(\mathbf{x}^2) > f_2(\mathbf{x}^1)$ 

# Multiobjective Optimization

Two vectors can be related to each other in two ways: either one **dominates** the other or **none of them is dominated** by the other.





# Multiobjective Optimization

Two vectors can be related to each other in two ways: either one **dominates** the other or **none of them is dominated** by the other.





### Multiobjective Optimization Pareto optimality

A decision vector  $\mathbf{x}^* \in X$  which is **nondominated** by any other vector is called **Pareto optimal** and a set of those vectors is called **Pareto set**.

The set of corresponding objective vectors is called **Pareto front**.



This definition says that  $\mathbf{x}^* \in X$  is **Pareto optimal** if there exists **no feasible vector** which would **decrease** some criterion without causing a simultaneous **increase** in at least one other criterion.

The determination of these sets is the main goal of MO!

### Multiobjective Optimization Pareto optimality

A decision vector  $\mathbf{x}^* \in X$  which is **nondominated** by any other vector is called **Pareto optimal** and a set of those vectors is called **Pareto set**.

The set of corresponding objective vectors is called **Pareto front**.



This definition says that  $\mathbf{x}^* \in X$  is **Pareto optimal** if there exists **no feasible vector** which would **decrease** some criterion without causing a simultaneous **increase** in at least one other criterion.

The determination of these sets is the main goal of MO!













Pareto set and Pareto front



(Multi-Objective Optimization (MOO))

## Multi-objective optimization problem

Pareto set and Pareto front





Objective space

Remigijus Paulavičius v@RemisPau Introduction to Bilevel Programming (Multi-Objective Optimization (MOO))

## Multi-objective optimization problem

Pareto set and Pareto front



Remigijus Paulavičius v@RemisPau Introduction to Bilevel Programming (Multi-Objective Optimization (MOO))

### Multi-objective optimization problem Pareto set and Pareto front



Better Pareto front approximation!

## Bilevel optimization problem

The feasible region is implicitly determined by an inner optimization problem



Bard, Jonathan F. (1998). Practical Bilevel Optimization, Springer US

### Origins of Bilevel Programming Mathematical programming generalization

### Consider traditional mathematical programming problem

- Suppose y to be an optimal solution of nested optimization problem
- This yields the Bilevel Programming Problem (BPP)[Bracken, 1973]:

$$\begin{split} \min_{\mathbf{x}, \mathbf{y}} \, F(\mathbf{x}, \mathbf{y}) \\ \text{s.t.} \, \, \mathbf{G}(\mathbf{x}, \mathbf{y}) &\leq 0 \\ \mathbf{x} \in X \subset \mathbb{R}^n, \mathbf{y} \in Y \subset \mathbb{R}^m \end{split}$$

Mathematical programming generalization

- Consider traditional mathematical programming problem
- Suppose y to be an optimal solution of nested optimization problem
- ► This yields the Bilevel Programming Problem (**BPP**)[Bracken, 1973]:



Mathematical programming generalization

- Consider traditional mathematical programming problem
- Suppose y to be an optimal solution of nested optimization problem
- ► This yields the Bilevel Programming Problem (BPP)[Bracken, 1973]:



Bracken J, McGill J (1973) Mathematical programs with optimization problems in the constraints. Operations Research 21: 37-44.

Stackelberg game generalization

 BPP can be viewed as a static version of (non-cooperative), two-person Stackelberg game [Stackelberg, H., 1934].



- In game theory terminology, bilevel programs consist of the following rules:
  - 1. The leader (the player that moves first) is minimizing his objective function F under his resource **G**. This sub-problem is called the leader's (outer or upperlevel) problem.
  - 2. The **follower** (the **player** that moves **second**) is **minimizing** his own objective function *f* under his resource constraints **g**. This **sub-problem** is called the **follower's** (**inner** or **lower-level**) **problem**.
  - 3. The leader's decision variables are x while the follower's are y.

Stackelberg game generalization

- BPP can be viewed as a static version of (non-cooperative), two-person Stackelberg game [Stackelberg, H., 1934].
- In game theory terminology, bilevel programs consist of the following rules:



- 1. The leader (the player that moves first) is minimizing his objective function F under his resource **G**. This sub-problem is called the leader's (outer or upper-level) problem.
- 2. The **follower** (the **player** that moves **second**) is **minimizing** his own objective function *f* under his resource constraints **g**. This **sub-problem** is called the **follower's** (**inner** or **lower-level**) **problem**.
- 3. The leader's decision variables are  $\mathbf{x}$  while the follower's are  $\mathbf{y}$ .

Stackelberg game generalization

- BPP can be viewed as a static version of (non-cooperative), two-person Stackelberg game [Stackelberg, H., 1934].
- In game theory terminology, bilevel programs consist of the following rules:



- 1. The leader (the player that moves first) is minimizing his objective function F under his resource **G**. This sub-problem is called the leader's (outer or upperlevel) problem.
- 2. The **follower** (the **player** that moves **second**) is **minimizing** his own objective function *f* under his resource constraints **g**. This **sub-problem** is called the **follower's** (**inner** or **lower-level**) **problem**.
- 3. The leader's decision variables are x while the follower's are y.

Stackelberg game generalization

- BPP can be viewed as a static version of (non-cooperative), two-person Stackelberg game [Stackelberg, H., 1934].
- In game theory terminology, bilevel programs consist of the following rules:



- 1. The leader (the player that moves first) is minimizing his objective function F under his resource **G**. This sub-problem is called the leader's (outer or upperlevel) problem.
- 2. The **follower** (the **player** that moves **second**) is **minimizing** his own objective function *f* under his resource constraints **g**. This **sub-problem** is called the **follower's** (inner or lower-level) problem.
- 3. The leader's decision variables are  $\mathbf{x}$  while the follower's are  $\mathbf{y}$ .

#### 4. The **game** between the leader and the follower consists of these steps:

#### 4.1 The leader chooses a value of $\hat{\mathbf{x}} \in X$ .

- 4.2 The follower formulates a best response based on  $\hat{\mathbf{x}}$  and generates the set  $\mathcal{Y}(\hat{\mathbf{x}})$  (each  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$  is an **optimal** response for the follower of given  $\hat{\mathbf{x}}$  from the leader).
- 4.3 The leader evaluates his objective function and constraints given  $\hat{\mathbf{x}}$  and the set  $\mathcal{Y}(\hat{\mathbf{x}})$  and chooses a pair  $(\hat{\mathbf{x}}, \mathbf{y})$ ,  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$ , that results in an optimal strategy for him.
- 4.4 The leader repeats steps (4.1)-(4.3) until his objective function is minimized across all values of  $x \in X$  in his search space.
- 5. The game ends when the leader has created a strategy (value of  $\mathbf{x} \in X$ ) that minimizes his best response across different values of  $\mathbf{x} \in X$  (his decision variable).



- 4. The game between the leader and the follower consists of these steps:
  - 4.1 The leader chooses a value of  $\hat{\mathbf{x}} \in X$ .
  - 4.2 The follower formulates a best response based on  $\hat{\mathbf{x}}$  and generates the set  $\mathcal{Y}(\hat{\mathbf{x}})$  (each  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$  is an optimal response for the follower of given  $\hat{\mathbf{x}}$  from the leader).
  - 4.3 The leader evaluates his objective function and constraints given  $\hat{\mathbf{x}}$  and the set  $\mathcal{Y}(\hat{\mathbf{x}})$  and chooses a pair  $(\hat{\mathbf{x}}, \mathbf{y})$ ,  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$ , that results in an **optimal strategy** for him.
  - 4.4 The leader repeats steps (4.1)-(4.3) until his objective function is minimized across all values of  $x \in X$  in his search space.
  - 5. The game ends when the leader has created a strategy (value of  $\mathbf{x} \in X$ ) that minimizes his best response across different values of  $\mathbf{x} \in X$  (his decision variable).



4. The **game** between the leader and the follower consists of these steps:

- 4.1 The leader chooses a value of  $\hat{\mathbf{x}} \in X$ .
- 4.2 The follower formulates a best response based on  $\hat{\mathbf{x}}$  and generates the set  $\mathcal{Y}(\hat{\mathbf{x}})$  (each  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$  is an optimal response for the follower of given  $\hat{\mathbf{x}}$  from the leader).
- 4.3 The leader evaluates his objective function and constraints given  $\hat{\mathbf{x}}$  and the set  $\mathcal{Y}(\hat{\mathbf{x}})$  and chooses a pair  $(\hat{\mathbf{x}}, \mathbf{y})$ ,  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$ , that results in an optimal strategy for him.
- 4.4 The leader repeats steps (4.1)-(4.3) until his objective function is minimized across all values of  $x \in X$  in his search space.
- 5. The game ends when the leader has created a strategy (value of  $\mathbf{x} \in X$ ) that minimizes his best response across different values of  $\mathbf{x} \in X$  (his decision variable).



4. The game between the leader and the follower consists of these steps:

- 4.1 The leader chooses a value of  $\hat{\mathbf{x}} \in X$ .
- 4.2 The follower formulates a best response based on  $\hat{\mathbf{x}}$  and generates the set  $\mathcal{Y}(\hat{\mathbf{x}})$  (each  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$  is an optimal response for the follower of given  $\hat{\mathbf{x}}$  from the leader).
- 4.3 The leader evaluates his objective function and constraints given  $\hat{\mathbf{x}}$  and the set  $\mathcal{Y}(\hat{\mathbf{x}})$  and chooses a pair  $(\hat{\mathbf{x}}, \mathbf{y})$ ,  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$ , that results in an **optimal strategy** for him.
- 4.4 The leader repeats steps (4.1)-(4.3) until his objective function is minimized across all values of  $x \in X$  in his search space.
  - 5. The game ends when the leader has created a strategy (value of  $\mathbf{x} \in X$ ) that minimizes his best response across different values of  $\mathbf{x} \in X$  (his decision variable).



4. The **game** between the leader and the follower consists of these steps:

- 4.1 The leader chooses a value of  $\hat{\mathbf{x}} \in X$ .
- 4.2 The follower formulates a best response based on  $\hat{\mathbf{x}}$  and generates the set  $\mathcal{Y}(\hat{\mathbf{x}})$  (each  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$  is an optimal response for the follower of given  $\hat{\mathbf{x}}$  from the leader).
- 4.3 The leader evaluates his objective function and constraints given  $\hat{\mathbf{x}}$  and the set  $\mathcal{Y}(\hat{\mathbf{x}})$  and chooses a pair  $(\hat{\mathbf{x}}, \mathbf{y})$ ,  $\mathbf{y} \in \mathcal{Y}(\hat{\mathbf{x}})$ , that results in an optimal strategy for him.
- 4.4 The leader repeats steps (4.1)-(4.3) until his objective function is minimized across all values of  $x \in X$  in his search space.
- 5. The game ends when the leader has created a strategy (value of  $\mathbf{x} \in X$ ) that minimizes his best response across different values of  $\mathbf{x} \in X$  (his decision variable).



## Bilevel optimization problem

Nested optimization problem within the constraints of another optimization problem

Ŀ,





## Bilevel optimization problem Constraint region

$$S = \{ (\mathbf{x}, \mathbf{y}) : \mathbf{x} \in X, \mathbf{y} \in Y, \mathbf{G}(\mathbf{x}, \mathbf{y}) \le 0, \mathbf{g}(\mathbf{x}, \mathbf{y}) \le 0 \}.$$



## Bilevel optimization problem Projection of S onto the leader's decision space

$$S(X) = \{ \mathbf{x} \in X : \exists \mathbf{y} \in Y, \mathbf{G}(\mathbf{x}, \mathbf{y}) \le 0, \mathbf{g}(\mathbf{x}, \mathbf{y}) \le 0 \}.$$



## Bilevel optimization problem Follower's feasible set for each fixed $\hat{x} \in S(X)$

$$S(\mathbf{\hat{x}}) = \{ \mathbf{y} \in Y : \mathbf{g}(\mathbf{\hat{x}}, \mathbf{y}) \le 0 \}.$$



### Bilevel optimization problem Follower's rational reaction set for each $\hat{x}$

$$\mathcal{Y}(\hat{\mathbf{x}}) = \left\{ \mathbf{y} \in Y : \min \left\{ f(\hat{\mathbf{x}}, \mathbf{y}) : \mathbf{y} \in S(\hat{\mathbf{x}}) \right\} \right\}.$$

#### Defines the follower's response



## Bilevel optimization problem

The feasible set of bilevel problem (Inducible region)

$$S^{I} = \{ (\mathbf{x}, \mathbf{y}) : (\mathbf{x}, \mathbf{y}) \in S, \mathbf{y} \in \mathcal{Y}(\mathbf{x}) \}.$$

The **induced region** is usually **nonconvex** and, in the presence of **upper level constraints**, can be **disconnected** or even **empty**.



## Bilevel optimization problem

Compact BPP formulation

Given these definitions the BPP can be reformulated as:

 $\min \{F(\mathbf{x}, \mathbf{y}) : (\mathbf{x}, \mathbf{y}) \in S^{I}\}.$ 



## Bilevel optimization problem Gradient $(-\nabla F) \uparrow \&$ Contour lines: F = -2; F = -7; F = -12; F = -21



## Bilevel optimization problem Optimal solution: $(x^*, y^*) = (4, 4)$



(Bilevel Optimization (BO))

## Bilevel optimization problem

Optimal solution:  $(x^{\ast},y^{\ast})=(4,4)$  & Optimal function values:  $F^{\ast}=-12$  and  $f^{\ast}=4$ 

0

-10

-20





## Summary: Single level vs Multi-objective vs Bilevel Optimization



 Single-level problems usually have a single optimal solution.







- Bilevel problems usually also have a single optimal solution.
- A bilevel solution is not necessarily a Pareto-optimal solution.
- It is not possible to directly use algorithms for multi-objective optimization for bilevel problems.

## Summary: Single level vs Multi-objective vs Bilevel Optimization



 Single-level problems usually have a single optimal solution.



 Multi-objective problems usually have multiple optimal solutions.



- Bilevel problems usually also have a single optimal solution.
- A bilevel solution is not necessarily a Pareto-optimal solution.
- It is not possible to directly use algorithms for multi-objective optimization for bilevel problems.

## Summary: Single level vs Multi-objective vs Bilevel Optimization



 Single-level problems usually have a single optimal solution.



 Multi-objective problems usually have multiple optimal solutions.



- Bilevel problems usually also have a single optimal solution.
- > A bilevel solution is not necessarily a Pareto-optimal solution.
- It is not possible to directly use algorithms for multi-objective optimization for bilevel problems.

## Thank You For Your Attention!

# Questions ?





remigijus.paulavicius@mii.vu.lt