An Introduction to Bilevel Programming

Chris Fricke

Department of Mathematics and Statistics

University of Melbourne

Outline

- What is Bilevel Programming?
- Origins of Bilevel Programming.
- Some Properties of Bilevel Programs.
- The Linear Bilevel Programming Problem.
- Applications.
- References.

What is Bilevel Programming?

What is Bilevel Programming?

- 'A mathematical program that contains an optimization problem in the constraints.' *
- Evolved in two ways:
  - A logical extension of mathematical programming.
  - Generalisation of a particular problem in game theory (Stackelberg Game).

*Bracken and McGill, *Operations Research*, Vol. 21 (1973)

General Bilevel Programming Problem (Bard, 1998)

$$\min_{x \in X} F(x, y)$$

s.t.

$$G(x, y) \leq 0$$

$$\min_{y \in Y} f(x, y)$$

s.t.

$$g(x, y) \leq 0$$

$$x, y \geq 0$$

Origins of Bilevel Programming

- Game Theory Approach
- Mathematical Programming Approach

Origins of Bilevel Programming

- Stackelberg
  - (*The Theory of the Market Economy*, Oxford University Press, 1952).
- Bracken and McGill
  - ("Mathematical Programs with Optimization Problems in the Constraints", *Operations Research* Vol. 21 No. 1, 1973).

Game Theory Approach

Elementary Game Theory

- Situations involving more than one decision maker (player).
- Leads to the notion of *conflict* or *competition*.
- Each player has a number of available strategies with corresponding payoffs.
- Simplest case: Two-person, zero-sum game.

Two-person Zero-sum Game: Assumptions

- One player's gain = other player's loss.
- Players make decisions simultaneously.
- Players have perfect information
  – Of both their own and their opponent's permissible strategies and corresponding payoff.
- Players do not cooperate.

Two-person Zero-sum Game

- Payoff matrix:

Player II

$$\begin{array}{cc} & \begin{array}{cccc} A_1 & A_2 & \dots & A_n \end{array} \\ \text{Player I} \quad \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_m \end{array} & \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} \end{array}$$

Stackelberg Game: Assumptions

- One player's gain $\neq$ the other player's loss.
- Players make decisions in specified order.
- Second player reacts rationally to first player's decision.
- Players have perfect information
  – Of both their own and their opponent's permissible strategies and consequent payoff.
- Players do not cooperate.

Stackelberg Game

- Payoff matrix:

(Bimatrix Game)

$$
\begin{array}{c c}
 & \begin{array}{cccc} A_1 & A_2 & \cdots & A_n \end{array} \\
\begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_m \end{array} &
\begin{bmatrix}
(v_{11}, w_{11}) & (v_{12}, w_{12}) & \cdots & (v_{1n}, w_{1n}) \\
(v_{21}, w_{21}) & (v_{22}, w_{22}) & \cdots & (v_{2n}, w_{2n}) \\
\vdots & \vdots & & \vdots \\
(v_{m1}, w_{m1}) & (v_{m2}, w_{m2}) & \cdots & (v_{mn}, w_{mn})
\end{bmatrix}
\end{array}
$$

Player I Payoff

Player II Payoff

Stackelberg Game: Definitions

- Player who moves first is called the LEADER.
- Player who reacts (rationally) to the leader's decision (strategy) is called the FOLLOWER.
- The actions of one affect the choices and payoffs available to the other, and vice-versa.

Stackelberg Game: Solution Methods

- Small instances can be solved analytically (standard game theory techniques
  e.g. graphical method).
- How to analyze/solve when each player has many available strategies?
- How to incorporate a complex relationship between the strategies and payoffs?

Extension to Bilevel Programming

- Also allows additional constraints to be placed on the player's strategies.
- Mathematical programming viewpoint:
  - **LEADER** moves first and attempts to minimize their own objective function.
  - **FOLLOWER** observes the leader's action and moves in a way that is personally optimal.

Mathematical Programming Approach

Some Distinctions

- General mathematical program:

$$\min_x f(x)$$

$$\text{s.t.}$$

$$Ax \leq b$$

$$x \geq 0$$

Some Distinctions

- Multiple objective program:

$$\min_x f(x)$$

$$\min_x g(x)$$

s.t.

$$Ax \leq b$$

$$x \geq 0$$

Some Distinctions

- Bilevel program:

$$\begin{aligned} \min_x \; & f(x,y) \\ \text{s.t.} \; & \\ & A(x,y) \le b \\ & \min_y \; g(x,y) \\ & \text{s.t.} \\ & \quad C(x,y) \le d \\ & \quad x, y \ge 0 \end{aligned}$$

General Bilevel Programming Problem (BLPP)

$$\min_{x \in X} F(x, y)$$

s.t.

$$G(x, y) \leq 0$$

$$\min_{y \in Y} f(x, y)$$

s.t.

$$g(x, y) \leq 0$$

$$x, y \geq 0$$

Properties of Bilevel Programs

- Existence of Solutions
- Order of Play

Existence of Solutions

- A BLPP need not have a solution.
- Restricting the functions F, G, f, g to be continuous and bounded DOES NOT guarantee the existence of a solution.

Example (Bard, 1998)

$$\min_{\mathbf{x}} \left\{ F = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : \begin{array}{l} x_1 \geq 0, x_2 \geq 0 \\ x_1 + x_2 = 1 \end{array} \right\}$$

s.t.

$$\min_{\mathbf{y}} \left\{ f = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -1 & -4 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} : \begin{array}{l} y_1 \geq 0, y_2 \geq 0 \\ y_1 + y_2 = 1 \end{array} \right\}$$

Example 1

- The solution $\hat{y}$ to the follower's problem as a function of x is:

$$\hat{\mathbf{y}}(\mathbf{x}) = \begin{cases} (1,0) & \text{for} \quad x_1 + 3x_2 > 4x_1 + 2x_2; \quad \text{i.e.} \quad x_1 < \frac{1}{4} \\ y_1 + y_2 = 1 & \text{for} \quad x_1 = \frac{1}{4} \\ (0,1) & \text{for} \quad x_1 > \frac{1}{4} \end{cases}$$

Example 1

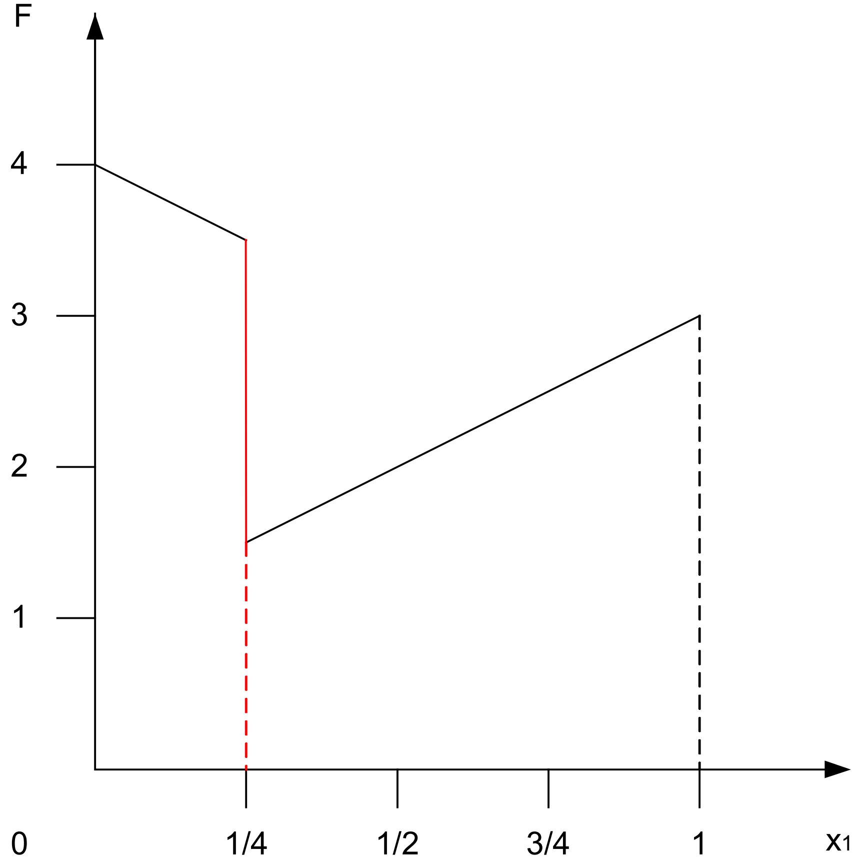
- Substituting these values into the leader's problem gives:

$$\min_{\mathbf{x}} F = \begin{cases} 2x_1 + 4x_2 & ; x_1 < \frac{1}{4} \\ 2y_1 + \frac{3}{2} \qquad (0 \leq y_1 \leq 1) & ; x_1 = \frac{1}{4} \\ 3x_1 + x_2 & ; x_1 > \frac{1}{4} \end{cases}$$

$$\text{s.t.} \qquad x_1 + x_2 = 1; \quad x_1 \geq 0; \quad x_2 \geq 0$$

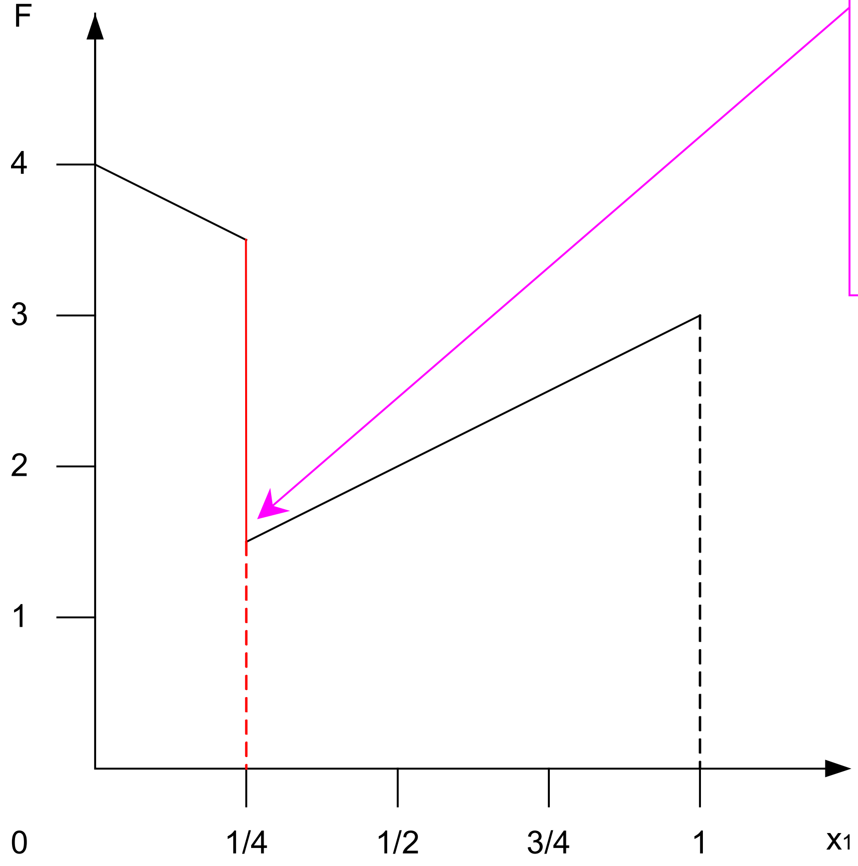
Example 1

- Solution space for leader's problem:



Example 1

- Leader's optimal solution:



$$F = 1.5$$
$$x = (\tfrac{1}{4}, \tfrac{3}{4})$$
$$y = (0, 1)$$

Example 1

- At $x = (\frac{1}{4}, \frac{3}{4})$ follower's optimal solution is $f = 1$ at any point on the line $y_1 + y_2 = 1$
- Corresponding solution for leader is
$$F = 2y_1 + \frac{3}{2} \quad \Rightarrow \quad F \in [1.5, 3.5]$$
- No way for the leader to guarantee they achieve their minimum payoff

$\Rightarrow$ No solution.

Order of Play

- The order in which decisions are made is important.
- The roles of leader and follower are NOT interchangeable (problem is not symmetric).

Example 2

- Reverse the previous example:

$$\min_{\mathbf{y}} f = \quad -(x_1 + 3x_2)y_1 \quad -(4x_1 + 2x_2)y_2$$

$$\text{s.t.}$$

$$y_1 + y_2 = 1; \qquad y_1 \geq 0, \quad y_2 \geq 0$$

$$\min_{\mathbf{x}} F = \qquad (2y_1 + 3y_2)x_1 + (4y_1 + y_2)x_2$$

$$\text{s.t.}$$

$$x_1 + x_2 = 1; \quad x_1 \geq 0; \quad x_2 \geq 0$$

Example 2

- The solution $\hat{x}$ to the follower's problem as a function of y is:

$$\hat{\mathbf{x}}(\mathbf{y}) = \begin{cases} (1,0) & \text{for} \quad 2y_1 + 3y_2 < 4y_1 + y_2; \quad \text{i.e.} \quad y_1 > \frac{1}{2} \\ x_1 + x_2 = 1 & \text{for} \quad y_1 = \frac{1}{2} \\ (0,1) & \text{for} \quad y_1 < \frac{1}{2} \end{cases}$$

Example 2

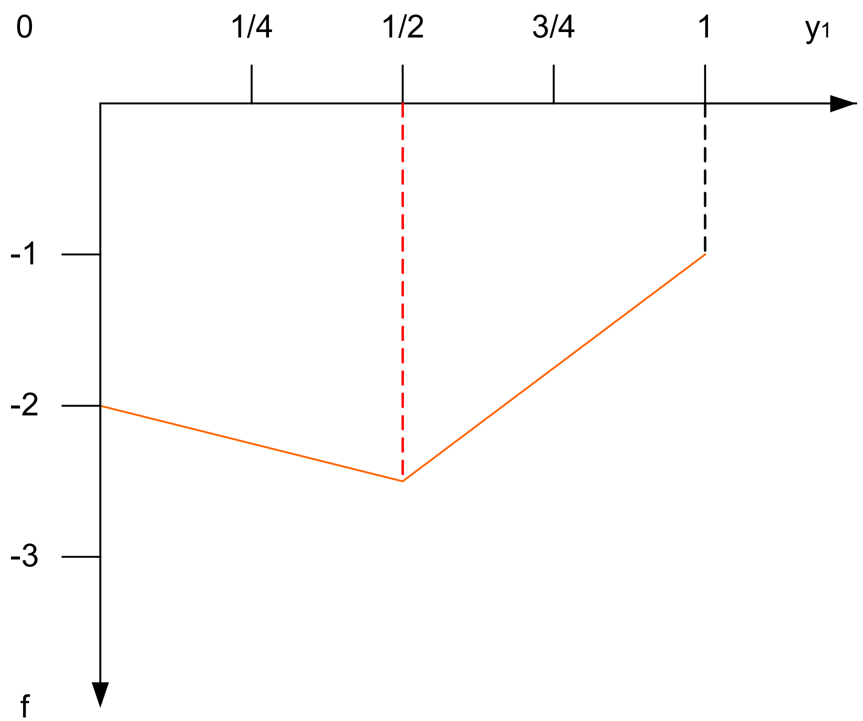
- Substituting these values into the leader's problem gives:

$$\min_{\mathbf{y}} f = \begin{cases} -y_1 - 4y_2 & ; y_1 > \frac{1}{2} \\ -(3 - 2x_1)y_1 - (2x_1 + 2)y_2 & ; y_1 = \frac{1}{2} \\ -3y_1 - 2y_2 & ; y_1 < \frac{1}{2} \end{cases}$$

$$\text{s.t.} \qquad y_1 + y_2 = 1; \quad y_1 \geq 0; \quad y_2 \geq 0$$

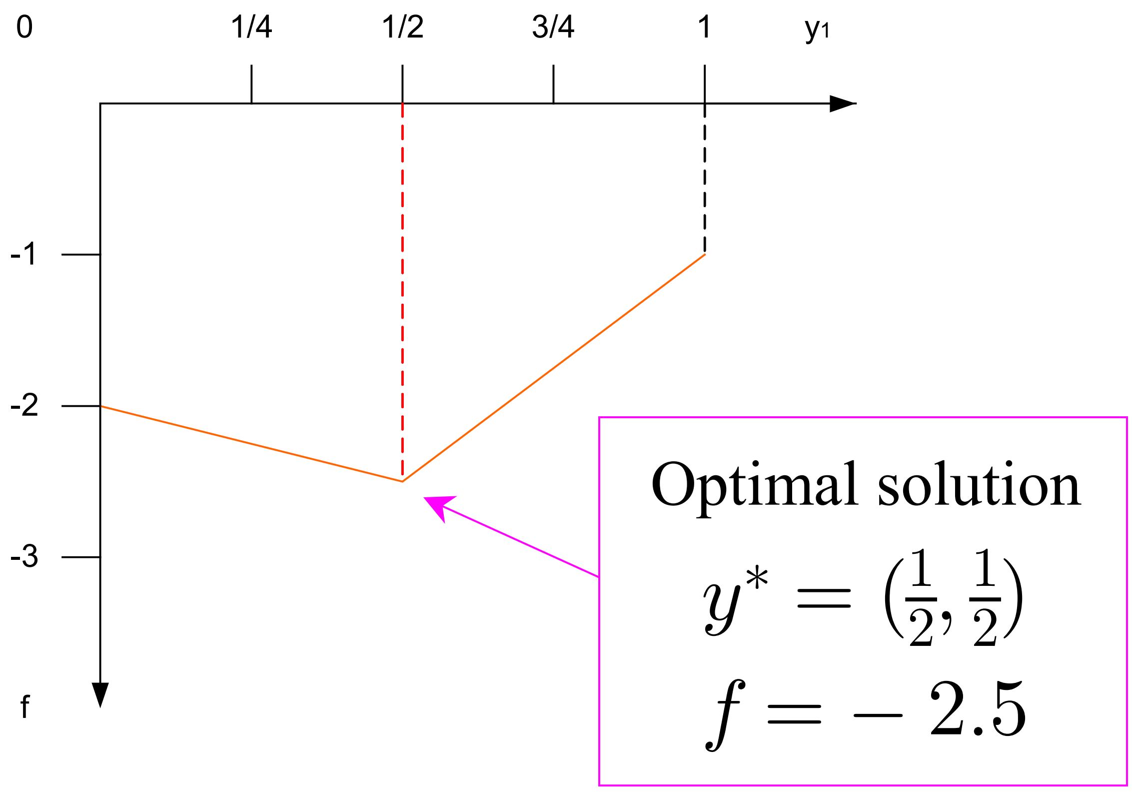
Example 2

- Solution space for leader's problem:



Example 2

- Solution space for leader's problem:



Example 2

- At $y^* = (\frac{1}{2}, \frac{1}{2})$ follower's optimal solution is $F = 2.5$ at any point on the line $x_1 + x_2 = 1$
- Comparison of solutions:

	Example 1	Example 2	Nash Equilibrium
Solution ($\mathbf{x}$)	$(\frac{1}{4}, \frac{3}{4})$	$x_1 + x_2 = 1$	$(\frac{1}{4}, \frac{3}{4})$
Cost (F)	1.5	2.5	2.5
Solution ($\mathbf{y}$)	$(0, 1)$	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2})$
Cost (f)	-2.5	-2.5	-2.5

The Linear Bilevel Programming Problem (LBLPP)

General Bilevel Programming Problem (Bard, 1998)

$$\min_{x \in X} F(x, y)$$

s.t.

$$G(x, y) \leq 0$$

$$\min_{y \in Y} f(x, y)$$

s.t.

$$g(x, y) \leq 0$$

$$x, y \geq 0$$

General LBLPP

$$\min_{x \in X} F(x, y) = c_1 x + d_1 y$$

s.t.

$$A_1 x + B_1 y \leq b_1$$

$$\min_{y \in Y} f(x, y) = c_2 x + d_2 y$$

s.t.

$$A_2 x + B_2 y \leq b_2$$

Definitions

- Constraint region of the BLPP:

$$S = \{(x, y) : x \in X, y \in Y, A_1 x + B_1 y \leq b_1, \\ A_2 x + B_2 y \leq b_2\}$$

- Follower's feasible set for each fixed $x \in X$:

$$S(x) = \{y \in Y : B_2 y \leq b_2 - A_2 x\}$$

- Follower's rational reaction set:

$$P(x) = \{y \in Y : y \in argmin[f(x, \hat{y}) : \hat{y} \in S(x)]\}$$

Definitions

- Inducible Region:

$$IR = \{(x, y) \in S, y \in P(x)\}$$

- When S and $P(x)$ are non-empty, the BLPP can be written as:

$$\min\{F(x, y) : (x, y) \in IR\}$$

Example (Bard, 1998)

$$\min_{x \geq 0} \quad F(x, y) = x - 4y$$

s.t.

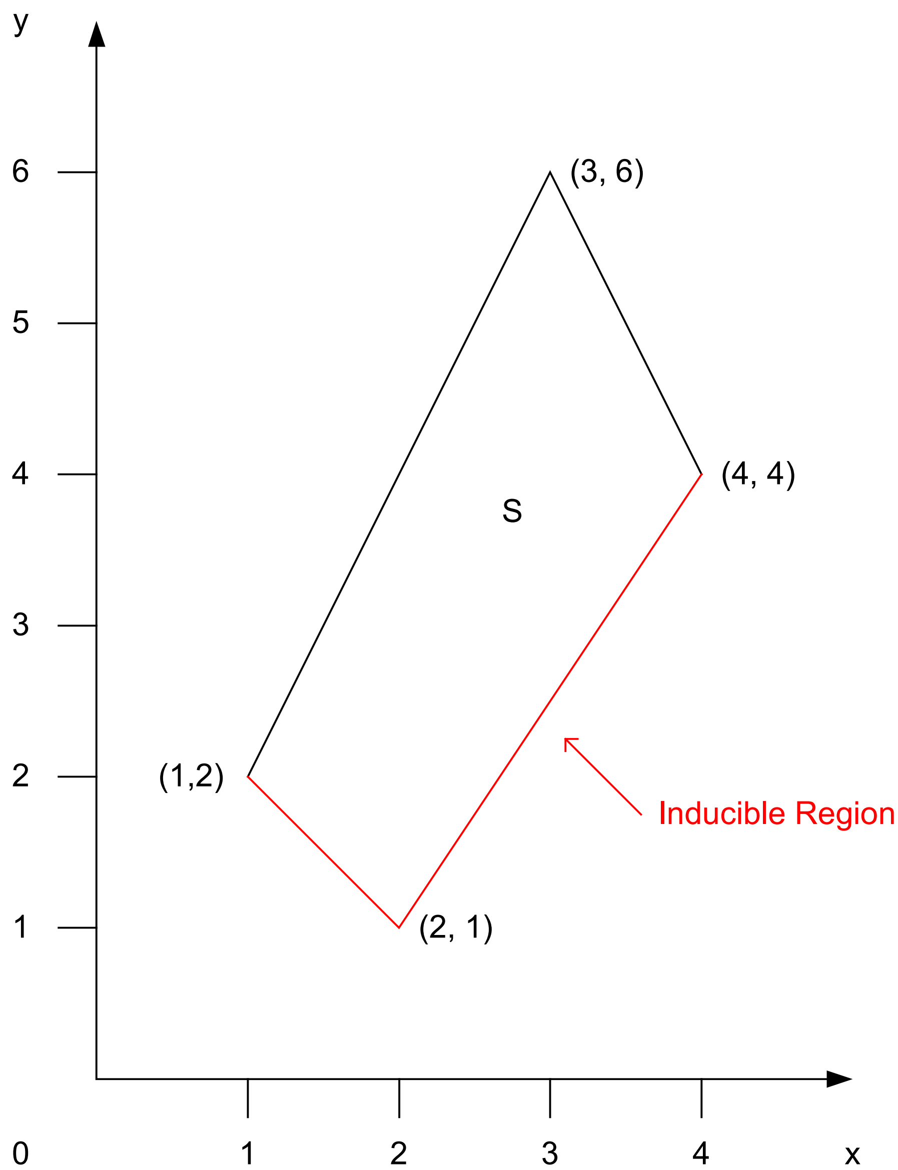
$$\min_{y \geq 0} \quad f(y) = y$$

s.t.

$$\begin{aligned} -x - y &\leq -3 \\ -2x + y &\leq 0 \\ 2x + y &\leq 12 \\ -3x + 2y &\leq -4 \end{aligned}$$

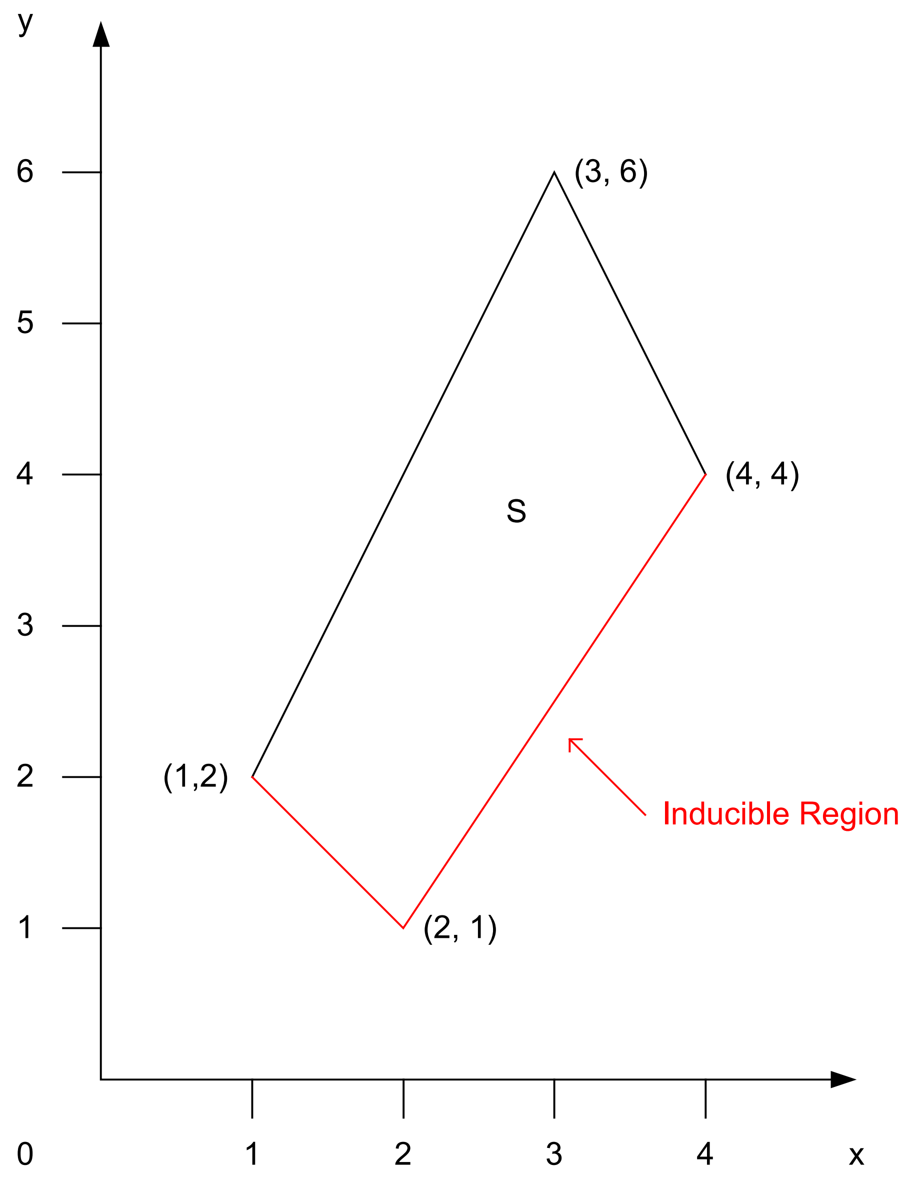
$$F(x,y) = x - 4y$$

$$f(y) = y$$



$$F(x, y) = x - 4y$$

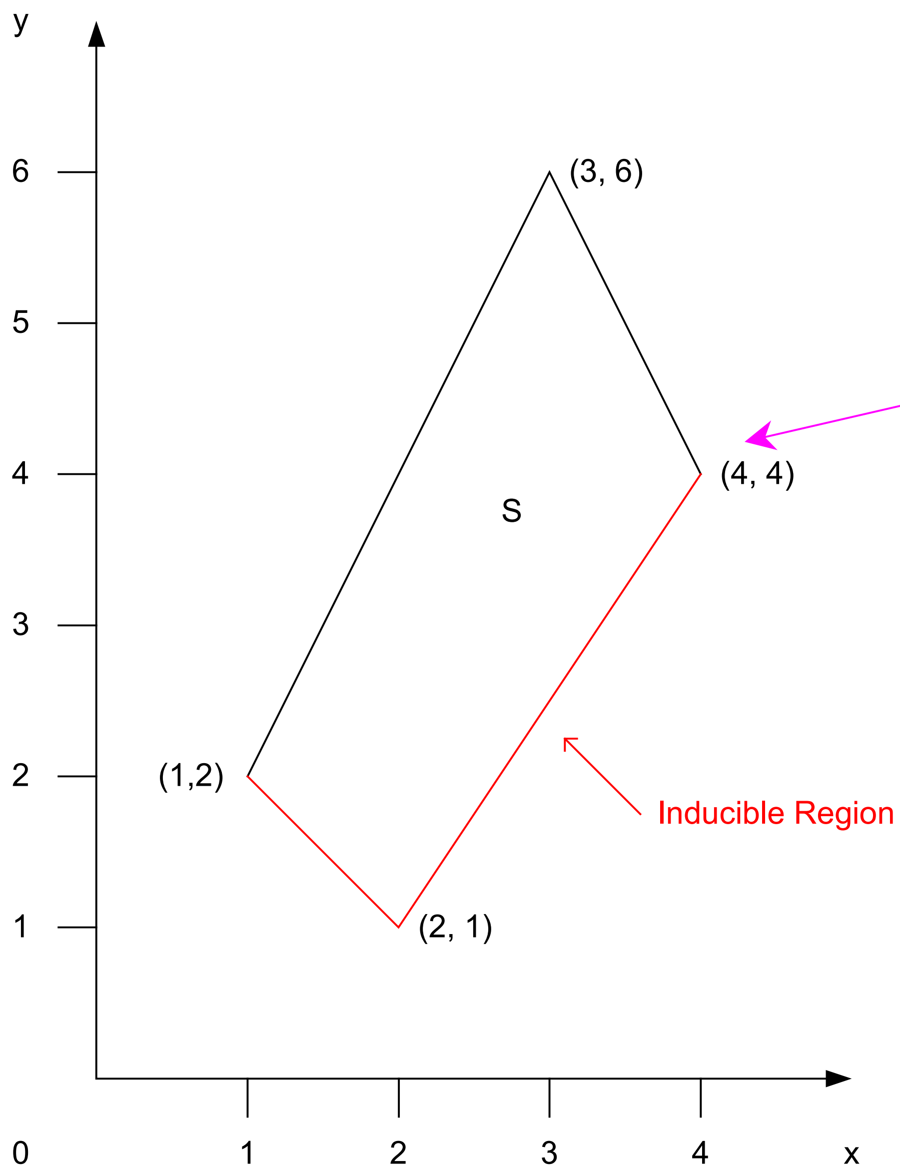
$$f(y) = y$$



Non-convex in general

$$F(x, y) = x - 4y$$

$$f(y) = y$$



Optimal solution

$$x^* = 4$$

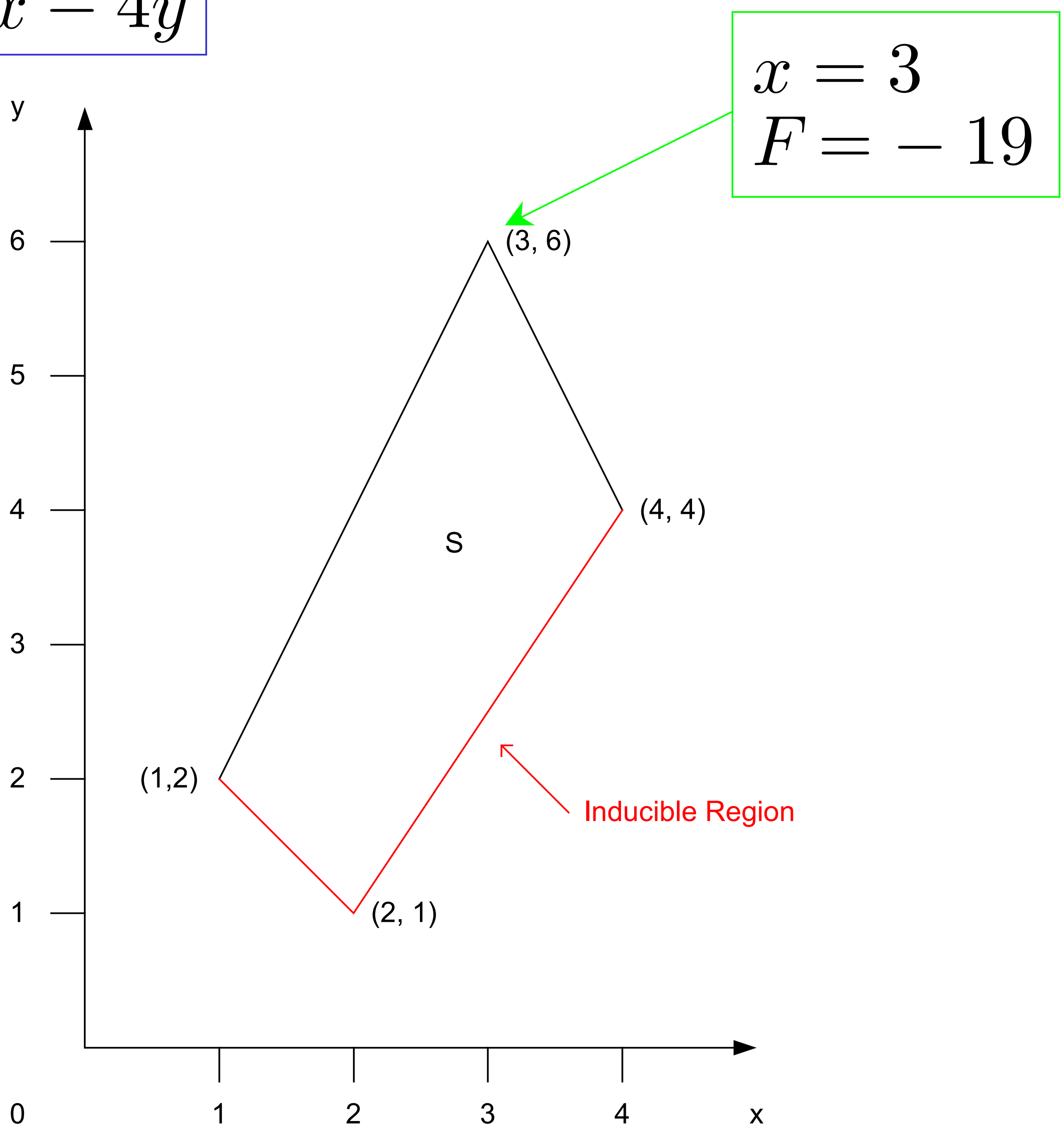
$$y^* = 4$$

$$F^* = -12$$

$$f^* = 4$$

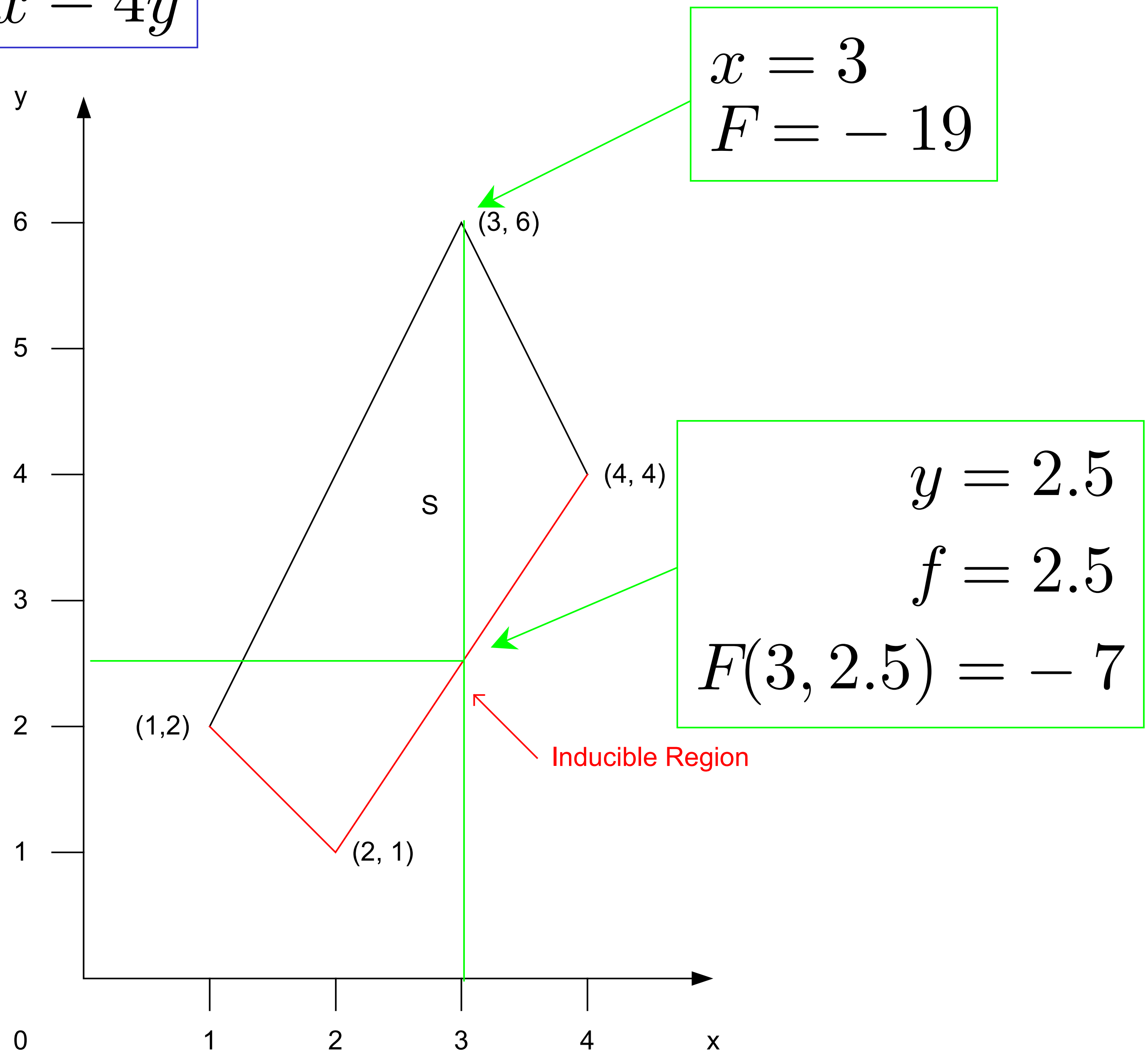
$$\boxed{F(x, y) = x - 4y}$$

$$\boxed{f(y) = y}$$



$$\boxed{F(x, y) = x - 4y}$$

$$\boxed{f(y) = y}$$

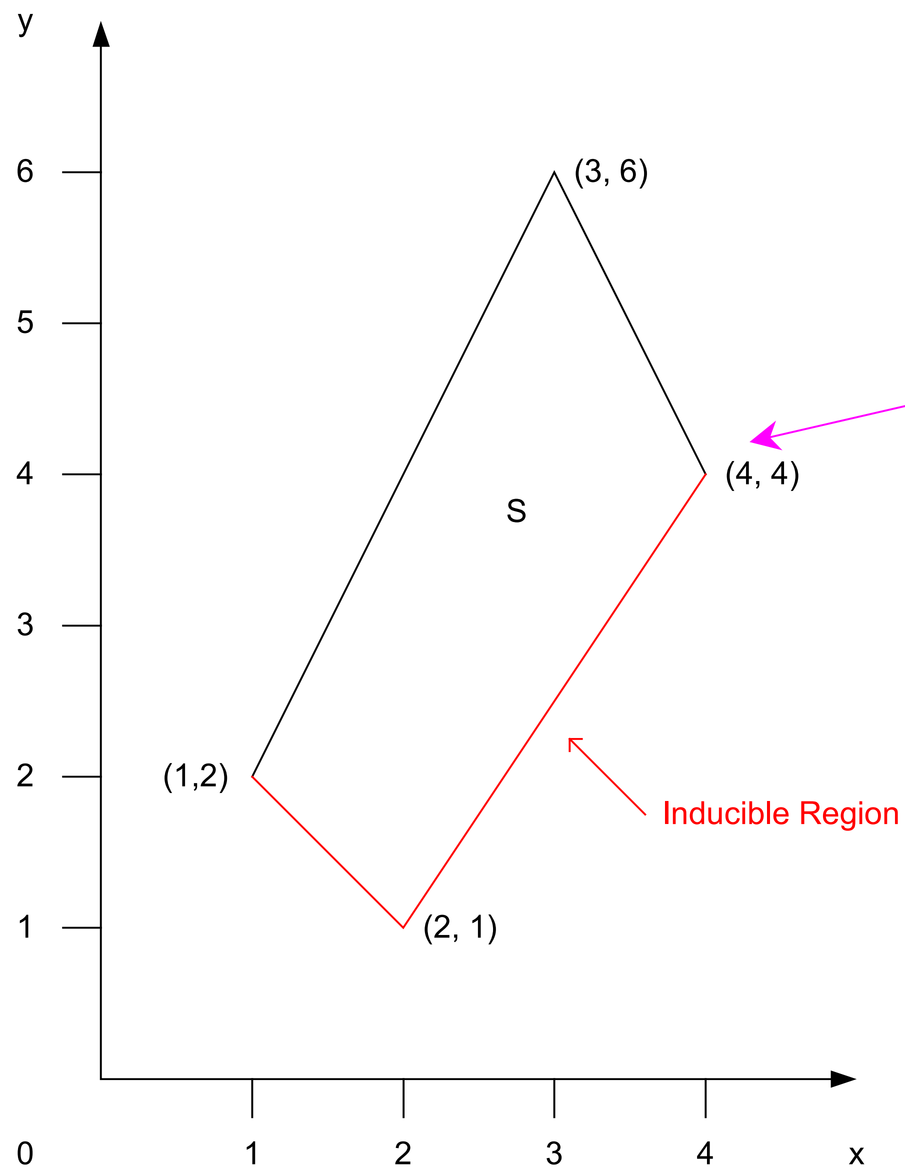


Pareto Optimality

- Multiple objective problem.
- Feasible solution B dominates feasible solution A:
  - B at least as good as A w.r.t. every objective,
  - B strictly better than A w.r.t. at least one objective.
- Pareto optimal solutions:
  - Set of all non-dominated feasible solutions.

$$F(x,y) = x - 4y$$

$$f(y) = y$$



Optimal solution

$$x^* = 4$$

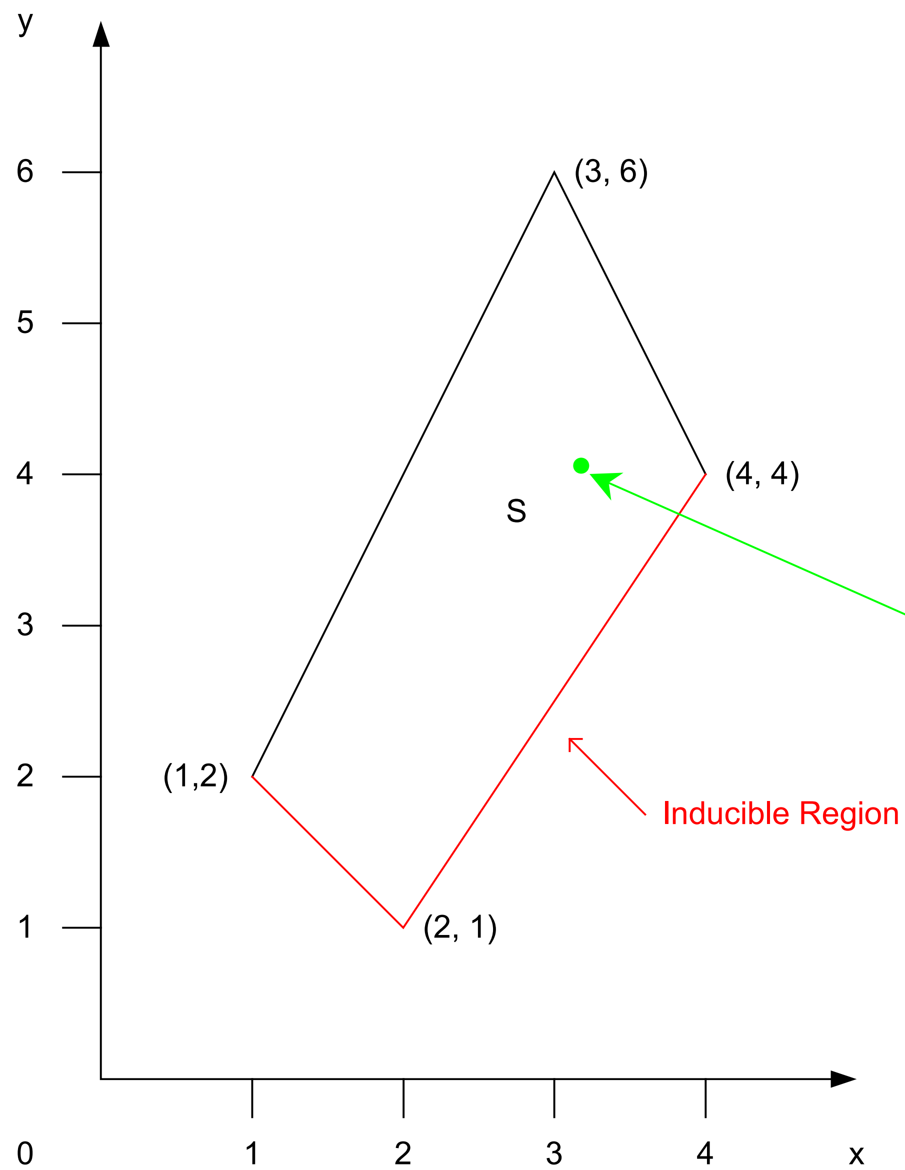
$$y^* = 4$$

$$F^* = -12$$

$$f^* = 4$$

$$\boxed{F(x,y) = x - 4y}$$

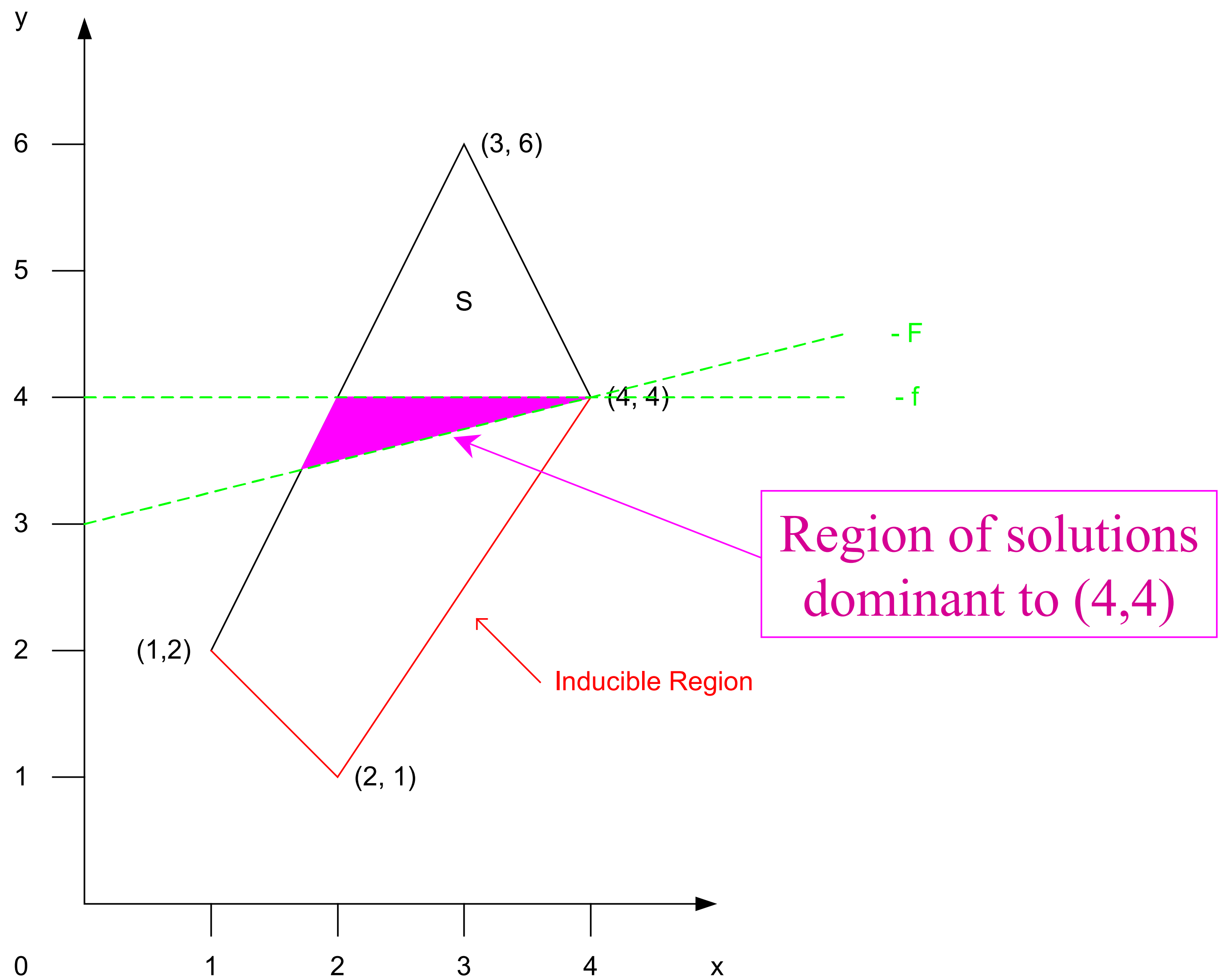
$$\boxed{f(y) = y}$$



$$x = 3$$
$$y = 4$$
$$\boxed{F = -13}$$
$$f = 4$$

$$F(x, y) = x - 4y$$

$$f(y) = y$$


y
6
5
4
3
2
1
0
1
2
3
4
x
(3, 6)
S
(4, 4)
(1,2)
(2, 1)
- F
- f
Inducible Region
Region of solutions dominant to (4,4)

Summary of Properties of Bilevel Programs

- No guarantee of solution.
- Order in which decisions are made is important.
- No guarantee of Pareto optimality.
- Non-convex optimization problem.
- All of the above can apply even when all functions are continuous and bounded.

Solution Methods for the LBLPP

- Vertex Enumeration
- Penalty Methods
- KKT Conditions

Represent the IR as a Piecewise Linear Function*

- **Theorem 1:**
  - The IR can be written equivalently as a piecewise linear equality constraint comprised of supporting hyperplanes of S.
- **Theorem 2:**
  - The solution (x^*, y^*) of the LBLPP occurs at a vertex of S.

*Bard, *Practical Bilevel Optimization*, (1998)

Solution Methods

- These Theorems are exploited in vertex enumeration algorithm of Candler and Townsley*.
- Approach based on Simplex method.
- Various penalty methods also developed.
- Most common approach via use of KKT conditions.

*Candler and Townsley, *Computers and Operations Research*, Vol. 9 (1982)

Karush-Kuhn-Tucker Conditions

- Given $x = \hat{x}$ the KKT conditions for a local optimum at the point y^* for

$$\begin{aligned} \min_y \quad & f(\hat{\mathbf{x}}, \mathbf{y}) \\ & g(\hat{\mathbf{x}}, \mathbf{y}) \geq 0 \quad (\mu) \end{aligned}$$

are:

$$\begin{aligned} \nabla_y f(\hat{\mathbf{x}}, \mathbf{y}^*) - \mu^{\mathbf{t}} \nabla_y g(\hat{\mathbf{x}}, \mathbf{y}^*) &= 0 \\ \mu^{\mathbf{t}} g(\hat{\mathbf{x}}, \mathbf{y}^*) &= 0 \\ \mu &\geq 0 \end{aligned}$$

Replace the Follower's Problem with its KKT Conditions

- Recall the follower's problem:

$$
\begin{aligned}
\min_{\mathbf{y} \in Y} \quad & f(\mathbf{x}, \mathbf{y}) = \mathbf{c_2 x} + \mathbf{d_2 y} \\
\text{s.t.} \quad & \\
& A_2 \mathbf{x} + B_2 \mathbf{y} \leq \mathbf{b_2} \\
& \mathbf{y} \geq 0
\end{aligned}
$$

KKT Conditions

- Let **u** and **v** be the dual variables associated with the two sets of constraints.

$$
\begin{aligned}
\min_{\mathbf{y} \in Y} \quad & f(\mathbf{x}, \mathbf{y}) = \mathbf{c_2 x} + \mathbf{d_2 y} \\
\text{s.t.} \quad & \\
& \mathbf{b_2} - A_2\mathbf{x} - B_2\mathbf{y} \geq 0 \quad (\mathbf{u}) \\
& \mathbf{y} \geq 0 \quad (\mathbf{v})
\end{aligned}
$$

KKT Conditions

- Applying the KKT conditions to the follower's problem gives:

$$
\begin{aligned}
\mathbf{d_2} + \mathbf{u}B_2 - \mathbf{v} &= 0 \\
\mathbf{u}(\mathbf{b_2} - A_2\mathbf{x} - B_2\mathbf{y}) + \mathbf{v}\mathbf{y} &= 0 \\
\mathbf{u}, \mathbf{v} &\geq 0
\end{aligned}
$$

Reformulation

- Solve for **x**, **y**, **u**, **v**

$$\begin{aligned}
\min_{\mathbf{x}\in X} \quad & \mathbf{c_1}\mathbf{x} + \mathbf{d_1}\mathbf{y} \\
\text{s.t.} \quad & \\
& A_1\mathbf{x} + \mathbf{B_1}\mathbf{y} && \leq \mathbf{b_1} \\
& \mathbf{u}B_2 - \mathbf{v} && = -\mathbf{d_2} \\
& \mathbf{u}(\mathbf{b_2} - A_2\mathbf{x} - B_2\mathbf{y}) + \mathbf{v}\mathbf{y} && = 0 \\
& A_2\mathbf{x} + B_2\mathbf{y} && \leq \mathbf{b_2} \\
& \mathbf{x} \geq 0, \quad \mathbf{y} \geq 0, \quad \mathbf{u} \geq 0, \quad \mathbf{v} \geq 0
\end{aligned}$$

Reformulation

- Complementarity condition is non-linear.

$$\mathbf{u}(\mathbf{b_2} - A_2\mathbf{x} - B_2\mathbf{y}) + \mathbf{v}\mathbf{y} \qquad = 0$$

- Requires implementation of non-linear programming methods.
- Problem is further complicated if IR is not convex.

Integer Programming Approach to Dealing with Non-Linearity

- Write all inequalities in follower's problem in the form:

$$g_i(x, y) \geq 0$$

- Complementary slackness gives:

$$u_i \, g_i(x, y) = 0$$

- Introduce binary variables $z_i \in \{0, 1\}$ and sufficiently large constant M

Integer Programming Approach to Dealing with Non-Linearity*

- Replace complementary slackness with the following inequalities:

$$u_i \leq M z_i, \quad g_i \leq M(1 - z_i)$$

- Now have a mixed-integer linear program

  → Apply standard IP solvers (e.g. CPLEX).

- Drawback: increased number of variables and constraints

  → Computational inefficiency.

*Fortuny-Amat and McCarl, *Journal of the Operational Research Society*, Vol. 32 (1981)

Alternative Approach to Dealing with Non-Linearity

- Rewrite complementary slackness term as sum of piecewise linear separable functions.
- Use globally convergent non-linear code to obtain solutions.
- Basis for Bard-Moore (branch-and-bound) algorithm for solving the LBLPP*

$$\sum_i u_i \, g_i(x, y) = 0$$

*Bard and Moore, *SIAM Journal of Scientific and Statistical Computing*, Vol. 11 (1990)

$$\sum_i u_i \, g_i(x,y) = 0$$

$$\Rightarrow \quad \sum_i \min\{u_i, \; g_i\} = 0$$

$$\Rightarrow \quad \sum_i (\min\{0, \; (g_i - u_i)\} + u_i) = 0$$

- Replace $g_i - u_i$ with new variables w_i to give equivalent set of constraints:

Piecewise linear separable term

$$\sum_i (\min\{0, \; w_i\} + u_i) = 0$$

Linear equalities

$$w_i - g_i + u_i = 0$$

Applications

- Economics
- Resource Allocation
- Transportation Network Design

Applications

- Multilevel systems
  - A high level decision maker is able to influence the decisions made at lower levels, without having complete control over their actions.
  - Objective function of one department is determined, in part, by variables controlled by other departments operating at higher or lower levels.

Economic Planning at the Regional or National Level

- Leader: Government
  - Controls policy variables e.g. tax rates, import quotas.
  - Maximize employment / Minimize use of a resource.
- Follower: Industry to be regulated
  - Maximize net income s.t. economic and governmental constraints.

Determining Price Support Levels for Biofuel Crops

- Leader: Government
  - Determine the level of tax credits for each biofuel product.
  - Minimize total outlays.
- Follower: Petro-chemical industry
  - Minimize costs.

Resource Allocation in a Decentralized Firm

- Leader: Central resource supplier
  - Allocates products to manufacturers.
  - Maximize profit of firm as a whole.
- Follower: Manufacturing facilities at different locations
  - Determines own production mix/output.
  - Maximize performance of own unit.

Transportation System Network Design

- Leader: Central planner
  - Controls investment costs e.g. which links to improve.
  - Influence users' preferences to minimize total costs.
- Follower: Individual users
  - Their route selection determines the traffic flows and therefore operational costs.
  - Seek to minimize cost of own route.

Summary of Bilevel Programming Problems

$$\min_{x \in X} F(x, y)$$

s.t.

$$G(x, y) \leq 0$$

$$\min_{y \in Y} f(x, y)$$

s.t.

$$g(x, y) \leq 0$$

$$x, y \geq 0$$

Summary of Bilevel Programming Problems

- No guarantee of solution.
- Order in which decisions are made is important.
- No guarantee of Pareto optimality.
- Non-convex optimization problem.
- In linear case, a number of possible reformulations exist to aid solution.
- Used for specific applications.

References

- Bard, J.F. *Practical Bilevel Optimization: Applications and Algorithms*, Kluwer Academic Press, 1998.
- Shimizu, K., Ishizuka, Y. and Bard, J.F., *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic Publishers, 1997.
- http://www.acsu.buffalo.edu/~bialas/