GLPK Case Study 6 - 60016 Operations Research

Assume we operate company that imports coffee from Brazil to Poland. First we ship the coffee to the port in Marseilles and then use the rail network in Europe to deliver it to Krakow, Poland. The task is to find the best route considering limitations on the amount of goods that can be shipped through the railway network. After doing some analyses we are left with a few options that can be represented via the main 9 cities laying on each route under consideration. Those cities are the following: Marseilles, Lyon, Basel, Verona, Graz, Brno, Bratislava, Prague and Krakow. For example, one route can be Marseilles - Lyon - Verona - Graz - Brno - Bratislava - Krakow.

This problem can be modeled with a directed weighted graph. Let $N = (V, E)$ be a network (directed graph) with no parallel edges (meaning that any two nodes can be connected only by one edge), where nodes represent cities and edges represent routes between them. Then s and t are called the source (Marseilles in our case) and the sink (Krakow) of N respectively. Each edge from node u to node v has a capacity limit c_{uv}, defined as the mapping $c : E \mapsto \mathbb{R}^+$. It represents the maximum amount of flow that can pass through an edge. Flow on the edge from u to v is the mapping $f : E \mapsto \mathbb{R}^+$, denoted by f_{uv}, such that the flow of each edge is bounded by its capacity, also the total amounts of incoming and outgoing flows of each node are equal. In our case the flow of each edge is the amount of coffee shipped through that edge. Clearly each route can be used only up to its capacity limit, thus our goal is find out how to ship all the coffee from Marseilles to Krakow through via the rail network so that we use as much of the network capacity as possible, aka achieve maximum flow.

Figure 1 displays the network N on the map. For the sake of simpler modeling we enumerate the cities: Marseilles is 1, Lyon is 2, Basel is 3, Verona is 4, Graz is 5, Brno is 6, Bratislava is 7, Prague is 8 and Krakow is 9. We also show the corresponding capacity limits on each edge in blue, for instance from Munich to Prague we can send at most 18 tones of coffee.

Question 4.1. Write a LP that determines a route from Marseilles to Krakow with the maximum

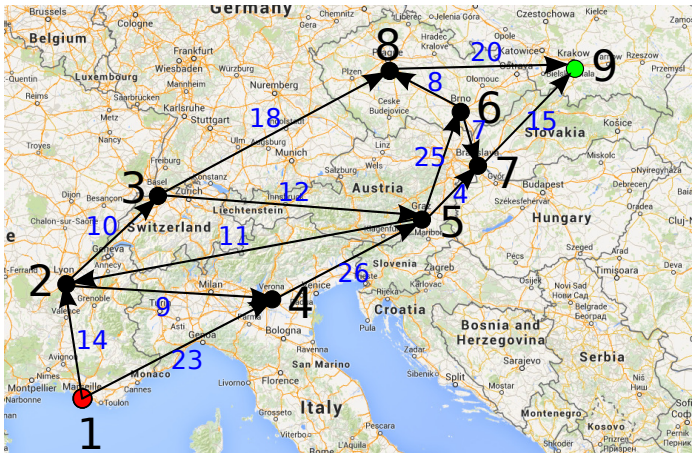

Figure 1: The network

possible flow. Write a GMPL program that solves this problem for *network.dat* file. *network.dat* contains three parameters: n for the number of cities in the network, E for the incidence matrix of the network and c containing the capacity limits for each edge in E.

Question 4.2. Now we consider a different criteria for the best route: the goal is to avoid *bottlenecks*. In order to incorporate the concept of a bottleneck into our model we define the $s-t$ cut as a partition $C = (S, T)$ of V such that $s \in S$ and $t \in T$. For example, with Marseilles, Lyon, Basel and Verona being in S, and Graz, Brno, Bratislava, Prague and Krakow in T we have an $s-t$ cut. Then the capacity of an $s-t$ cut is defined by $C(S,T) = \sum_{(u,v)\in S\times T} c_{uv}$, which reflects how much total capacity we have for routs exiting S and entering T. Thus we can use it as a measure of bottleneck.

- Write a LP that finds an $s-t$ cut with minimum capacity.
- Write a GMPL program that solves this problem for *network.dat*.

Question 4.3. Compare the solutions of Question 1 and Question 2.

Solution

Solution 4.1. The decision variables are the flows (amount of coffee sent) of each edge, i.e. f_{uv} shows how much should be sent from node u to v for all $(u, v) \in E$. Then the objective function is the value of flow:

$$F = \sum_{v:(s,v)\in E} f_{sv}, \tag{1}$$

which represents the amount of flow passing from the source to the sink.

Following the problem definitions we have the following constraints:

1. Capacity constraints, representing the fact that we can't ship via any path more than its capacity:
$$f_{uv} \le c_{uv}, \quad \forall (u,v) \in E \tag{2}$$
2. Conservation of flows, meaning that for all nodes, except the source and target, the total amount of flow entering it is equal to the total amount of flow leaving from it. For example, if we send 8 tones from Lyon to Basel, we must send a cumulative 8 tones from Basel to Prague and Graz.
$$\textstyle\sum_{\{u:(u,v)\in E\}} f_{uv} = \sum_{\{u:(v,u)\in E\}} f_{vu}, \quad \forall v \in V \setminus \{s,t\} \tag{3}$$
Thus the amount of flow leaving from the source node is the same as the amount of flow entering the target node, which is the value of flow defined above. Thus we add the following two constraints correspondingly for the flow leaving the source s and entering target t:
$$F + \sum_{\{u:(u,s)\in E\}} f_{us} - \sum_{\{u:(s,u)\in E\}} f_{su} = 0 \tag{4}$$
$$-F + \sum_{\{u:(u,t)\in E\}} f_{ut} - \sum_{\{u:(t,u)\in E\}} f_{tu} = 0 \tag{5}$$

Now we can formulate the maximum flow problem as a LP:

$$\begin{array}{lll} \max & F = \sum_{v:(s,v)\in E} f_{sv} & \\ \textit{subject to} & \sum_{\{u:(u,v)\in E\}} f_{uv} - \sum_{\{u:(v,u)\in E\}} f_{vu} \le 0 & \forall v \in V \setminus \{s,t\} \\ & F + \sum_{\{u:(u,s)\in E\}} f_{us} - \sum_{\{u:(s,u)\in E\}} f_{su} \le 0 & \\ & -F + \sum_{\{u:(u,t)\in E\}} f_{ut} - \sum_{\{u:(t,u)\in E\}} f_{tu} \le 0 & \\ & f_{uv} \le c_{uv}, & \forall (u,v) \in E \\ & f_{uv} \ge 0 & \forall (u,v) \in E \end{array} \tag{P}$$

Listing 1: maxflow.mod

```
/* MAXFLOW, Maximum Flow Problem */

param n, integer, >= 2;

set V, default {1..n};
set E, within V cross V;

param c{(i,j) in E}, > 0;
param s, in V, default 1;
param t, in V, != s, default n;

var f{(i,j) in E}, >= 0, <= c[i,j];
var flow, >= 0; /* total flow from s to t */


maximize maxflow: flow;
s.t.
node{i in V}: sum{(j,i) in E} f[j,i]+ (if i = s then flow) = sum{(i,j) in E} f[i,j]+ (if i = t then flow);

solve;

printf maxflow;
```

Solution 4.2. For a partition (S, T) we define the following decision variables:

$$d_{uv} = \begin{cases} 1 & (u,v) \text{ is such that } u \in S, \text{ and } v \in T \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

and

$$p_u = \begin{cases} 0 & u \in T \\ 1 & u \in S \end{cases} \tag{7}$$

We use $d(u, v)$ to indicate if nodes u and v are in different partitions of V, and $p(u)$ represents whether u is in the same partition with source s or target t.

Then we can define the objective function as $\sum_{(u,v)\in E} c_{uv} d_{uv}$, sum of all capacities from nodes in partition S to nodes in partition T, which is the capacity of the $s - t$ cut. Furthermore, we must make sure that our two groups of binary variables, namely d_{uv} and p_u, are consistent with each other. It is easy to check that for all cases of nodes u and v being in either S or T the constraint $d_{uv} - p_u + p_v \geq 0$ enforces that the definitions of d_{uv} and p_u remain valid. Last, we add a constraint $p_s - p_t \geq 1$ to ensure that source s is in partition S and target t is in partition T.

Therefore we can write the corresponding relaxed LP:

$$\begin{array}{lll} \textit{minimize} & \sum_{(u,v)\in E} c_{uv} d_{uv} & \\ \textit{subject to} & d_{uv} - p_u + p_v \geq 0 & \forall (u,v) \in E \\ & p_s - p_t \geq 1 & \\ & p_u \geq 0 & \forall u \in V \\ & d_{uv} \geq 0 & \forall (u,v) \in E \end{array} \tag{D}$$

This is called *minimum cut problem.*

Listing 2: mincut.mod

```
/* MINCUT, Minimum Cut Problem */

param n, integer, >= 2;

set V, default {1..n};
set E, within V cross V;

param c{(i,j) in E}, > 0;
param s, symbolic, in V, default 1;
param t, symbolic, in V, != s, default n;
```

```
var d{(i,j) in E}, >= 0;
var p{i in V}, >= 0;


minimize mincut: sum{(i,j) in E} c[i,j]*d[i,j];
s.t. edge{(i,j) in E}: d[i,j]-p[i]+p[j] >= 0;
     source_sink:      p[s]-p[t] >= 1;

solve;

printf mincut;
```

Solution 4.3. As we can see from the outputs of *maxflow.mod* and *mincut.mod* the optimal values of the two problems are equal. This is due the strong duality, as *minimum cut problem* is in fact the dual of *maximum flow problem*. In order to check it, note that

- the primal (P) is a max problem and the dual (D) is a min problem
- (D) has as many constraints (namely $|E| + 1$) as (P) has variables and (D) has as many variables (namely $vertE| + |V|$) as P has constraints.
- the dual coefficient matrix is the transpose of the primal coefficient matrix. The right hand side c_{uv} of (P) is the cost of (D) and the cost 1 of (P) is the right hand side of (D).
- positive dual variables p_u correspond to the first three primal $\geq$ constraint groups of (P) and d_{uv} correspond to the fourth primal $\geq$ constraint group of (P),
- primal positive variables f_{uv} correspond to the first dual $\geq$ constraint group of (D) and F corresponds to the second dual $\geq$ constraint of (D).