What is a MIP model

Outline

Data, variables, objective, constraints

Why integrality matters

The LP relaxation and the sandwich of bounds

Three patterns that keep coming back

What the solver does with the two bounds

What this chapter has put in your hands

Learning objectives (1/2)

At the end of this chapter you will be able to:

- write a model with its data, variables, objective function, constraints and domains;
- tell the classes LP, ILP, BIP and MILP apart from the domains of the variables;
- explain with an example why rounding the LP optimum does not solve an ILP;

Learning objectives (2/2)

At the end of this chapter you will be able to:

- say on which side of $z(MILP)$ a feasible solution and a relaxation lie, and compute the gap;
- recognise set partitioning, set packing and set covering constraints.

Prerequisites

- From the module *Integer Linear Programming*, Chapter 1 «Integer linear programming»: integer linear programming problems, their integer feasible points and the LP relaxation
- From the module *Linear Programming*, Chapter 4 «Linear Programming problems»: the form of a Linear Programming problem: objective function, linear constraints and domains of the variables

Data, variables, objective, constraints

Introduction: Data, variables, objective, constraints

- We see which ingredients a **mathematical programming** model is made of
- We tell apart the four classes of models: **LP**, **ILP**, **BIP** and **MILP**
- We fix the notation for optimal values, solutions and **bounds**, used throughout the course

The four ingredients of a model

A model translates a decision into four ingredients, always in the same order:

1. the **data**: the numbers describing the problem, known before deciding (costs, capacities, demands, times. . .)
2. the **variables**: the quantities the decision maker controls
3. the **objective**: a function of the variables to minimise or maximise
4. the **constraints**: equations and inequalities the variables must satisfy for the solution to be feasible

An *optimisation* problem asks for the feasible solution (satisfying all constraints) with the best objective value.

The classes of models

- **The acronyms**: upright LP is the *class* of problems, italic *LP* is *one* problem of that class, the one whose $\tilde{x}(LP)$ and $z(LP)$ are written

  The same for ILP, BIP and MILP

- **The size and the data**:
  - n variables, with $j \in \{1, 2, \ldots, n\}$
  - m constraints, with $i \in \{1, 2, \ldots, m\}$
  - the costs c_j, the coefficients a_{ij} and the right-hand sides b_i

What a model can contain

- a **linear objective function** of the variables, to be minimised or maximised
- **linear constraints** of three types:
  - $\geq$, for a requirement to be met
  - $\leq$, for a capacity to be respected
  - equality, for a balance or an exact assignment
- **variables**, in one or more families, each with its own domain, declared at the end of the model

None of the three has a fixed shape: the problem decides.

What tells the four classes apart: the domains

- **LP**: all the variables are continuous, $x_j \in \mathbb{R}$, usually non-negative

  In this course an LP always arrives as the *relaxation* of a MIP model, when the integrality constraints are dropped

- **ILP**: all the variables are integer, $x_j \in \mathbb{Z}$
- **BIP**: all the variables are binary, $x_j \in \{0, 1\}$
- **MILP**: some variables integer or binary, others continuous

That is where the difference in difficulty lives: an LP is solved in polynomial time, an ILP and a MILP in general are not.

Example: a model of ILP type

How many units to buy from two suppliers, x_1 and x_2, in whole lots.

$$\begin{array}{rrcrcl} \min & 4x_1 & + & 7x_2 & & \\ \text{subject to} & 2x_1 & + & 3x_2 & \geq & 12 \\ & x_1 & + & x_2 & \leq & 5 \\ & x_1, & & x_2 & \in & \mathbb{Z}_{\geq 0} \end{array}$$

- The objective minimises the spending
- The first constraint ($\geq$) requires the demand of 12 units to be covered
- The second ($\leq$) caps the total lots at five
- Last row: the domain. All variables integer: it is an **ILP**

Example: a model of BIP type

Which projects to fund: $y_k = 1$ if project k is funded, 0 otherwise.

$$
\begin{array}{rrcrcrcl}
\max & 5y_1 & + & 4y_2 & + & 6y_3 & & \\
\text{subject to} & y_1 & + & y_2 & + & y_3 & \leq & 2 \\
 & y_1 & & & + & y_3 & \geq & 1 \\
 & y_1, & & y_2, & & y_3 & \in & \{0,1\}
\end{array}
$$

- The objective maximises the value of the projects chosen
- First constraint ($\leq$): at most two of them can be funded
- Second ($\geq$): at least one of the first and the third
- Last row: all variables binary, it is a **BIP**

Example: a model of MILP type (1/2)

Two products, with x_1 and x_2 the quantities produced, and a plant to switch on or not: $y = 1$ if it is switched on, $y = 0$ otherwise.

$$
\begin{array}{rrcrcrcl}
\max & 3x_1 & + & 8x_2 & - & 10y & & \\
\text{subject to} & x_1 & + & x_2 & & & = & 6 \\
 & & & x_2 & - & 4y & \leq & 0 \\
 & x_1 & & & & & \geq & 1 \\
 & x_1, & & x_2 & & & \geq & 0 \\
 & & & & & y & \in & \{0,1\}
\end{array}
$$

Example: a model of MILP type (2/2)

- The objective maximises the revenue of the two products minus the cost of the plant
- First constraint ($=$): produce exactly the six units demanded
- Second ($\leq$): at most four units of the second product, and only with the plant on, because with $y = 0$ it leaves $x_2 \leq 0$
- Third ($\geq$): at least one unit of the first product
- Last two rows: two continuous variables and one binary, it is a **MILP**

How every model of the course is written

Every datum, every set and every symbol is *defined before it is used*. In every model:

1. the variables are introduced explicitly before the formulation
2. the objective and each family of constraints carry an equation number
3. the domain constraints, which declare the variables, close the model
4. after the model a list explains objective and constraints family by family

Notation for optimal values (1/2)

- **Optimal values** name the model they optimise: $z(MILP)$ is the optimum of a model of MILP type; the relaxations and their duals follow the same rule
- **Feasible** solutions carry a bar, however they were built, **optimal** ones a tilde:

$$(\bar{x}, \bar{\pi}) \qquad \text{and} \qquad (\tilde{x}, \tilde{\pi})$$

- The **bounds** are called LB and UB, whatever the direction of the objective:

$$LB \leq z(MILP) \leq UB$$

Notation for optimal values (2/2)

- When the solution a bound comes from matters, it is written in parentheses: a feasible solution $(\bar{x}_1, \ldots, \bar{x}_n)$ of a minimisation model gives

$$UB(\bar{x}) = \sum_{j=1}^{n} c_j \, \bar{x}_j$$

- In a maximisation model the same sum is $LB(\bar{x})$
- In a minimisation: feasible solution $\Rightarrow$ UB, relaxation $\Rightarrow$ LB
- In a maximisation it is the other way round

What this course uses

This course works almost exclusively with **MILPs**:

binary or integer variables for "yes/no" or counting decisions,

continuous variables for quantities that are measured (time, money, flow).

Summary

Ingredients of a model:

- Data, variables, objective, constraints; the domain of the variables closes the model
- Linear constraints of type $\geq$, $\leq$, $=$

Classes:

- LP (continuous), ILP (integer), BIP (binary), MILP (mixed)
- LP polynomial, ILP and MILP in general not

Notation:

- $\bar{x}$ feasible, $\tilde{x}$ optimal; $LB \leq z(MILP) \leq UB$

Why integrality matters

Introduction: Why integrality matters

- An ILP or a MILP is **not** solved by "rounding" the solution of its continuous relaxation
- We show it with a minimal counterexample, with only two binary variables
- We will draw two distinct lessons on the relaxation and on rounding

Example: rounding fails (1/5)

$$\begin{array}{lrcrcl} \max & x_1 & + & x_2 & & \\ \text{subject to} & 2x_1 & + & 2x_2 & \leq & 3 \\ & x_1, & & x_2 & \in & \{0,1\} \end{array}$$

- The LP relaxation replaces $x_1, x_2 \in \{0,1\}$ by

$$0 \leq x_1 \leq 1, \qquad 0 \leq x_2 \leq 1$$

- and has optimal value

$$z(LP^+) = \frac{3}{2}$$

Example: rounding fails (2/5)

- Careful: 3/2 is attained by *infinitely many* optimal solutions, every point of the segment

$$x_1 + x_2 = \frac{3}{2}, \qquad 0 \le x_1 \le 1, \qquad 0 \le x_2 \le 1$$

  among them (3/4, 3/4), (1, 1/2) and (1/2, 1)

- Which one the solver returns depends on the algorithm, not on the model: none of them is *the* solution of the relaxation

- On our installation Gurobi returns the vertex (1/2, 1)

Example: rounding fails (3/5)

Rounding gives different outcomes depending on the starting point *and* on the direction:

- from (3/4, 3/4):
  - upwards (1, 1): violates the constraint, $2 + 2 = 4 > 3$
  - downwards (0, 0): feasible but of value 0
  - *neither* direction finds the optimum
- from (1, 1/2):
  - upwards (1, 1) again, infeasible
  - downwards (1, 0), feasible with value 1: exactly the integer optimum $z(MILP) = 1$, attained at $(1, 0)$ and at $(0, 1)$

Example: rounding fails (4/5)

- So there is no "right" direction: here downwards works, and only from one of the two points
- Nothing, before solving, told us which point the solver would return

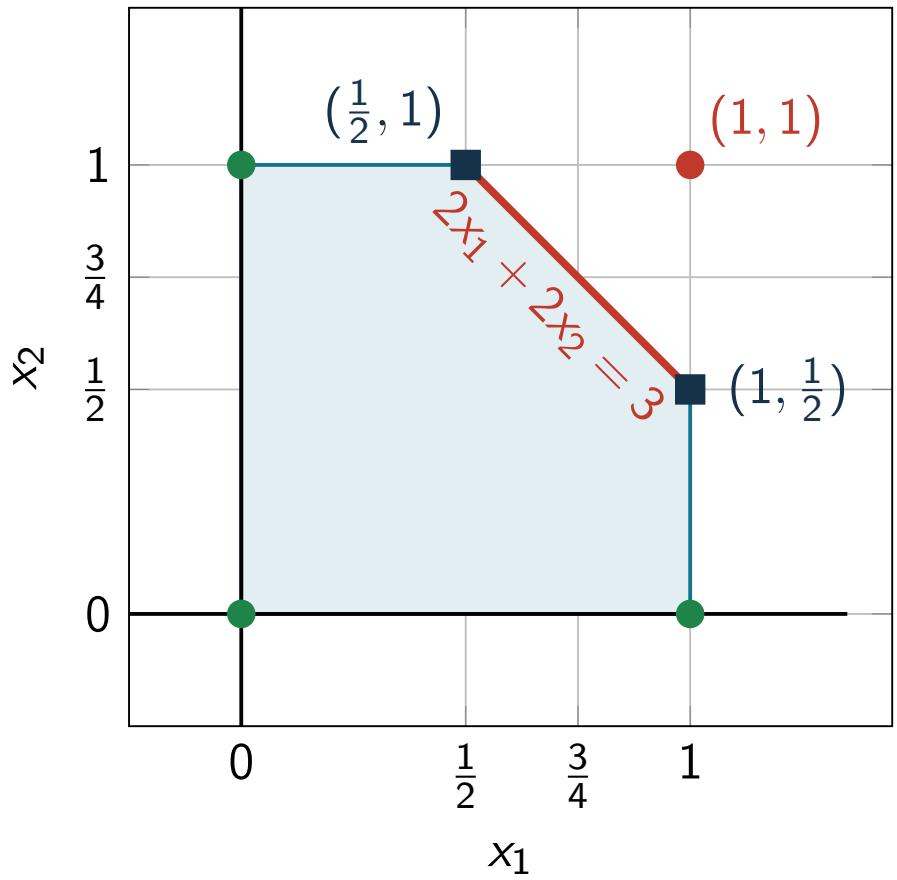
Two distinct lessons:

- Rounding can produce *infeasible* points, and when it produces feasible ones there is no guarantee whatsoever on their value
- The difference

$$\frac{3}{2} - 1 = \frac{1}{2}$$

  is not the fault of rounding: *no* feasible integer point is worth more than 1

Example: rounding fails (5/5)



Blue: feasible region of the LP^+ relaxation. Green: feasible integer points; red: $(1, 1)$, infeasible. Blue squares: two of the infinitely many optimal solutions of the relaxation, of value $3/2$.

Summary

The lesson:

- The continuous relaxation gives an indication (a **bound**), not a solution
- Rounding can give infeasible points, or feasible ones with no guarantee on their value
- The difference between relaxation and integer optimum may exist without being the fault of rounding

Hence:

- MILPs are solved with dedicated algorithms, not with naive heuristics on the continuous problem

The LP relaxation and the sandwich of bounds

Introduction: The LP relaxation and the sandwich of bounds

- We define the two versions of the **LP relaxation** used in the course
- We see on which side of the integer optimum each **bound**, primal and dual, lies
- We tell apart the three meanings of the word **gap**
- We understand what the bounds are actually for when the solver does not terminate

LP relaxation: two versions not to be confused (1/2)

The **LP relaxation** of a MILP is obtained by dropping the integrality and binary requirements and keeping everything else. Two ways:

- **relaxation with the bounds**, $z(LP^+)$:
  - $x \in \{0,1\}$ becomes $0 \leq x \leq 1$, $x \in \mathbb{Z}_{\geq 0}$ becomes $x \geq 0$
  - what Gurobi solves with `relax()`, the first relaxation the solver solves
- **relaxation without the bounds**, $z(LP)$:
  - $x \in \{0,1\}$ becomes $x \geq 0$ alone
  - weaker (or equal), but with a more compact dual: the one whose dual is written by hand

LP relaxation: two versions not to be confused (2/2)

- The relaxation without the bounds is the weaker of the two, that is the bound *furthest* from the integer optimum
- The two coincide when the other constraints already imply $x \leq 1$
- This happens, for instance, with an assignment constraint

$$\sum_m x_{jm} = 1$$

The relaxation is a bound: on which side, and why

- The relaxation *removes* constraints: every feasible solution of the MILP stays feasible for the relaxation, not conversely
- Optimising over a larger set cannot give a worse value:

$$\text{maximisation:} \quad z(LP) \geq z(LP^+) \geq z(MILP)$$

$$\text{minimisation:} \quad z(LP) \leq z(LP^+) \leq z(MILP)$$

- In a maximisation it is an *upper* bound, in a minimisation a *lower* bound
- Always **optimistic**: never a solution, and the value it promises may be unreachable

Which side each bound comes from (1/3)

Observation 1: Which side each bound comes from

- ***Primal bounds*** *come from feasible integer solutions, built by hand or found by the solver: the* ***pessimistic*** *side*
- ***Dual bounds*** *come from the optimal solution of the LP relaxation or from any feasible solution of its dual: the* ***optimistic*** *side*

Which side each bound comes from (2/3)

- The dual of the relaxation does *not* give a bound "from the other side"
- By **weak duality**, in a minimisation every feasible dual solution is worth at most $z(LP)$, hence at most $z(MILP)$
- It sits on the *same* side as the relaxation, in general weaker
- Let $\bar{x}$ be feasible for the MILP, of value $\sum_{j=1}^{n} c_j \bar{x}_j$, and $\bar{\pi}$ feasible for the dual of the relaxation, of value $\sum_{i=1}^{m} b_i \bar{\pi}_i$

Which side each bound comes from (3/3)

Minimisation: $\bar{x}$ gives $UB(\bar{x})$, $\bar{\pi}$ gives $LB(\bar{\pi})$

$$LB(\bar{\pi}) \leq z(D(LP)) = z(LP) \leq z(LP^+) \leq z(MILP) \leq UB(\bar{x})$$

Maximisation: roles swapped, $\bar{x}$ gives $LB(\bar{x})$, $\bar{\pi}$ gives $UB(\bar{\pi})$

$$LB(\bar{x}) \leq z(MILP) \leq z(LP^+) \leq z(LP) = z(D(LP)) \leq UB(\bar{\pi})$$

- The equality in the middle is **strong duality**: the dual of the relaxation is worth exactly the relaxation
- The real inequality is between the *value of a feasible dual solution* and that common value

The sandwich of bounds

The two bounds meet in the *sandwich*

$$LB \leq z(MILP) \leq UB$$

that runs through the whole course.

A feasible solution gives the pessimistic side,
a relaxation or a solution of its dual
gives the optimistic one.

Three "gaps" not to be confused (1/3)

The word *gap* denotes three different quantities. Conventions:

- denominator always $|z(MILP)|$ or $|UB|$, as indicated
- numerator in absolute value
- if the denominator is zero, the relative gap is not written (the absolute difference is reported)

1. Heuristic gap. If $\bar{x}$ is the heuristic solution, its value is $UB(\bar{x})$ in a minimisation ($LB(\bar{x})$ in a maximisation), and when the optimum is known

$$\text{gap} = \frac{\left|UB(\bar{x}) - z(MILP)\right|}{|z(MILP)|}$$

Three "gaps" not to be confused (2/3)

- The heuristic gap is the one in the exercise tables: the course instances are solved to optimality
- When the optimum is *not* known it is computed against a dual bound, which takes the place of $z(MILP)$: the number is larger, and it is the only one that can be claimed

2. Certified gap between two known bounds, without knowing $z(MILP)$:

$$\frac{UB - LB}{|UB|} \qquad \text{(minimisation, } UB > 0\text{)}$$

- It guarantees that $z(MILP) \in [LB, UB]$, not that it is close to either end

Three "gaps" not to be confused (3/3)

3. **The solver's** `MIPGap`:
   - the same idea, computed by Gurobi from $UB =$ `ObjVal` of the incumbent and $LB =$ `ObjBound`
   - convention $|UB - LB|/|UB|$, with its tolerances
   - it is 0 when optimality has been proved

The formulas in items 2 and 3 require the incumbent to exist and the denominator to be non-zero: they are **conventions**, not mathematical identities, and must be stated every time a number is reported.

What the two bounds are actually for (1/2)

- In this course's exercises the instances are small on purpose: Gurobi finds $z(MILP)$ in a fraction of a second
- LB and UB serve to *understand* where a bound comes from and to check the reasoning behind the heuristic and the dual
- In practice a real MILP can have thousands of variables and the solver may not terminate in useful time
- It is stopped: what is left is the best solution found and the best bound proved at the time of the stop

What the two bounds are actually for (2/2)

- The certified gap between the two is all one can say with certainty about the distance from the optimum, *without* knowing $z(MILP)$
- Building a feasible solution quickly, and an optimistic bound by hand, is what remains when the exact optimum is out of reach
- It is also what makes it possible to judge the numbers the solver reports

Summary (1/2)

LP relaxation:

- With the bounds ($z(LP^+)$) and without the bounds ($z(LP)$, weaker, more compact dual)
- Upper bound in a maximisation, lower bound in a minimisation: always optimistic

Bounds:

- Primal from feasible solutions (pessimistic)
- Dual from the relaxation or from feasible dual solutions (optimistic)
- Sandwich: $LB \leq z(MILP) \leq UB$

Summary (2/2)

Three gaps:

- Heuristic gap against $z(MILP)$ (or against a dual bound)
- Certified gap $(UB - LB)/|UB|$
- Gurobi's `MIPGap`

In practice:

- When the solver does not terminate, the certified gap is all that can be claimed

Three patterns that keep coming back

Introduction: Three patterns that keep coming back

- Many MILP constraints with binary selection variables, indexed by a set I, belong to one of three families
- The family depends only on the direction of the constraint
- Recognising the pattern immediately tells how the constraint behaves

Partitioning, packing, covering

Name	Constraint
set partitioning	$\sum_{i \in I} x_i = 1$
set packing	$\sum_{i \in I} x_i \leq 1$
set covering	$\sum_{i \in I} x_i \geq 1$

- **partitioning**: exactly one element of I is chosen
- **packing**: at most one element of I is chosen
- **covering**: at least one element of I is chosen

Recognising the pattern

- "Every job to exactly one machine" is a **set partitioning**
- "At most one job per time slot" is a **set packing**
- "Every zone covered by at least one station" is a **set covering**, the direct translation of an OR clause
- A *covering* allows several elements to be chosen together, a *packing* does not, a *partitioning* imposes both at once

Summary

Three patterns:

- $\sum_{i \in I} x_i = 1$: set partitioning (exactly one)
- $\sum_{i \in I} x_i \leq 1$: set packing (at most one)
- $\sum_{i \in I} x_i \geq 1$: set covering (at least one, OR clause)

Why they matter:

- The direction of the constraint immediately tells how many elements can be chosen together

What the solver does with the two bounds

Introduction: What the solver does with the two bounds

- A MILP with a bounded feasible set is solved with dedicated algorithms
- How they work (exploration, cuts, branching) is the subject of a separate course
- Only one thing matters here: the two bounds we build by hand are the two sides of the same sandwich the solver tries to close
- They are not handed to it: the solver builds its own, with its own heuristics and relaxations

The role of the two bounds in the solver

- The **primal bound** is a feasible solution, and gives a value to beat: any part of the space that cannot do better is discarded without being explored
- The **dual bound** comes from the relaxation, and says how much one may hope for at most: the closer it is to the integer optimum, that is the **tighter** the formulation, the less work is left
- The gap between the two is the `MIPGap`, the only thing one can claim with certainty when the optimum is not reached

Summary

The solver and the bounds:

- The primal bound allows parts of the space to be discarded
- The dual bound measures how much one may still hope for
- The `MIPGap` is the gap between the two

What is of use from this course:

- Writing the model with a tight formulation
- Being able to produce the two bounds by hand

What this chapter has put in your hands

Introduction: What this chapter has put in your hands

- We line up again the objects every MIP model is written with
- and the notions of solution, relaxation, bound and gap used in the rest of the course

Final summary (1/2)

A MIP model is always written with the same five objects:

- the **data**, known and not decided
- the **variables**, one per decision, with their domain
- the **objective function**, which says what is maximised or minimised
- the **constraints**, one for every sentence of the statement
- the **domain** of the variables, a constraint too, and the easiest to forget

Final summary (2/2)

- **Feasible solution**: satisfies every constraint; **optimal solution**: the best among the feasible ones
- Relaxing integrality gives the **linear relaxation**, an easier problem whose optimum is a **bound**
- Pessimistic bound from a feasible solution, optimistic from a relaxation
- The distance between the two is the **gap**: the rest of the course is about building the two sides and tightening them