Solving debts problems with Linear Programming

MOPSI 2020-2021

Adrien Sadé

École nationale des ponts et chaussées - Département IMI
Supervised by Vincent Leclere, researcher at CERMICS

February 27, 2021

Overview

1. Mathematical modelization and introduction of an equivalent problem

2. An alternative and effective heuristic method

1. Mathematical modelization and introduction of an equivalent problem

2. An alternative and effective heuristic method

Context

We consider a group of n persons. For some reasons, they all owe money to the other members. Thus, one can wonder, *what is the best way to balance the finances of the group ?*

What does 'best way' mean ?

This problem is common and has been studied before. For example, **Tricount** is a pretty complete mobile app that came up as a solution to this kind of problems.

Mathematical model

We introduce the following variables:

- $D \in \mathbb{R}_+^{n\times n}$: the matrix of debts (the input), $D_{i,j}$ is the amount i owes to j.
- $T \in \mathbb{R}_+^{n\times n}$: the matrix of transactions (the output), $T_{i,j}$ is the amount i must reimburse to j.

We can then model our situation with the following problem:

$$
\begin{aligned}
\text{Min } & ||T||_0 \\
\text{s.t. } & \sum_j T_{i,j} - T_{j,i} = \sum_j D_{i,j} - D_{j,i} \qquad \forall i \\
& T_{i,j} \geq 0 \qquad \forall i,j
\end{aligned}
$$

The objective function counts the number of exchanges (nonzeros in the transactions matrix). The main constraint ensure that the transactions allow everybody to reimburse what they owed or receive what they lent: the sums are on the columns.

MILP-problem

In order to compute the exact solution of the problem using linear solvers, we introduce a new formulation of the problem that is linear. To this aim, we introduce a new binary matrix $X \in \{0,1\}^{n \times n}$ which coordinates are 1 if i reimburse a positive amount to j, 0 otherwise. We can then model the problem as the following:

NP-problem (exchanges)

$$\begin{aligned}
\text{Min} \quad & \sum_{i,j} X_{i,j} \\
\text{s.t.} \quad & \sum_{j} T_{i,j} - T_{j,i} = \sum_{j} D_{i,j} - D_{j,i} \qquad \forall i \\
& T_{i,j} \geq 0 \qquad \forall i,j \\
& X_{i,j} \geq T_{i,j}/||D||_1 \qquad \forall i,j \\
& X_{i,j} \in \{0,1\} \qquad \forall i,j
\end{aligned}$$

Continuous **LP**-problem

It's interesting to compare the optimal solution(s) in terms of exchanges computed thanks to the previous **NP** formulation with the optimal solution(s) in terms of flows. In order to do that we model the new problem with the following continuous linear formulation:

LP-problem (flow)

$$\begin{aligned} \text{Min} \quad & \sum_{i,j} T_{i,j} \\ \text{s.t.} \quad & \sum_{j} T_{i,j} - T_{j,i} = \sum_{j} D_{i,j} - D_{j,i} \qquad \forall i \\ & T_{i,j} \geq 0 \qquad \forall i, j \end{aligned}$$

The **LP**-solutions are optimal in terms of flows (total amount exchanged) whereas the **NP**-solutions are optimal in terms of exchanges.

NP/LP-equivalence ?

Definition (NP/LP-equivalence)

In the sequel, a debts problem is said to be '**NP**/**LP**-equivalent' if the optimal solutions of both **NP** and **LP**-problems imply the same number of exchanges.

Theorem (Donoho - Tanner (2005))

Let $y \in \mathbb{R}^i$, $A \in \mathbb{R}^{i \times j}$ and $x \in \mathbb{R}^j$ a variable. Let

$$\begin{array}{lll} (\text{NP*}) & Min\ ||x||_0 & s.t.\ \ y = Ax,\ \ x \geq 0 \\ (\text{LP*}) & Min\ ||x||_1 & s.t.\ \ y = Ax,\ \ x \geq 0 \end{array}$$

If $j > i > 2$, then NP/LP*-equivalence holds as long as the number of nonzeros of an optimal solution of* NP* *is*

$$||x||_0 \leq \left\lfloor \frac{i}{2} \right\rfloor$$

NP/**LP**-equivalence ?

Implementation and resolution of both problems using *Julia* and *JuMP*, and the solver *Gurobi*. The tests were done using random debts matrices. The results are summed up in the table below.

NP/**LP**-equivalence								
n	3	4	5	6	7	8	9	10
true (%)	1.00	1.00	1.00	1.00	0.97	0.95	0.93	0.9

The tests where run 100 times, except for the case $n = 10$ where it was only run 10 times.

What makes the equivalence break ?

The heuristic

Why should we seek a heuristic ?

Algorithm 1 Heuristic to solve debts problems

$D \equiv$ list of indebted people by increasing order
$L \equiv$ list of lender people by increasing order
$V \equiv$ vector of algebraic debts
while $len(D) > 0$ **do**
    $indebted \leftarrow pop(D)$
    $lender \leftarrow pop(L)$
    $V[indebted] \leftarrow \max(0, V[indebted] + V[lender])$
    $V[lender] \leftarrow \min(0, V[lender] + V[indebted])$
    if $V[indebted] > 0$ **then**
        $D \leftarrow D + [indebted]$
    else if $V[lender] < 0$ **then**
        $L \leftarrow L + [lender]$
    end if
end while

Theoretical effectiveness

Theorem

The heuristic converges in at most $n-1$ steps, yielding a solution to the debts problem of at most $n-1$ exchanges.

Proof. Let's write $L = [L_1, \dots L_k]$ and $D = [D_1, \dots D_p]$ with

$$V[L_k] < \dots < V[L_1] < 0 < V[D_1] < \dots < V[D_p]$$

and $k + p = n$.

At each step, we eliminate at least one person from one of the lists.

The worst case scenario is, after $n-2$ steps we have

$$L = [L_1]$$
$$D = [D_1]$$

As we are in a debts problem (as define before), we have $V[L_1] = -V[D_1]$. Therefore, the very last step consists in eliminating both L_1 and D_1 simultaneously.

The finances of the n-group are balanced in at most $n-1$ steps.

Theoretical effectiveness

We can go further.

- On a continuous space the probability to have an exact matching between the amounts of a lender and an indebted is null a.s.. In this case, the heuristic converges in exactly $n-1$ steps a.s..
- Even if this equality arise, we are not sure that these particular lender and indebted will ever be confronted. However, we can improve the heuristic to run an equality search at every step.
- The solution given by the heuristic is actually a solution to the **LP**-problem. Indeed, the flow is bounded by the sum of the actual debts (sum of the positive debts of the algebraic vector of debts). And this heuristic gives a solution with this exact flow.

Practical effectiveness

We can compare the results given by the heuristic and the optimal flow solution.

Heuristic/**LP**-equivalence									
n	3	4	5	6	7	8	9	10	50
true (%)	1.00	1.00	1.00	0.998	0.998	0.997	0.996	0.995	0.955

The tests were run 10^4 times.

This slight difference is explained by a sparse gain of an exchange by the heuristic (we consider the equivalence about the number of exchanges). As explained before, this difference tends to 0 when the maximum amount possible of the random debts (uniformly distributed) tends toward $+\infty$.

Effectiveness on an incomplete graph

We consider a particular case, were in a n-group of indebted people, there are two smaller k-group with $n \equiv 0$ [2] (even) and $k = n/2$. We suppose the reimbursement graph is not complete anymore: in each smaller group it is complete (everyone can reimburse everyone), but between the two groups there is only **one** connection.

We adapt the **LP**-problem by adding the associated constraints, and the heuristic:

- On each smaller group, we compute the algebraic vector of debts we setting the debt of the connected person to $-\sum_{i \neq \text{connected}} V[i]$. We solve this problem with the former heuristic.
- To finish, we compute the new debts of both connected people (what they owed - what they paid during the last step). The last transaction is their exchange: the finances are balanced.

Notice that the number of exchanges is at most $(k-1) + (k-1) + 1 = n - 1$.

Effectiveness on an incomplete graph

The results are summed up below.

Heuristic/**LP**-equivalence (non-complete graph)									
n	4	6	8	10	12	14	16	18	20
true	1.00	0.999	1.00	0.999	0.997	0.996	0.993	0.991	0.990

The tests were run 10^4 times.

The End

References:

Miguel Biron Lattes (Feb. 2018)
Simplifying payments with linear programming
https://miguelbiron.github.io/post/2018-02-09-simplifying-pmts-with-lp/

David L. Donoho and Jared Tanner (Mar. 2005)
Sparse nonnegative solution of undetermined linear equations by linear programming
PNAS, National Academy of Sciences of the United States of America