

# Chapter 1

## Exponential Size Integer Programming Formulations

Although typically models need to be a reasonable size in order for us to code them and send them to a solver, there are some ways that we can allow having models of exponential size. The first example here is the cutting stock problem, where we will model with exponentially many variables. The second example is the traveling salesman problem, where we will model with exponentially many constraints. We will also look at some other models for the traveling salesman problem.

### 1.1 Cutting Stock

This is a classic problem that works excellent for a technique called *column generation*. We will discuss two versions of the model and then show how we can use column generation to solve the second version more efficiently. First, let's describe the problem.

**Cutting Stock:**

You run a company that sells of pipes of different lengths. These lengths are  $L_1, \dots, L_k$ . To produce these pipes, you have one machine that produces pipes of length  $L$ , and then cuts them into a collection of shorter pipes as needed.

You have an order come in for  $d_i$  pipes of length  $i$  for  $i = 1, \dots, k$ . How can you fill the order while cutting up the fewest number of pipes?

**Example 1:** A plumber stocks standard lengths of pipe, all of length 19 m. An order arrives for:

- 12 lengths of 4m
- 15 lengths of 5m
- 22 lengths of 6m

How should these lengths be cut from standard stock pipes so as to minimize the number of standard pipes used?



An initial model for this problem could be constructed as follows:

- Let  $N$  be an upper bound on the number of pipes that we may need.
- Let  $z_j = 1$  if we use pipe  $j$  and  $z_j = 0$  if we do not use pipe  $j$ , for  $j = 1, \dots, N$ .
- Let  $x_{ij}$  be the number of cuts of length  $L_i$  in pipe  $j$  that we use.

Then we have the following model

$$\begin{aligned}
 \min \quad & \sum_{j=1}^N z_j \\
 \text{s.t.} \quad & \sum_{i=1}^k L_i x_{ij} \leq L z_j \quad \text{for } j = 1, \dots, N \\
 & \sum_{j=1}^N x_{ij} \geq d_i \quad \text{for } i = 1, \dots, k \\
 & z_j \in \{0, 1\} \quad \text{for } j = 1, \dots, N \\
 & x_{ij} \in \mathbb{Z}_+ \quad \text{for } i = 1, \dots, k, j = 1, \dots, N
 \end{aligned} \tag{1.1.1}$$

**Exercise 1.** In the example above, show that we can choose  $N = 16$ .

For our example above, using  $N = 16$ , we have

$$\begin{aligned}
 \min \quad & \sum_{j=1}^{16} z_j \\
 \text{s.t.} \quad & 4x_{1j} + 5x_{2j} + 6x_{3j} \leq 19z_j \\
 & \sum_{j=1}^{16} x_{1j} \geq 12 \\
 & \sum_{j=1}^{16} x_{2j} \geq 15 \\
 & \sum_{j=1}^{16} x_{3j} \geq 22 \\
 & z_j \in \{0,1\} \text{ for } j = 1, \dots, 16 \\
 & x_{ij} \in \mathbb{Z}_+ \text{ for } i = 1, \dots, 3, j = 1, \dots, 16
 \end{aligned} \tag{1.1.2}$$

Additionally, we could break the symmetry in the problem. That is, suppose the solution uses 10 of the 16 pipes. The current formulation does not restrict which 10 pipes are used. Thus, there are many possible solutions. To reduce this complexity, we can state that we only use the first 10 pipes. We can write a constraint that says *if we don't use pipe  $j$ , then we also will not use any subsequent pipes*. Hence, by not using pipe 11, we enforce that pipes 11, 12, 13, 14, 15, 16 are not used. This can be done by adding the constraints

$$z_1 \geq z_2 \geq z_3 \geq \dots \geq z_N. \tag{1.1.3}$$

See ?? for code for this formulation.

Unfortunately, this formulation is slow and does not scale well with demand. In particular, the number of variables is  $N + kN$  and the number of constraints is  $N$  (plus integrality and non-negativity constraints on the variables). The solution times for this model are summarized in the following table:

#### Solver comparison:

	Cbc solver	Gurobi solver
standard formulation	0.05 sec	0.02 sec
with symmetry broken	0.02 sec	0.0055 sec

I give up...

#### Using 10 times more demand:

	Cbc solver	Gurobi solver
standard formulation	0.6 sec	3.3 sec
with symmetry broken	6.5 sec	19 sec

<sup>1</sup>This table was taken from notes by Laurent Lessard.

### 1.1.1 Pattern formulation

We could instead list all patterns that are possible to cut each pipe. A pattern is an vector  $a \in \mathbb{Z}_+^k$  such that for each  $i$ ,  $a_i$  lengths of  $L_i$  can be cut from a pipe of length  $L$ . That is

$$\sum_{i=1}^k L_i a_i \leq L \quad (1.1.4)$$

$$a_i \in \mathbb{Z}_+ \text{ for all } i = 1, \dots, k$$

In our running example, we have

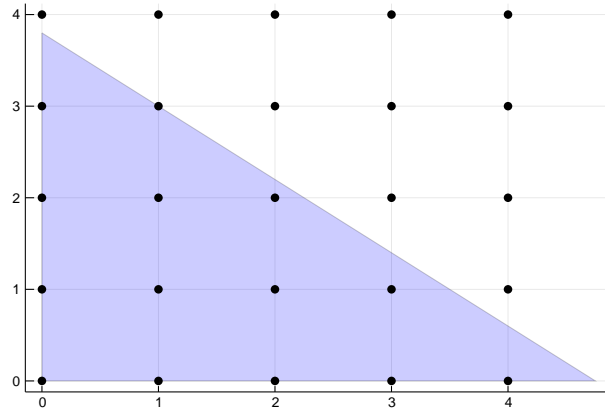
$$4a_1 + 5a_2 + 6a_3 \leq 19 \quad (1.1.5)$$

$$a_i \in \mathbb{Z}_+ \text{ for all } i = 1, \dots, 3$$

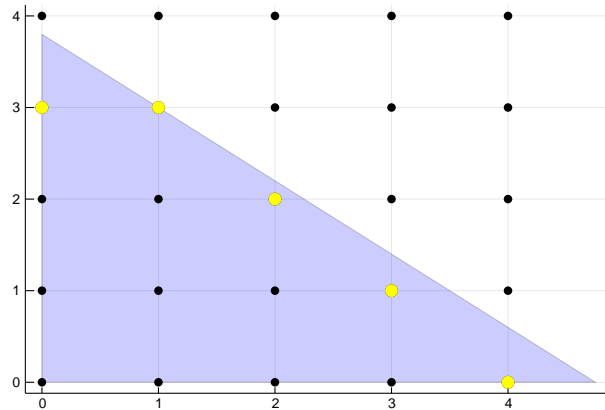
For visualization purposes, consider the patterns where  $a_3 = 0$ . That is, only patterns with cuts of length 4m or 5m. All patterns of this type are represented by an integer point in the polytope

$$P = \{(a_1, a_2) : 4a_1 + 5a_2 \leq 19, a_1 \geq 0, a_2 \geq 0\} \quad (1.1.6)$$

which we can see here:



where  $P$  is the blue triangle and each integer point represents a pattern. Feasible patterns lie inside the polytope  $P$ . Note that we only need patterns that are maximal with respect to number of each type we cut. Pictorially, we only need the patterns that are integer points represented as yellow dots in the picture below.



For example, the pattern  $[3, 0, 0]$  is not needed (only cut 3 of length 4m) since we could also use the pattern  $[4, 0, 0]$  (cut 4 of length 4m) or we could even use the pattern  $[3, 1, 0]$  (cut 3 of length 4m and 1 of length 5m).

### 1.1.2 Column Generation

Consider the (??), but in this case we are instead minimizing. Thus we can write it as

$$\begin{aligned} \min \quad & (c_N - c_B B^{-1} N) x_N + c_B B^{-1} b \\ \text{s.t.} \quad & x_B + B^{-1} N x_N = B^{-1} b \\ & x \geq 0 \end{aligned} \tag{1.1.7}$$

In our LP we have  $c = \mathbb{1}$ , that is,  $c_i = 1$  for all  $i = 1, \dots, k$ . Hence, we can write it as

$$\begin{aligned} \min \quad & (\mathbb{1}_N - \mathbb{1}_B B^{-1} N) x_N + \mathbb{1}_B B^{-1} b \\ \text{s.t.} \quad & x_B + B^{-1} N x_N = B^{-1} b \\ & x \geq 0 \end{aligned} \tag{1.1.8}$$

Now, if there exists a non-basic variable that could enter the basis and improve the objective, then there is one with a reduced cost that is negative. For a particular non-basic variable, the coefficient on it is

$$\left(1 - \mathbb{1}_B B^{-1} N^i\right) x_i \tag{1.1.9}$$

where  $N^i$  is the  $i$ -th column of the matrix  $N$ . Thus, we want to look for a column  $a$  of  $N$  such that

$$1 - \mathbb{1}_B B^{-1} a < 0 \Rightarrow 1 < \mathbb{1}_B B^{-1} a \tag{1.1.10}$$

#### Pricing Problem:

(knapsack problem!)

Given a current basis  $B$  of the *master* linear program, there exists a new column to add to the basis that improves the LP objective if and only if the following problem has an objective value strictly larger than 1

$$\begin{aligned} \max \quad & \mathbb{1}_B B^{-1} a \\ \text{s.t.} \quad & \sum_{i=1}^k L_i a_i \leq L \\ & a_i \in \mathbb{Z}_+ \text{ for } i = 1, \dots, k \end{aligned} \tag{1.1.11}$$

**Example 2:** Pricing Problem(knapsack problem!) After choosing the initial columns ....

we can find the objective function....

$$\begin{array}{ll}
 \max & \dots\dots a \\
 \text{s.t.} & 4a_1 + 5a_2 + 6a_3 \leq 19 \\
 & a_i \in \mathbb{Z}_+ \text{ for } i = 1, \dots, k
 \end{array} \tag{1.1.12}$$

### 1.1.3 Cutting Stock - Multiple widths

Gurobi has an excellent demonstration application to look at: <https://www.gurobi.com/cutting-stock-problem-with-multiple-master-rolls/>

<https://demos.gurobi.com/cutstock/>.

Here are some solutions:

- <https://github.com/fzsun/cutstock-gurobi>.
- <http://www.dcc.fc.up.pt/~jpp/mpa/cutstock.py>

Here is an AIMMS description of the problem:

[https://download.aimms.com/aimms/download/manuals/AIMMS30M\\_CuttingStock.pdf](https://download.aimms.com/aimms/download/manuals/AIMMS30M_CuttingStock.pdf)

## 1.2 Spanning Trees

Volume 10, Number 3

OPERATIONS RESEARCH LETTERS

April 1991

### 3.1. Minimum spanning tree

An LP formulation for the minimum spanning tree (MST) problem on  $G$  is

$$\min \left\{ \sum_{e \in E} c_e x_e \mid x \in P_{\text{MST}}(G) \right\},$$

where

$$P_{\text{MST}}(G) = \{ x \mid x(E) = n - 1, x(\gamma(S)) \leq |S| - 1 \text{ for all } S \subseteq V, x_e \geq 0 \text{ for all } e \in E \}.$$

Cunningham [9], Padberg and Wolsey [15] and Picard and Queyranne [16] show that it is possible to test if  $0 \leq \bar{x} \leq 1$  violates a constraint  $x(\gamma(S)) \leq |S| - 1$ , with  $k \in S$  by solving a max flow problem. A linear program corresponding to the max flow problem (see [19]) is  $(\text{SP}_k(x))$  defined by

$$\max \left\{ \sum_{e \in E} x_e \alpha_e^k - \sum_{i=1, i \neq k}^n \theta_i^k \mid \alpha_e^k \leq \theta_i^k \text{ for all } e \in \delta(i), i \in V, \theta_i^k \geq 0 \right\}.$$

The constraint matrix of  $(\text{SP}_k(x))$  is totally unimodular and the extreme rays may be taken to have 0-1 components. Then, for  $0 \leq \bar{x}$ , there is always an optimal solution where  $\alpha_{(i,j)}^k = \min\{\theta_i^k, \theta_j^k\}$ . It follows that for  $0 \leq \bar{x} \leq 1$ ,  $v(\text{SP}_k(\bar{x})) = 0$  if and only if  $\bar{x}(\gamma(S)) \leq |S| - 1$ ,  $k \in S$ . Therefore  $(\text{SP}_k(x))$  is a valid separation LP for all cuts which contain vertex  $k$ .

Using Proposition 1, model (MST) is reformulated below as model (RMST) and is  $O(n^3)$  in both variables and constraints while the number of rows in (MST) is  $O(2^n)$ .

$$(\text{RMST}) \quad \text{minimize} \quad \sum_{e \in E} c_e x_e \tag{10}$$

$$\text{subject to} \quad \sum_{e \in E} x_e = n - 1, \tag{11}$$

$$z_{kij} + z_{kji} = x_e, \quad k = 1, \dots, n, \quad e \in \gamma(\{i, j\}), \tag{12}$$

$$\sum_{s>i} z_{kis} + \sum_{h<i} z_{kih} \leq 1, \quad k = 1, \dots, n, \quad i \neq k, \tag{13}$$

$$\sum_{s>k} z_{kks} + \sum_{h<k} z_{khh} \leq 0, \quad k = 1, \dots, n, \tag{14}$$

$$x_e \geq 0 \text{ for all } e \in E, \quad z_{kij} \geq 0 \text{ for all } k, i, j. \tag{15}$$

[?]

$$\text{Minimize } \sum_{d \in \mathcal{D}} l_d z_d \quad (34)$$

Subject to:

$$\sum_{(i,j) \in \mathcal{D}} z_{(i,j)} = 1 \quad \forall i \in \mathcal{N} - \{\omega\} \quad (35)$$

$$\sum_{(\omega,j) \in \mathcal{D}} z_{(\omega,j)} = 0 \quad (36)$$

$$V_i \geq V_j + z_{(i,j)} - n(1 - z_{(i,j)}) \quad \forall (i,j) \in \mathcal{D} \quad \forall i \in \mathcal{N} - \{\omega\} \quad (37)$$

$$V_\omega = 0 \quad (38)$$

$$z_d \in \{0,1\} \quad \forall d \in \mathcal{D} \quad (39)$$

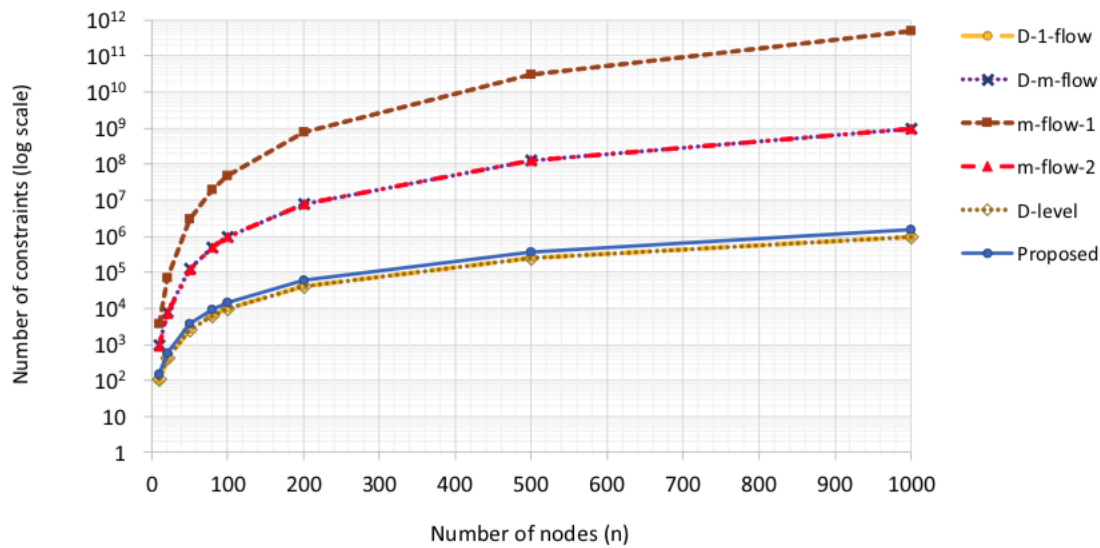


Figure 3. Comparison between selected models based on the number of constraints.

Table 1. Model characteristic and size comparison.

Model No.	Model Name Abbreviation	LP Relax. Is Integer	Number of Variables	Number of Integer Variables	Number of Constraints
1st	Cycle	yes	$A$	$A$	$2^n - n - 1$
2nd	1-cut	no	$A$	$A$	$2^{n-1}$
3rd	m-cut	yes	$A$	$A$	$\left\{ \frac{n}{K} \right\} + 1$
4th	D-cycle	yes	$3A$	$A$	$2^n + A - 2$
5th	D-1-cut	yes	$3A$	$A$	$2^{n-1} + A$
6th	D-1-flow	no	$3A$	$A$	$2A + n + 1$
7th	D-m-flow	yes	$A(2n + 1)$	$3A$	$2An + A + 1$
8th	m-flow-1	yes	$A(2n - 1)$	$A$	$(A + 1)(n - 1)^2 + n$
9th	m-flow-2	no	$A(2n - 1)$	$A$	$2An + 1$
10th	D-level	no	$2A + n$	$2A$	$2A + 2n$
11th	Proposed	no	$2A + n$	$2A$	$3A + n + 2$

See [?]

for a list of 11 models for the minimum spanning tree and a comparison using CPLEX.



## 1.3 Traveling Salesman Problem

See <http://www.math.uwaterloo.ca/tsp/index.html> for excellent material on the TSP.

See also this chapter [https://www.math.uwaterloo.ca/~bico/papers/comp\\_chapter1.pdf](https://www.math.uwaterloo.ca/~bico/papers/comp_chapter1.pdf).

Also, watch this excellent talk by Bill Cook "Postcards from the Edge of Possibility": <https://m.youtube.com/watch?v=5VjphFYQKj8>

We consider a directed graph, graph  $G = (N, A)$  of nodes  $N$  and arcs  $A$ . Arcs are directed edges. Hence the arc  $(i, j)$  is the directed path  $i \rightarrow j$ .

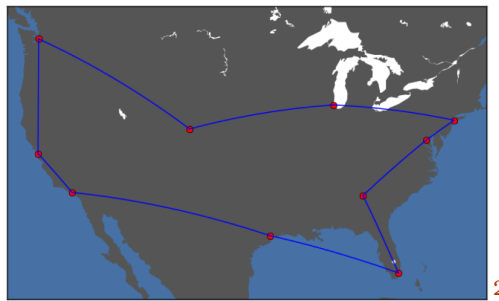
A *tour* is a cycle that visits all the nodes in  $N$  exactly once and returns back to the starting node.

Given costs  $c_{ij}$  for each arc  $(i, j) \in A$ , the goal is to find a minimum cost tour.

### Traveling Salesman Problem:

*NP-Hard*

Given a directed graph  $G = (N, A)$  and costs  $c_{ij}$  for all  $(i, j) \in A$ , find a tour of minimum cost.



In the figure, the nodes  $N$  are the cities and the arcs  $A$  are the directed paths  $\text{city}_i \rightarrow \text{city}_j$ .

**Models** When constructing an integer programming model for TSP, we define variables  $x_{ij}$  for all  $(i, j) \in A$  as

$x_{ij} = 1$  if the arc  $(i, j)$  is used and  $x_{ij} = 0$  otherwise.

We want the model to satisfy the fact that each node should have exactly one incoming arc and one leaving arc. Furthermore, we want to prevent self loops. Thus, we need the constraints:

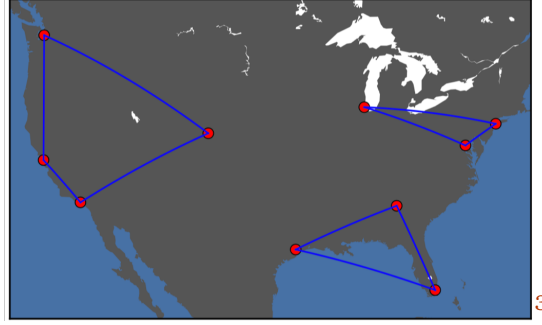
$$\sum_{j \in N} x_{ij} = 1 \quad \text{for all } i \in N \quad [\text{outgoing arc}] \quad (1.3.1)$$

$$\sum_{i \in N} x_{ij} = 1 \quad \text{for all } j \in N \quad [\text{incoming arc}] \quad (1.3.2)$$

$$x_{ii} = 0 \quad \text{for all } i \in N \quad [\text{no self loops}] \quad (1.3.3)$$

Unfortunately, these constraints are not enough to completely describe the problem. The issue is that *subtours* may arise. For instance

<sup>2</sup>This figure was taken from notes by Laurent Lessard.



### MTZ Model

- This model adds variables  $u_i \in \mathbb{Z}$  with  $1 \leq u_i \leq n$  that decide the order in which nodes are visited.
- We set  $u_1 = 1$  to set a starting place.
- Crucially, this model relies on the following fact

*Let  $x$  be a solution to (??)-(??) with  $x_{ij} \in \{0, 1\}$ . If there exists a subtour in this solution that contains the node 1, then there also exists a subtour that does not contain the node 1.*

The following model adds constraints

$$\text{If } x_{ij} = 1, \text{ then } u_i + 1 \leq u_j. \quad (1.3.4)$$

This if-then statement can be modeled with a big-M, choosing  $M = n$  is a sufficient upper bound. Thus, it can be written as

$$u_i + 1 \leq u_j + n(1 - x_{ij}) \quad (1.3.5)$$

Setting these constraints to be active enforces the order  $u_i < u_j$ .

Consider a subtour now  $2 \rightarrow 5 \rightarrow 3 \rightarrow 2$ . Thus,  $x_{25} = x_{53} = x_{32} = 1$ . Then using the constraints from (??), we have that

$$u_2 < u_5 < u_3 < u_2, \quad (1.3.6)$$

but this is infeasible since we cannot have  $u_2 < u_2$ .

As stated above, if there is a subtour containing the node 1, then there is also a subtour not containing the node 1. Thus, we can enforce these constraints to only prevent subtours that don't contain the node 1. Thus, the full tour that contains the node 1 will still be feasible.

This is summarized in the following model:

---

<sup>3</sup>This figure was taken from notes by Laurent Lessard.

**Traveling Salesman Problem - MTZ Model:**

$$\min \sum_{i,j \in N} c_{ij} x_{ij} \quad (1.3.7)$$

$$\sum_{j \in N} x_{ij} = 1 \quad \text{for all } i \in N \quad [\text{outgoing arc}] \quad (1.3.8)$$

$$\sum_{i \in N} x_{ij} = 1 \quad \text{for all } j \in N \quad [\text{incoming arc}] \quad (1.3.9)$$

$$x_{ii} = 0 \quad \text{for all } i \in N \quad [\text{no self loops}] \quad (1.3.10)$$

$$u_i + 1 \leq u_j + n(1 - x_{ij}) \quad \text{for all } i, j \in N, i, j \neq 1 \quad [\text{prevents subtours}] \quad (1.3.11)$$

$$u_1 = 1 \quad (1.3.12)$$

$$2 \leq u_i \leq n \quad \text{for all } i \in N, i \neq 1 \quad (1.3.13)$$

$$u_i \in \mathbb{Z} \quad \text{for all } i \in N \quad (1.3.14)$$

$$x_{ij} \in \{0, 1\} \quad \text{for all } i, j \in N \quad (1.3.15)$$

**Example 3:**

Distance Matrix:

A \ B	1	2	3	4
1	0	1	2	3
2	1	0	1	2
3	2	1	0	4
4	3	2	4	0

$$\begin{aligned}
& \min \quad x_{1,2} + 2x_{1,3} + 3x_{1,4} + x_{2,1} + x_{2,3} + 2x_{2,4} + \\
& 2x_{3,1} + x_{3,2} + 4x_{3,4} + 3x_{4,1} + 2x_{4,2} + 4x_{4,3} \\
& \text{Subject to} \\
& \quad x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4} = 1 \quad \text{outgoing from node 1} \\
& \quad x_{2,1} + x_{2,2} + x_{2,3} + x_{2,4} = 1 \quad \text{outgoing from node 2} \\
& \quad x_{3,1} + x_{3,2} + x_{3,3} + x_{3,4} = 1 \quad \text{outgoing from node 3} \\
& \quad x_{4,1} + x_{4,2} + x_{4,3} + x_{4,4} = 1 \quad \text{outgoing from node 4} \\
& \quad x_{1,1} + x_{2,1} + x_{3,1} + x_{4,1} = 1 \quad \text{incoming to node 1} \\
& \quad x_{1,2} + x_{2,2} + x_{3,2} + x_{4,2} = 1 \quad \text{incoming to node 2} \\
& \quad x_{1,3} + x_{2,3} + x_{3,3} + x_{4,3} = 1 \quad \text{incoming to node 3} \\
& \quad x_{1,4} + x_{2,4} + x_{3,4} + x_{4,4} = 1 \quad \text{incoming to node 4} \\
& \quad x_{1,1} = 0 \quad \text{No self loop with node 1} \\
& \quad x_{2,2} = 0 \quad \text{No self loop with node 2} \\
& \quad x_{3,3} = 0 \quad \text{No self loop with node 3} \\
& \quad x_{4,4} = 0 \quad \text{No self loop with node 4} \\
& \quad u_1 = 1 \quad \text{Start at node 1} \\
& \quad 2 \leq u_i \leq 4, \quad \forall i \in \{2, 3, 4\} \\
& \quad u_2 + 1 \leq u_3 + 4(1 - x_{2,3}) \\
& \quad u_2 + 1 \leq u_4 + 4(1 - x_{2,4}) \leq 3 \\
& \quad u_3 + 1 \leq u_2 + 4(1 - x_{3,2}) \leq 3 \\
& \quad u_3 + 1 \leq u_4 + 4(1 - x_{3,4}) \leq 3 \\
& \quad u_4 + 1 \leq u_2 + 4(1 - x_{4,2}) \leq 3 \\
& \quad u_4 + 1 \leq u_3 + 4(1 - x_{4,3}) \leq 3 \\
& \quad x_{i,j} \in \{0, 1\} \quad \forall i \in \{1, 2, 3, 4\}, j \in \{1, 2, 3, 4\} \\
& \quad u_i \in \mathbb{Z}, \quad \forall i \in \{1, 2, 3, 4\}
\end{aligned}$$

**Example 4:** 5 nodes

$$\begin{aligned}
\min \quad & x_{1,2} + 2x_{1,3} + 3x_{1,4} + 4x_{1,5} + x_{2,1} + x_{2,3} + 2x_{2,4} + 2x_{2,5} + 2x_{3,1} + \\
& x_{3,2} + 4x_{3,4} + x_{3,5} + 3x_{4,1} + 2x_{4,2} + 4x_{4,3} + 2x_{4,5} + \\
& 4x_{5,1} + 2x_{5,2} + x_{5,3} + 2x_{5,4} \\
\text{Subject to} \quad & x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4} + x_{1,5} = 1 \\
& x_{2,1} + x_{2,2} + x_{2,3} + x_{2,4} + x_{2,5} = 1 \\
& x_{3,1} + x_{3,2} + x_{3,3} + x_{3,4} + x_{3,5} = 1 \\
& x_{4,1} + x_{4,2} + x_{4,3} + x_{4,4} + x_{4,5} = 1 \\
& x_{5,1} + x_{5,2} + x_{5,3} + x_{5,4} + x_{5,5} = 1 \\
\\
& x_{1,1} + x_{2,1} + x_{3,1} + x_{4,1} + x_{5,1} = 1 \\
& x_{1,2} + x_{2,2} + x_{3,2} + x_{4,2} + x_{5,2} = 1 \\
& x_{1,3} + x_{2,3} + x_{3,3} + x_{4,3} + x_{5,3} = 1 \\
& x_{1,4} + x_{2,4} + x_{3,4} + x_{4,4} + x_{5,4} = 1 \\
& x_{1,5} + x_{2,5} + x_{3,5} + x_{4,5} + x_{5,5} = 1 \\
\\
& x_{1,1} = 0 \\
& x_{2,2} = 0 \\
& x_{3,3} = 0 \\
& x_{4,4} = 0 \\
& x_{5,5} = 0 \\
\\
& u_1 = 1 \\
& 2 \leq u_i \leq 5 \quad \forall i \in \{1, 2, 3, 4, 5\} \\
& u_2 + 1 \leq u_3 + 5(1 - x_{2,3}) \\
& u_2 + 1 \leq u_4 + 5(1 - x_{2,4}) \\
& u_2 + 1 \leq u_5 + 5(1 - x_{2,5}) \\
& u_3 + 1 \leq u_2 + 5(1 - x_{3,2}) \\
& u_3 + 1 \leq u_4 + 5(1 - x_{3,4}) \\
& u_4 + 1 \leq u_2 + 5(1 - x_{4,2}) \\
& u_4 + 1 \leq u_3 + 5(1 - x_{4,3}) \\
& u_3 + 1 \leq u_5 + 5(1 - x_{3,5}) \\
& u_4 + 1 \leq u_5 + 5(1 - x_{4,5}) \\
& u_5 + 1 \leq u_2 + 5(1 - x_{5,2}) \\
& u_5 + 1 \leq u_3 + 5(1 - x_{5,3}) \\
& u_5 + 1 \leq u_4 + 5(1 - x_{5,4}) \\
& x_{i,j} \in \{0, 1\} \quad \forall i \in \{1, 2, 3, 4, 5\}, j \in \{1, 2, 3, 4, 5\} \\
& u_i \in \mathbb{Z}, \quad \forall i \in \{1, 2, 3, 4, 5\}
\end{aligned}$$

**Pros of this model**

- Small description
- Easy to implement

**Cons of this model**

- Linear relaxation is not very tight. Thus, the solver may be slow when given this model.

**1.3.1 Dantzig-Fulkerson-Johnson Model**

This model does not add new variables. Instead, it adds constraints that conflict with the subtours. For instance, consider a subtour

$$2 \rightarrow 5 \rightarrow 3 \rightarrow 2. \quad (1.3.16)$$

We can prevent this subtour by adding the constraint

$$x_{25} + x_{53} + x_{32} \leq 2 \quad (1.3.17)$$

meaning that at most 2 of those arcs are allowed to happen at the same time. In general, for any subtour  $S$ , we can have the *subtour elimination constraint*

$$\sum_{(i,j) \in S} x_{ij} \leq |S| - 1 \quad \text{Subtour Elimination Constraint.} \quad (1.3.18)$$

In the previous example with  $S = \{(2,5), (5,3), (3,2)\}$  we have  $|S| = 3$ , where  $|S|$  denotes the size of the set  $S$ .

This model suggests that we just add all of these subtour elimination constraints.

**Traveling Salesman Problem - DFJ Model:**

$$\min \sum_{i,j \in N} c_{ij} x_{ij} \quad (1.3.19)$$

$$\sum_{j \in N} x_{ij} = 1 \quad \text{for all } i \in N \quad [\text{outgoing arc}] \quad (1.3.20)$$

$$\sum_{i \in N} x_{ij} = 1 \quad \text{for all } j \in N \quad [\text{incoming arc}] \quad (1.3.21)$$

$$x_{ii} = 0 \quad \text{for all } i \in N \quad [\text{no self loops}] \quad (1.3.22)$$

$$\sum_{(i,j) \in S} x_{ij} \leq |S| - 1 \quad \text{for all subtours } S \quad [\text{prevents subtours}] \quad (1.3.23)$$

$$x_{ij} \in \{0,1\} \quad \text{for all } i,j \in N \quad (1.3.24)$$

**Example 5:** DFJ Model for  $n = 4$  nodes

Distance Matrix:

A \ B	1	2	3	4
1	0	1	2	3
2	1	0	1	2
3	2	1	0	4
4	3	2	4	0

$$\begin{aligned} 1 \text{ min } & x_{1,2} + 2x_{1,3} + 3x_{1,4} + x_{2,1} + x_{2,3} + 2x_{2,4} \\ & + 2x_{3,1} + x_{3,2} + 4x_{3,4} + 3x_{4,1} + 2x_{4,2} + 4x_{4,3} \end{aligned}$$

Subject to

$$\begin{aligned} x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4} &= 1 && \text{outgoing from node 1} \\ x_{2,1} + x_{2,2} + x_{2,3} + x_{2,4} &= 1 && \text{outgoing from node 2} \\ x_{3,1} + x_{3,2} + x_{3,3} + x_{3,4} &= 1 && \text{outgoing from node 3} \\ x_{4,1} + x_{4,2} + x_{4,3} + x_{4,4} &= 1 && \text{outgoing from node 4} \end{aligned}$$

$$\begin{aligned} x_{1,1} + x_{2,1} + x_{3,1} + x_{4,1} &= 1 && \text{incoming to node 1} \\ x_{1,2} + x_{2,2} + x_{3,2} + x_{4,2} &= 1 && \text{incoming to node 2} \\ x_{1,3} + x_{2,3} + x_{3,3} + x_{4,3} &= 1 && \text{incoming to node 3} \\ x_{1,4} + x_{2,4} + x_{3,4} + x_{4,4} &= 1 && \text{incoming to node 4} \end{aligned}$$

$$\begin{aligned} x_{1,1} &= 0 && \text{No self loop with node 1} \\ x_{2,2} &= 0 && \text{No self loop with node 2} \\ x_{3,3} &= 0 && \text{No self loop with node 3} \\ x_{4,4} &= 0 && \text{No self loop with node 4} \end{aligned}$$

$$\begin{aligned} x_{1,2} + x_{2,1} &\leq 1 && S = [(1,2), (2,1)] \\ x_{1,3} + x_{3,1} &\leq 1 && S = [(1,3), (3,1)] \\ x_{1,4} + x_{4,1} &\leq 1 && S = [(1,4), (4,1)] \\ x_{2,3} + x_{3,2} &\leq 1 && S = [(2,3), (3,2)] \\ x_{2,4} + x_{4,2} &\leq 1 && S = [(2,4), (4,2)] \\ x_{3,4} + x_{4,3} &\leq 1 && S = [(3,4), (4,3)] \\ x_{2,1} + x_{1,3} + x_{3,2} &\leq 2 && S = [(2,1), (1,3), (3,2)] \\ x_{1,2} + x_{2,3} + x_{3,1} &\leq 2 && S = [(1,2), (2,3), (3,1)] \\ x_{3,1} + x_{1,4} + x_{4,3} &\leq 2 && S = [(3,1), (1,4), (4,3)] \\ x_{1,3} + x_{3,4} + x_{4,1} &\leq 2 && S = [(1,3), (3,4), (4,1)] \\ x_{2,1} + x_{1,4} + x_{4,2} &\leq 2 && S = [(2,1), (1,4), (4,2)] \\ x_{1,2} + x_{2,4} + x_{4,1} &\leq 2 && S = [(1,2), (2,4), (4,1)] \\ x_{3,2} + x_{2,4} + x_{4,3} &\leq 2 && S = [(3,2), (2,4), (4,3)] \\ x_{2,3} + x_{3,4} + x_{4,2} &\leq 2 && S = [(2,3), (3,4), (4,2)] \end{aligned}$$

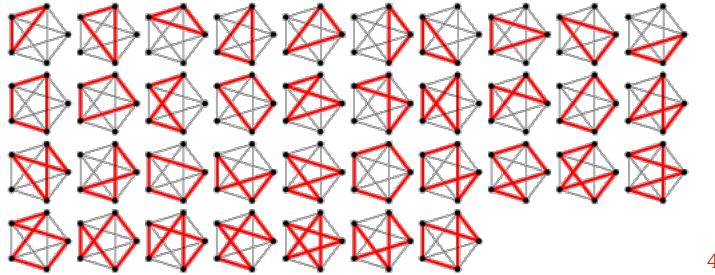
$$x_{i,j} \in \{0,1\} \quad \forall i \in \{1,2,3,4\}, j \in \{1,2,3,4\}$$



**Example 6:**

Consider a graph on 5 nodes.

Here are all the subtours of length at least 3 and also including the full length tours.



Hence, there are many subtours to consider.

**Pros of this model**

- Very tight linear relaxation

**Cons of this model**

- Exponentially many subtours  $S$  possible, hence this model is too large to write down.

**Solution: Add subtour elimination constraints as needed. We will discuss this in a future section on *cutting planes* .**

**1.3.2 Traveling Salesman Problem - Branching Solution**

We will see in the next section

1. That the constraint (??)-(??) always produce integer solutions as solutions to the linear relaxation.
2. A way to use branch and bound (the topic of the next section) in order to avoid subtours.

**1.4 Google maps data**

<https://www.geeksforgeeks.org/python-calculate-distance-duration-two-places-using-google-distan>

**1.5 Literature**

Gilmore-Gomory Cutting Stock [?]

[http://www.optimization-online.org/DB\\_HTML/2018/06/6648.html](http://www.optimization-online.org/DB_HTML/2018/06/6648.html)

[http://www.optimization-online.org/DB\\_HTML/2018/06/6670.html](http://www.optimization-online.org/DB_HTML/2018/06/6670.html)

[http://www.optimization-online.org/DB\\_FILE/2017/11/6331.pdf](http://www.optimization-online.org/DB_FILE/2017/11/6331.pdf)