

Dirk Schumacher

Mixed integer linear programming in R with *ompr*

R [r] is a popular language and computational environment among statisticians. In recent years R has gained a lot of popularity in other disciplines as well. With over 11,000 packages on CRAN [cran], the R package repository, there is a package for almost every problem. And of course, packages for solving mixed integer linear programs are also available. R as a language makes it easy to create domain specific languages through metaprogramming using so-called non-standard evaluation [advancedr].

ompr (the optimization modelling package) [ompr] is an attempt to develop an algebraic modelling language within R. The goal is to create a modelling API that uses idiomatic R and integrates well with other popular packages, such as *dplyr* for data processing. *ompr* models are solver-independent and currently support mixed integer linear programming problems. It works on all platforms supported by R (Linux, MacOS and Windows). Compared to other modelling languages it comes closest to *JuMP* [jump] for *julia* but is currently more limited in terms of features and scope. In fact *JuMP* inspired me to build *ompr*.

The features of *ompr* are best described by modelling a mixed integer linear program in R. As an example for this article we model a *warehouse location problem* where we would like to find the cost optimal location and number of warehouses and the assignment of customers to those hubs.

$$\begin{aligned}
 \min \quad & \sum_{i=1}^n \sum_{j=1}^m tc_{i,j} \cdot x_{i,j} + \sum_{j=1}^m fc_j \cdot y_j \\
 \text{subject to} \quad & \sum_{j=1}^m x_{i,j} = 1 \quad \forall i = 1, \dots, n \\
 & x_{i,j} \leq y_j, \quad \forall i = 1, \dots, n, j = 1, \dots, m \\
 & x_{i,j} \in \{0, 1\} \quad \forall i = 1, \dots, n, j = 1, \dots, m \\
 & y_j \in \{0, 1\} \quad \forall j = 1, \dots, m
 \end{aligned}$$

Figure 1. A warehouse location problem as an illustrative example. n customers need to be assigned to exactly one of at most m warehouses. Setting up a warehouse $y_j = 1$ has a fixed cost of $fc_j > 0$ and assigning a customer to a warehouse $x_{i,j} = 1$ leads to travel cost of $tc_{i,j} > 0$.

Every *ompr* model starts with an empty model to which necessary elements (like variables, constraints and an objective function) are added. Expressions can be mixed with variables defined in the general environment. For example in the model below the functions *tc* and *fc* are not part of the optimization problem, but are regular R functions defined in the environment. Like other algebraic domain specific languages, equations can be written down directly and the package takes care of transforming these to the required format for the API of the solver.

```
library(ompr)
model <- MIPModel() %>%
  add_variable(x[i, j], i=1:n, j=1:m, type="binary") %>%
  add_variable(y[j], j=1:m, type="binary") %>%
  set_objective(
    sum_expr(tc(i, j) * x[i, j], i=1:n, j=1:m) +
    sum_expr(fc(j) * y[j], j=1:m),
    sense = "min") %>%
  add_constraint(sum_expr(x[i, j], j=1:m) == 1, i=1:n) %>%
  add_constraint(x[i, j] <= y[j], i=1:n, j=1:m)
```

Figure 2. The warehouse location problem modeled with *ompr* using so called pipes to chain together successive functions.

After an *ompr* model is built, it can be passed to a solver. Within *ompr* a solver is simply a function mapping a model to a solution. In this case the GNU Linear Programming Kit [**glpk**] is used to find an optimal solution. To support most popular solvers out of the box, the package *ompr.roi* can be used. It makes it possible to use any solver supported by the *R optimization infrastructure (ROI)* packages [**roi**]. *ROI* gives access to many open source (*CLP*, *GLPK*, *lp-solve*, *Symphony*) and commercial solvers (*CPLEX*, *Gurobi*) through a standardized interface.

```
library(ompr.roi)
library(ROI.plugin.glpk)
result <- solve_model(model, with_ROI(solver = "glpk"))
assignment <- get_solution(result, x[i, j])
warehouses <- get_solution(result, y[i])
```

Figure 3. The model defined earlier is solved with the GNU Linear Programming Kit using the solver package *ompr.roi* which offers access to a variety of open source and commercial solvers.

The result can then be further processed within the R ecosystem. For example using R's plotting facilities, we can plot the assignment (Figure 4).

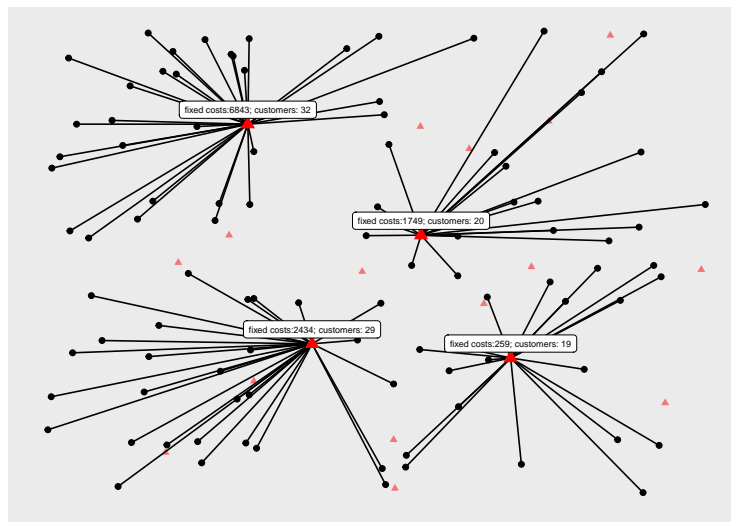


Figure 4. A solution to the warehouse location problem generated by *ggplot2* and *ompr*.

Currently *ompr*'s cost of abstraction can be rather high for larger models, but this is something I plan to work on in the future. Up to now the focus has been on creating an idiomatic API for modelling mixed integer linear programs directly in R. *ompr* as well as *ompr.roi* are published on CRAN and developed on GitHub. The packages are open source and I encourage anyone to send feedback, ideas or code contributions. On the project's website [**ompr**] I have compiled some articles showcasing the modelling features of the package and tutorials on how to model selected optimization problems in R with *ompr*.

In my opinion a big advantage of modelling optimization problems in R is the ability to easily utilize other packages within the R ecosystem for data wrangling, statistics, (interactive) visualizations and reproducible research. For example, this article was generated in a completely reproducible manner with *knitr* [**knitr**] and the code shown here is the actual code that led to the output in Figure 4; it is published on GitHub (<https://github.com/dirkschumacher/ompr-optima>) as well.

References

- [1] Hadley Wickham. Advanced R. CRC Press, 2014.
- [2] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <https://www.R-project.org>, 2008.
- [3] Dirk Schumacher. ompr: optimization modelling package. R package version 0.6, <https://cran.r-project.org/package=ompr>, 2017.
- [4] CRAN: The Comprehensive R Archive Network. <https://cran.r-project.org/>, 2017-07-16.
- [5] Kurt Hornik and David Meyer and Florian Schwendinger and Stefan Theussl. ROI: R Optimization Infrastructure. R package version 0.2-1, <https://cran.r-project.org/package=ROI>, 2016.
- [6] Dunning, Iain and Joey Huchette and Miles Lubin. JuMP: A modeling language for mathematical optimization. SIAM Review 59.2 (2017): 295-320.
- [7] GNU Linear Programming Kit. <http://www.gnu.org/software/glpk/glpk.html>, 2017.
- [8] Yihui Xie. knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.1, 2017.