Combinatorial Optimization @ Work 2015

Modelling

Christina Burt, Stephen J. Maher, Jakob Witzig

Zuse Institute Berlin
Berlin, Germany

29th September 2015

Research Center MATHEON
Mathematics for Key Technologies

Berlin Mathematical School

MODAL
Mathematical Optimization and Data Analysis Laboratories

Modelling Reformulations

Christina Burt

Comparing Formulations

Two comparisons are most interesting to us:

- showing that two formulations are equivalent; and
- showing that one formulation dominates the other with respect to linear relaxation.

The first of these tells us that if we solve the same instance with two equivalent models, we will get the same answer from both models.

The second tells us which model is likely to be faster.

Equivalent Formulations

To show that two formulations are equivalent, we must show:

1. that there exists a bijective mapping from one feasible region to the other; and
2. that the optimal solutions from both models will be the same point or region.

Dominant Formulations

To show that formulation A dominates formulation B, we must show for the linear relaxation:

- that no part of the feasible region (of the LP) for A steps outside the feasible region (of the LP) for B; and
- that some part of the feasible region (of the LP) for A is strictly inside the feasible region (of the LP) for B.

Good Formulations

Simple approaches that are likely to obtain a good formulation include:

- finding a compact formulation with respect to number of integer/binary variables, and constraints;
- tightening the linear relaxation.

However, this is not the whole picture, and these suggestions are not guaranteed to give fast models.

Study: Travelling Salesman Problem

Recall the classical combinatorial optimisation problem introduced in Martin Grötschel's lecture on linear and mixed integer linear programming. Here, we consider the *assymetric* version of that problem.

Definition (Travelling Salesman Problem)

Given a weighted di-graph, $D = (V, A)$, where each node $i \in V$ corresponds to a city and each arc (i, j) corresponds to paths between i and j. Let c_{ij} be the arc weights representing the distance between i and j, such that $c_{ij} = c_{ji} \; \forall \; (i, j) \in A$. Find a tour of minimal cost that visits each city exactly once.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} \\
\text{s.t. } & x(\delta^+(i)) = 1 && \forall\, i \in V, && \text{(out-degree)} \\
& x(\delta^-(i)) = 1 && \forall\, i \in V, && \text{(in-degree)} \\
& u_i - u_j + (|V|-1)x_{i,j} \leq |V| - 2 && \forall\, i \neq j, j > 1, && \text{(sub-tour elim)} \\
& u_{i-1} \leq u_i && \forall\, i \in V \setminus \{1\}, \\
& \boxed{x_{i,j} \in \{0,1\}} && \forall\, (i,j) \in A, \\
& u_i \in \mathbb{Z}^+ && \forall\, i \in V.
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} \\
\text{s.t. } & x(\delta^+(i)) = 1 && \forall\, i \in V, && \text{(out-degree)} \\
& x(\delta^-(i)) = 1 && \forall\, i \in V, && \text{(in-degree)} \\
& u_i - u_j + (|V|-1)x_{i,j} \leq |V| - 2 && \forall\, i \neq j, j > 1, && \text{(sub-tour elim)} \\
& u_{i-1} \leq u_i && \forall\, i \in V \backslash \{1\}, \\
& x_{i,j} \in \{0,1\} && \forall\, (i,j) \in A, \\
& \boxed{u_i \in \mathbb{Z}^+} && \forall\, i \in V.
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad \min & \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & \\
\text{s.t.} \; & x(\delta^+(i)) = 1 & \forall\, i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\, i \in V, & \quad \text{(in-degree)} \\
& u_i - u_j + (|V|-1)x_{i,j} \le |V| - 2 & \forall\, i \neq j, j > 1, & \quad \text{(sub-tour elim)} \\
& u_{i-1} \le u_i & \forall\, i \in V \backslash \{1\}, & \\
& x_{i,j} \in \{0,1\} & \forall\, (i,j) \in A, & \\
& u_i \in \mathbb{Z}^+ & \forall\, i \in V. &
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & \boxed{x(\delta^+(i)) = 1} & \forall\, i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\, i \in V, & \quad \text{(in-degree)} \\
& u_i - u_j + (|V|-1)x_{i,j} \leq |V| - 2 & \forall\, i \neq j, j > 1, & \quad \text{(sub-tour elim)} \\
& u_{i-1} \leq u_i & \forall\, i \in V \setminus \{1\}, & \\
& x_{i,j} \in \{0,1\} & \forall\, (i,j) \in A, & \\
& u_i \in \mathbb{Z}^+ & \forall\, i \in V. &
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & x(\delta^+(i)) = 1 & \forall\, i \in V, & \quad \text{(out-degree)} \\
& \boxed{x(\delta^-(i)) = 1} & \forall\, i \in V, & \quad \text{(in-degree)} \\
& u_i - u_j + (|V|-1)x_{i,j} \le |V| - 2 & \forall\, i \ne j, j > 1, & \quad \text{(sub-tour elim)} \\
& u_{i-1} \le u_i & \forall\, i \in V \setminus \{1\}, & \\
& x_{i,j} \in \{0,1\} & \forall\, (i,j) \in A, & \\
& u_i \in \mathbb{Z}^+ & \forall\, i \in V. &
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & x(\delta^+(i)) = 1 & \forall\ i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\ i \in V, & \quad \text{(in-degree)} \\
& \boxed{u_i - u_j + (|V|-1)x_{i,j} \leq |V| - 2} & \forall\ i \neq j, j > 1, & \quad \text{(sub-tour elim)} \\
& u_{i-1} \leq u_i & \forall\ i \in V \backslash \{1\}, & \\
& x_{i,j} \in \{0,1\} & \forall\ (i,j) \in A, & \\
& u_i \in \mathbb{Z}^+ & \forall\ i \in V. &
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A compact formulation

$$
\begin{aligned}
MTZ: \quad & \min \sum_{(i,j) \in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & x(\delta^+(i)) = 1 & \forall\, i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\, i \in V, & \quad \text{(in-degree)} \\
& u_i - u_j + (|V| - 1) x_{i,j} \leq |V| - 2 & \forall\, i \neq j, j > 1, & \quad \text{(sub-tour elim)} \\
& \boxed{u_{i-1} \leq u_i} & \forall\, i \in V \backslash \{1\}, & \\
& x_{i,j} \in \{0, 1\} & \forall\, (i,j) \in A, & \\
& u_i \in \mathbb{Z}^+ & \forall\, i \in V. &
\end{aligned}
$$

This is one of the most compact formulations of the TSP. However, it is generally not used because it has a loose linear relaxation.

TSP: A tight formulation

$$
\begin{aligned}
DFJ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & x(\delta^+(i)) = 1 & \forall\, i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\, i \in V, & \quad \text{(in-degree)} \\
& x(A(W)) \le |W| - 1 & \forall\, \emptyset \subsetneq W \subsetneq V, & \quad \text{(sub-tour elim)} \\
& \boxed{x_{i,j} \in \{0,1\}} & \forall\, (i,j) \in A. &
\end{aligned}
$$

This formulation has a tighter linear relaxation than MTZ.

However, this formulation has exponentially many subtour elimination constraints, since we must search the power set of V. The model is not compact, and suffers computationally from this flaw if the complete model is explicitly given to a solver.

TSP: A tight formulation

$$
\begin{aligned}
DFJ: \quad \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} & & & \\
\text{s.t. } x(\delta^+(i)) = 1 & \quad \forall\, i \in V, & & \text{(out-degree)} \\
x(\delta^-(i)) = 1 & \quad \forall\, i \in V, & & \text{(in-degree)} \\
x(A(W)) \leq |W| - 1 & \quad \forall\, \emptyset \subsetneq W \subsetneq V, & & \text{(sub-tour elim)} \\
x_{i,j} \in \{0,1\} & \quad \forall\, (i,j) \in A. & &
\end{aligned}
$$

This formulation has a tighter linear relaxation than MTZ.

However, this formulation has exponentially many subtour elimination constraints, since we must search the power set of V. The model is not compact, and suffers computationally from this flaw if the complete model is explicitly given to a solver.

TSP: A tight formulation

$$
\begin{aligned}
DFJ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} \\
\text{s.t. } & \boxed{x(\delta^+(i)) = 1} && \forall\, i \in V, && \text{(out-degree)} \\
& x(\delta^-(i)) = 1 && \forall\, i \in V, && \text{(in-degree)} \\
& x(A(W)) \le |W| - 1 && \forall\, \emptyset \subsetneq W \subsetneq V, && \text{(sub-tour elim)} \\
& x_{i,j} \in \{0,1\} && \forall\, (i,j) \in A.
\end{aligned}
$$

This formulation has a tighter linear relaxation than MTZ.

However, this formulation has exponentially many subtour elimination constraints, since we must search the power set of V. The model is not compact, and suffers computationally from this flaw if the complete model is explicitly given to a solver.

TSP: A tight formulation

$$
\begin{aligned}
DFJ: \quad & \min \sum_{(i,j)\in A} c_{i,j} x_{i,j} \\
\text{s.t. } & x(\delta^+(i)) = 1 && \forall\, i \in V, && \text{(out-degree)} \\
& \boxed{x(\delta^-(i)) = 1} && \forall\, i \in V, && \text{(in-degree)} \\
& x(A(W)) \le |W| - 1 && \forall\, \emptyset \subsetneq W \subsetneq V, && \text{(sub-tour elim)} \\
& x_{i,j} \in \{0,1\} && \forall\, (i,j) \in A.
\end{aligned}
$$

This formulation has a tighter linear relaxation than MTZ.

However, this formulation has exponentially many subtour elimination constraints, since we must search the power set of V. The model is not compact, and suffers computationally from this flaw if the complete model is explicitly given to a solver.

TSP: A tight formulation

$$
\begin{aligned}
DFJ: \quad & \min \sum_{(i,j) \in A} c_{i,j} x_{i,j} & & \\
\text{s.t. } & x(\delta^+(i)) = 1 & \forall\ i \in V, & \quad \text{(out-degree)} \\
& x(\delta^-(i)) = 1 & \forall\ i \in V, & \quad \text{(in-degree)} \\
& \boxed{x(A(W)) \leq |W| - 1} & \forall\ \emptyset \subsetneq W \subsetneq V, & \quad \text{(sub-tour elim)} \\
& x_{i,j} \in \{0, 1\} & \forall\ (i,j) \in A. &
\end{aligned}
$$

This formulation has a tighter linear relaxation than MTZ.

However, this formulation has exponentially many subtour elimination constraints, since we must search the power set of V. The model is not compact, and suffers computationally from this flaw if the complete model is explicitly given to a solver.

Study: Travelling Salesman Problem

The best way to use the DFJ formulation is to avoid writing the complete model explicitly.

We will look at how to do this, and implement it, in the lecture and exercise on constraint integer programming by Ambros Gleixner, Gregor Hendel and Felipe Serrano.

Good Formulations

The algorithms that solve mathematical programming problems are complex. Some heuristics within the algorithms may work better on less compact formulations.

For example:

- Primal Heuristics search for feasible solutions.
- Polishing Heuristics search for improved solutions from incumbent solutions (e.g., using local search)
- Cut Generator Heuristics will search for particular structure.

Martin Grötschel introduced us to the power of cuts in the lecture on Basics of Polyhedral Theory.

In the lecture on Integer Programming, Tobias Achterberg will discuss some of these heuristics.

However, as a general rule, using the guidelines on the previous slide is a good start.

The Equitable Coach Problem: base formulation

Consider the equitable coach time allocation problem where we wish to determine the **time** each player can spend on the field. Let us define the variables as follows:

- $x_{r,i}$ [integer] is the amount of time [minutes] r plays in position i;
- ϵ_r^+ [continuous] is the number of minutes player r plays above the average;
- ϵ_r^- [continuous] is the number of minutes player r plays below the average.
- R is the set of players;
- N is the set of positions;
- T is the total number of game minutes.

The Equitable Coach Problem: base formulation

$$
\begin{aligned}
ECPt: \quad & \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_r x_{r,i} = T & \forall\ i \in N, & \quad (1) \\
& \sum_i x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} = \epsilon_r^+ - \epsilon_r^- & \forall\ r \in R, & \quad (2) \\
& \boxed{x_{r,i} \in \mathbb{Z}_{\geq 0}} & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

This model has $|R||N| + 2|R|$ variables and $|N| + |R|$ constraints.

The Equitable Coach Problem: base formulation

$$ECPt: \quad \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_{r} x_{r,i} = T \qquad \forall\ i \in N, \qquad (1)$$

$$\sum_{i} x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \qquad (2)$$

$$x_{r,i} \in \mathbb{Z}_{\geq 0} \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\boxed{\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.}$$

This model has $|R||N| + 2|R|$ variables and $|N| + |R|$ constraints.

The Equitable Coach Problem: base formulation

$$
\begin{aligned}
ECPt: \quad & \min \sum_{r} \boxed{\left(\epsilon_r^+ + \epsilon_r^-\right)} \\
& \sum_{r} x_{r,i} = T && \forall\ i \in N, && (1) \\
& \sum_{i} x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} = \epsilon_r^+ - \epsilon_r^- && \forall\ r \in R, && (2) \\
& \quad x_{r,i} \in \mathbb{Z}_{\geq 0} && \forall r \in R,\ i \in N,\ t \in T, \\
& \quad \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.
\end{aligned}
$$

This model has $|R||N| + 2|R|$ variables and $|N| + |R|$ constraints.

The Equitable Coach Problem: base formulation

$$ECPt: \quad \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_{r} \boxed{x_{r,i} = T} \qquad \forall\ i \in N, \qquad (1)$$

$$\sum_{i} x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \qquad (2)$$

$$x_{r,i} \in \mathbb{Z}_{\geq 0} \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

This model has $|R||N| + 2|R|$ variables and $|N| + |R|$ constraints.

The Equitable Coach Problem: base formulation

$$ECPt: \quad \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_r x_{r,i} = T \qquad \forall\ i \in N, \tag{1}$$

$$\sum_i \boxed{x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} = \epsilon_r^+ - \epsilon_r^-} \qquad \forall\ r \in R, \tag{2}$$

$$x_{r,i} \in \mathbb{Z}_{\geq 0} \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

This model has $|R||N| + 2|R|$ variables and $|N| + |R|$ constraints.

The Equitable Coach Problem: noncompact

Consider the equitable coach time allocation problem where we wish to determine the time each player can spend on the field. Let us define the variables as follows:

$x_{r,i,t}$ [binary] is 1 if player r plays in position i in minute t;

ϵ_r^+ [continuous] is the number of minutes player r plays above the average;

ϵ_r^- [continuous] is the number of minutes player r plays below the average.

R is the set of players;

N is the set of positions;

T is the total number of game minutes.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_{t,r} x_{r,i,t} = M_i & \forall\, i \in N, & \quad (3)\\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \le \epsilon_r^+ & \forall\, r \in R, & \quad (4)\\
& \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \le \epsilon_r^- & \forall\, r \in R, & \quad (5)\\
& \boxed{x_{r,i,t} \in [0,1]} & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\ge 0}. & &
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) \\
& \sum_{t,r} x_{r,i,t} = M_i && \forall\ i \in N, && (3) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \leq \epsilon_r^+ && \forall\ r \in R, && (4) \\
& \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \leq \epsilon_r^- && \forall\ r \in R, && (5) \\
& x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, \\
& \boxed{\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.}
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_{r} \boxed{\left(\epsilon_r^+ + \epsilon_r^-\right)} \\
& \sum_{t,r} x_{r,i,t} = M_i && \forall\ i \in N, && (3) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \leq \epsilon_r^+ && \forall\ r \in R, && (4) \\
& \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \leq \epsilon_r^- && \forall\ r \in R, && (5) \\
& x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_{t,r} x_{r,i,t} = M_i & \forall\ i \in N, & \quad (3) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \leq \epsilon_r^+ & \forall\ r \in R, & \quad (4) \\
& \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \leq \epsilon_r^- & \forall\ r \in R, & \quad (5) \\
& x_{r,i,t} \in [0,1] & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_{t,r} x_{r,i,t} = M_i & \forall\, i \in N, & \quad (3) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \leq \epsilon_r^+ & \forall\, r \in R, & \quad (4) \\
& \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \leq \epsilon_r^- & \forall\, r \in R, & \quad (5) \\
& x_{r,i,t} \in [0,1] & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

The Equitable Coach Problem: noncompact

$$
\begin{aligned}
ncECPt: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) \\
& \sum_{t,r} x_{r,i,t} = M_i && \forall\ i \in N, && (3) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} \le \epsilon_r^+ && \forall\ r \in R, && (4) \\
& \sum_{t,i,r} \boxed{\frac{x_{r',i,t}}{|R|} - \sum_{t,i} x_{r,i,t} \le \epsilon_r^-} && \forall\ r \in R, && (5) \\
& x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\ge 0}.
\end{aligned}
$$

This model has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

Proposition

ECPt and ncECPt are equivalent models.

Proof.

Since $x_{r,i} \in \mathbb{Z}_{\geq 0}$, $\sum_t x_{r,i,t} = x_{r,i}$. With this substitution, it is clear that constraint (1) is equivalent to constraint (3). Using the substitution into constraints (4) and (5) gives:

$$\sum_i x_{r,i} - \sum_{i,r'} \frac{x_{r',i}}{|R|} \leq \epsilon_r^+ \qquad \forall\, r \in R, \tag{3}$$

$$\sum_{i,r'} \frac{x_{r',i}}{|R|} - \sum_i x_{r,i} \leq \epsilon_r^- \qquad \forall\, r \in R. \tag{4}$$

Since we minimise ϵ_r^+ and ϵ_r^-, ϵ_r^+ will take on the positive difference or 0, and ϵ_r^- will take on the negative difference or 0. Thus constraint (2) is equivalent to constraints (4) and (5). The feasible regions are therefore equivalent. The objective functions are the same, so the optimal solutions will be equivalent. □

ncECPt and ECPt Equivalence

So, we have shown that $ECPt$ and $ncECPt$ are equivalent.

However, the original model $ECPt$ has only $|N||R| + 2|R|$ variables and $|N| + |R|$ constraints.

This is significantly less variables and constraints than the $ncECPt$ model, which has $|R||N||T| + 2|R|$ variables and $|N| + 2|R|$ constraints.

Is it significant enough to make a difference to solve times?

The Equitable Coach Problem: aggregated model

We now consider the version of the ECP where we allocate players to positions at particular times.

$x_{r,i,t}$ [binary] is 1 if player r plays in position class i in play period t;

ϵ_r^+ [continuous] is the number of periods player r plays above the average;

ϵ_r^- [continuous] is the number of periods player r plays below the average.

R is the set of players;

N is the set of position classes;

T is the total number of game minutes.

The Equitable Coach Problem: aggregated model

$$
\begin{aligned}
aggECPd: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_{r} x_{r,i,t} = R_i & \forall\, i \in N,\ t \in T, & \quad (5)\\
& \sum_{i} x_{r,i,t} \leq 1 & \forall\, r \in R,\ t \in T, & \quad (6)\\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- & \forall\, r \in R, & \quad (7)\\
& \boxed{x_{r,i,t} \in [0,1]} & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

The Equitable Coach Problem: aggregated model

$$
\begin{aligned}
aggECPd: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) \\
& \sum_{r} x_{r,i,t} = R_i && \forall\, i \in N,\ t \in T, && (5) \\
& \sum_{i} x_{r,i,t} \leq 1 && \forall\, r \in R,\ t \in T, && (6) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- && \forall\, r \in R, && (7) \\
& x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, \\
& \boxed{\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.}
\end{aligned}
$$

The Equitable Coach Problem: aggregated model

$$
\begin{aligned}
aggECPd: \quad & \min \sum_r \boxed{(\epsilon_r^+ + \epsilon_r^-)} & & \\
& \sum_r x_{r,i,t} = R_i & \forall\, i \in N,\ t \in T, & \quad (5) \\
& \sum_i x_{r,i,t} \leq 1 & \forall\, r \in R,\ t \in T, & \quad (6) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- & \forall\, r \in R, & \quad (7) \\
& x_{r,i,t} \in [0,1] & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

The Equitable Coach Problem: aggregated model

$$
\begin{aligned}
aggECPd: \quad & \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right) & & \\
& \sum_{r} \boxed{x_{r,i,t} = R_i} & \forall\ i \in N,\ t \in T, & \quad (5) \\
& \sum_{i} x_{r,i,t} \leq 1 & \forall\ r \in R,\ t \in T, & \quad (6) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- & \forall\ r \in R, & \quad (7) \\
& x_{r,i,t} \in [0,1] & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

The Equitable Coach Problem: aggregated model

$$aggECPd: \quad \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_r x_{r,i,t} = R_i \qquad \forall\, i \in N,\; t \in T, \qquad (5)$$

$$\sum_i \boxed{x_{r,i,t} \le 1} \qquad \forall\, r \in R,\; t \in T, \qquad (6)$$

$$\sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\, r \in R, \qquad (7)$$

$$x_{r,i,t} \in [0,1] \qquad \forall r \in R,\; i \in N,\; t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\ge 0}.$$

The Equitable Coach Problem: aggregated model

$$
\begin{aligned}
aggECPd: \quad & \min \sum_{r} \left(\epsilon_r^{+} + \epsilon_r^{-}\right) & & \\
& \sum_{r} x_{r,i,t} = R_i & \forall\ i \in N,\ t \in T, & \quad (5)\\
& \sum_{i} x_{r,i,t} \leq 1 & \forall\ r \in R,\ t \in T, & \quad (6)\\
& \sum_{t,i} \boxed{x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^{+} - \epsilon_r^{-}} & \forall\ r \in R, & \quad (7)\\
& x_{r,i,t} \in [0,1] & \forall r \in R,\ i \in N,\ t \in T, & \\
& \epsilon_r^{+}, \epsilon_r^{-} \in \mathbb{R}_{\geq 0}. & &
\end{aligned}
$$

The Equitable Coach Problem: disaggregated model

We now consider the version of the ECP where we allocate players to positions at particular times.

- $x_{r,i,t}$ [binary] is 1 if player r plays in position i in play period t;
- ϵ_r^+ [continuous] is the number of periods player r plays above the average;
- ϵ_r^- [continuous] is the number of periods player r plays below the average.
- R is the set of players;
- N is the set of positions;
- T is the total number of game minutes.

The Equitable Coach Problem: disaggregated model

$$ECPd: \quad \min \sum_{r} \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_{r} x_{r,i,t} = 1 \qquad \forall\ t \in t, i \in N, \qquad (8)$$

$$\sum_{i} x_{r,i,t} \leq 1 \qquad \forall\ t \in T, r \in R, \qquad (9)$$

$$\sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \qquad (10)$$

$$\boxed{x_{r,i,t} \in [0,1]} \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

The Equitable Coach Problem: disaggregated model

$$
\begin{aligned}
ECPd: \quad & \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right) && \\
& \sum_r x_{r,i,t} = 1 && \forall\ t \in t, i \in N, && (8) \\
& \sum_i x_{r,i,t} \leq 1 && \forall\ t \in T, r \in R, && (9) \\
& \sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- && \forall\ r \in R, && (10) \\
& \quad x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, && \\
& \quad \boxed{\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.} && &&
\end{aligned}
$$

The Equitable Coach Problem: disaggregated model

$$ECPd: \quad \min \sum_r \boxed{\left(\epsilon_r^+ + \epsilon_r^-\right)}$$

$$\sum_r x_{r,i,t} = 1 \qquad \forall\ t \in t, i \in N, \qquad (8)$$

$$\sum_i x_{r,i,t} \leq 1 \qquad \forall\ t \in T, r \in R, \qquad (9)$$

$$\sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \qquad (10)$$

$$x_{r,i,t} \in [0,1] \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

The Equitable Coach Problem: disaggregated model

$$ECPd: \quad \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_r x_{r,i,t} = 1 \qquad \forall\ t \in t, i \in N, \qquad (8)$$

$$\sum_i x_{r,i,t} \leq 1 \qquad \forall\ t \in T, r \in R, \qquad (9)$$

$$\sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \qquad (10)$$

$$x_{r,i,t} \in [0,1] \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

The Equitable Coach Problem: disaggregated model

$$ECPd: \quad \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right)$$

$$\sum_r x_{r,i,t} = 1 \qquad \forall\ t \in t, i \in N, \tag{8}$$

$$\sum_i \boxed{x_{r,i,t} \leq 1} \qquad \forall\ t \in T, r \in R, \tag{9}$$

$$\sum_{t,i} x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^- \qquad \forall\ r \in R, \tag{10}$$

$$x_{r,i,t} \in [0,1] \qquad \forall r \in R,\ i \in N,\ t \in T,$$

$$\epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.$$

The Equitable Coach Problem: disaggregated model

$$
\begin{aligned}
ECPd: \quad & \min \sum_r \left(\epsilon_r^+ + \epsilon_r^-\right) \\
& \sum_r x_{r,i,t} = 1 && \forall\ t \in t, i \in N, && (8) \\
& \sum_i x_{r,i,t} \leq 1 && \forall\ t \in T, r \in R, && (9) \\
& \sum_{t,i} \boxed{x_{r,i,t} - \sum_{t,i,r'} \frac{x_{r',i,t}}{|R|} = \epsilon_r^+ - \epsilon_r^-} && \forall\ r \in R, && (10) \\
& x_{r,i,t} \in [0,1] && \forall r \in R,\ i \in N,\ t \in T, \\
& \epsilon_r^+, \epsilon_r^- \in \mathbb{R}_{\geq 0}.
\end{aligned}
$$

The Equitable Coach Problem: disaggregated model

This model performs better because the algorithms detected a Set-Packing constraint structure in the model. This can be seen by the number of setppc constraints in the SCIP output statistics.

This structure is well studied, and there are many known cuts for this type of structure. The topics of structure and cuts will be discussed in more detail in the coming days.