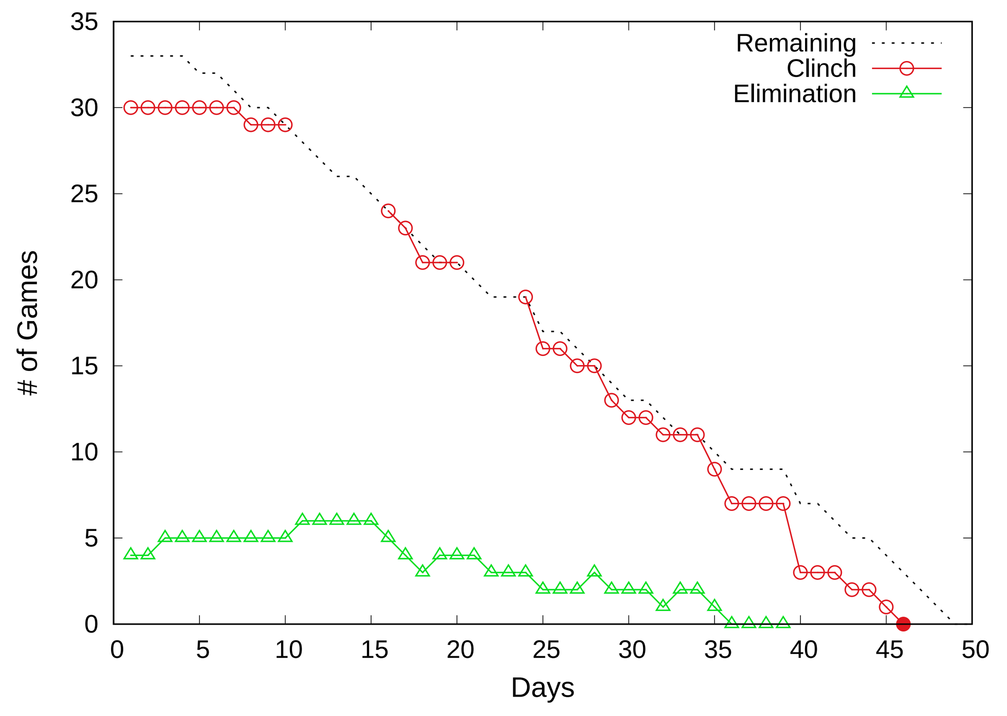
2020年10月27日　統計数理研究所　オープンハウス

整数計画によるクリンチ／エリミネーション数の計算

伊藤　聡　数理・推論研究系　教授

リーグスポーツのシーズン中のどの時点においても，最終的にリーグ優勝やプレーオフ出場権など特定の状況（指標）が達成されることが確定する最小の勝ち試合数（クリンチ数），もしくは逆にその状況（指標）に届かないことが確定する最小の負け試合数（エリミネーション数）が存在する（右の図に挙げた例では，エリミネーション数を，それぞれの時点で今後引分がないとしたときに，残りゲーム数から最小の負け試合数を引いた最大の勝ち試合数として表現している）．本研究は，順位決定に係る複数の判定基準が存在する場合のクリンチおよびエリミネーション数の計算を，多層の整数計画問題を解くことにより高速に行う汎用的な枠組みを開発することを目的としている．



キーワード：リーグスポーツ，総当たり戦，クリンチ数，エリミネーション数，順位判定基準，非線形整数計画，多層整数計画

シナリオ集合

リーグに属するチーム集合を L とし，そのチーム数を n とする．リーグ L の全チーム間のこれまでの勝敗記録と残り試合数が与えられているとし，w_{ij} をチーム $i \in L$ のチーム $j \in L$ に対する現時点での勝数，g_{ij} をチーム i, j 間の残り試合数とする（w および g は対角成分が0である n 次正方行列，特に後者は対称行列と考えてよい）．いま，チーム $i \in L$ のチーム $j \in L$ に対する今後の勝数を x_{ij} と表すことにすると，

$$x_{ij} + x_{ji} \le g_{ij}, \quad x_{ii} = 0, \quad x_{ij} \ge 0 \quad (\forall i, j \in L)$$

を満たす $x = (x_{ij}) \in \mathbb{Z}^{n \times n}$ は今後起こり得る勝敗に関するシナリオを過不足なく与える（$\mathbb{Z}$ は整数の集合であり，規則上引分がない場合は第1式の不等号を等号に変えるものとする）．残り試合数 g が与えられたとき，これらの条件を満たすシナリオ集合を X で表す．

クリンチ／エリミネーション問題の構造

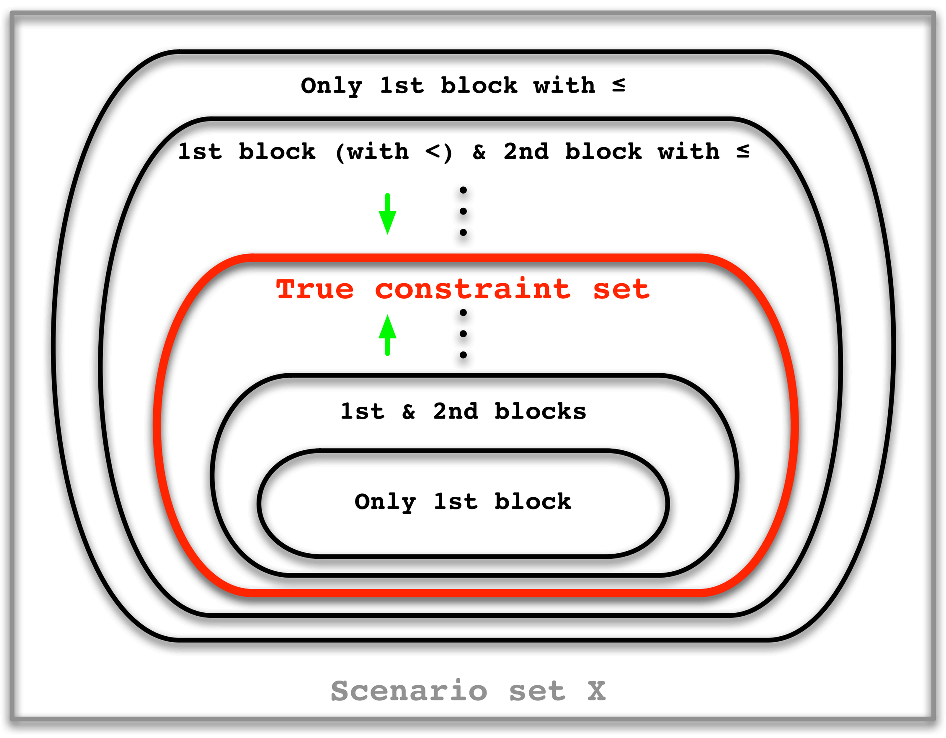
順位判定基準が複数あり，そのうち最初の m 個が勝敗数のみに基づく基準（すなわち $m+1$ 番目の基準は得失点差など勝敗数以外に基づくもの）であるとする．このとき簡単な例として，チーム $a \in L$ の第 k 位クリンチ数は，チーム a が今後引き分けることなく，かつ m 個のうちのいずれかの基準でチーム a より上位の成績を持つチームが k 個存在するという制約条件のもとで，チーム a の今後の勝数 $\sum_{j \in L} x_{aj}$ の最大値 $\bar{z}$ を求めることにより得られる（この最大化問題に許容解がなければ既に第 k 位以上が確定していることになるし，$\bar{z}$ が残り試合数と等しければ今後全勝しても第 $k+1$ 位以下の可能性があることになり，そうでなければ第 k 位クリンチ数は $\bar{z}+1$ となる）．最大値を与えるシナリオ $\bar{x} \in X$ は一般に唯一ではないが，最適解自体はこの枠組みでは不要である．ここで注意すべきことは，m 個の順位判定のうち最初の $m-1$ 個は等号なし（$<$）であり，最後の m 番目のみが等号つき（$\le$）の順位判定となることである．

例えば $m = 4$ の場合，チーム $i \in L$ に対する各順位判定基準の値を $(\mathrm{C1})_i \sim (\mathrm{C4})_i$ で表すことにすると，チーム $a \in L$ の第 k 位クリンチ数を求める問題（第 k 位クリンチ問題）は次のような構造を持つ。双対な概念である第 k 位エリミネーション数は，k 位以上になるという制約条件の下での今後の最大敗数をもとめる問題（第 k 位エリミネーション問題）を解くことにより，同様に求めることができる。

$$\left\{\begin{array}{l}
\displaystyle\max_{\substack{x \in X \\ \alpha,\beta,\gamma,\delta \in \{0,1\}^n}} \sum_{j \in L} x_{aj} \\
\text{subj. to } \; x_{aj} + x_{ja} = g_{aj} \;\; \forall j \in L \\
\qquad\quad \boxed{n - \displaystyle\sum_{i \in L} \alpha_i \text{ teams exist st } (\mathrm{C1})_a < (\mathrm{C1})_i} \\
\qquad\quad \boxed{n - \displaystyle\sum_{i \in L} \beta_i \text{ teams exist st } (\mathrm{C1})_a = (\mathrm{C1})_i,\, (\mathrm{C2})_a < (\mathrm{C2})_i} \\
\qquad\quad \boxed{n - \displaystyle\sum_{i \in L} \gamma_i \text{ teams exist st } (\mathrm{C1})_a = (\mathrm{C1})_i,\, (\mathrm{C2})_a = (\mathrm{C2})_i,\, (\mathrm{C3})_a < (\mathrm{C3})_i} \\
\qquad\quad \boxed{n - \displaystyle\sum_{i \in L} \delta_i \text{ teams exist st } (\mathrm{C1})_a = (\mathrm{C1})_i,\, (\mathrm{C2})_a = (\mathrm{C2})_i,\, (\mathrm{C3})_a = (\mathrm{C3})_i,\, (\mathrm{C4})_a \le (\mathrm{C4})_i} \\
\qquad\quad \alpha_a = \beta_a = \gamma_a = \delta_a = 1, \;\; \displaystyle\sum_{i \in L} (\alpha_i + \beta_i + \gamma_i + \delta_i) = 4n - k
\end{array}\right.$$

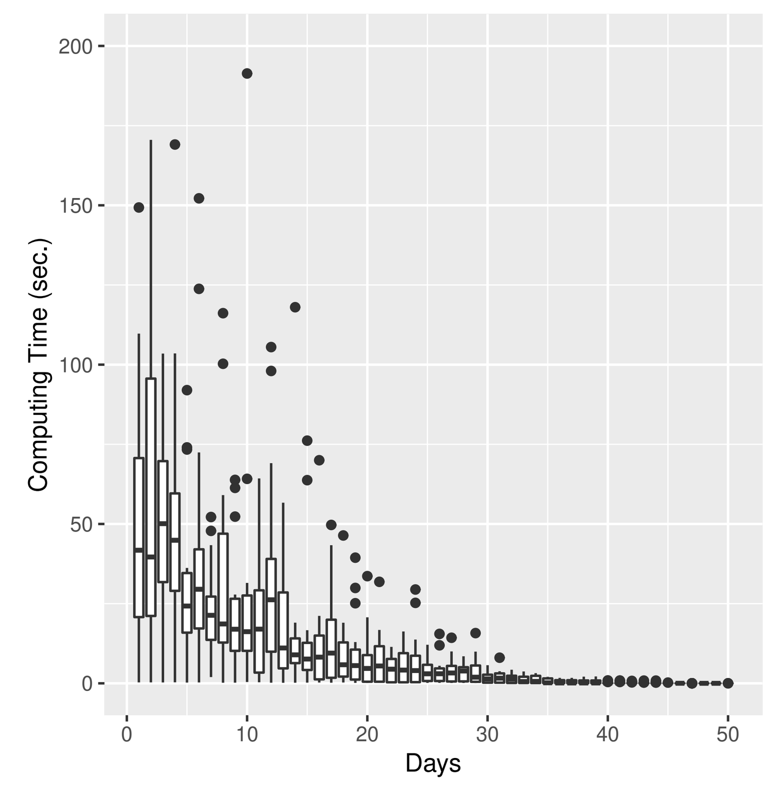
上下界の多層構造の活用

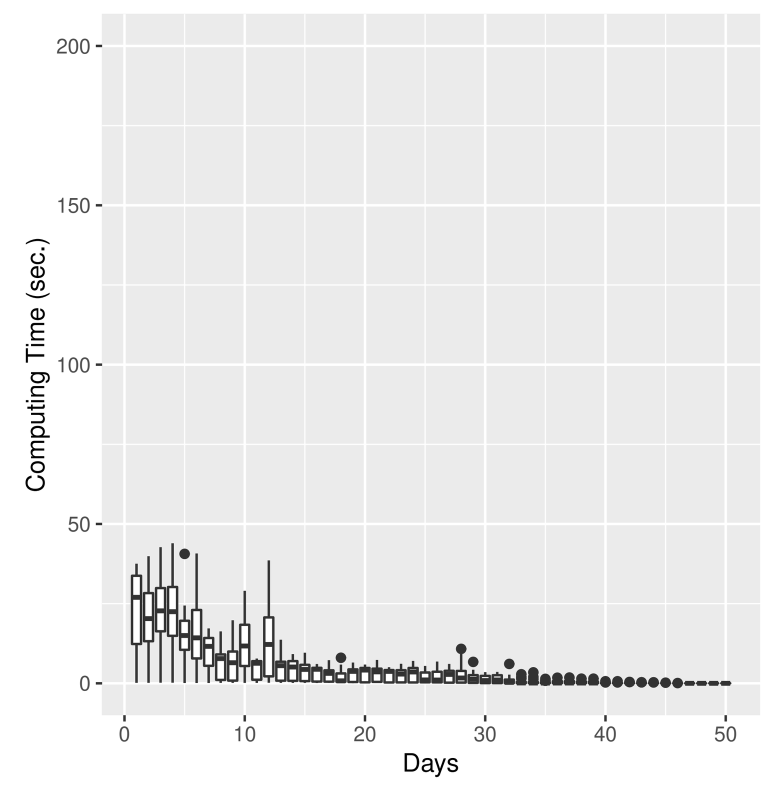
このように構成された整数計画モデルは多くの場合非線形となり，汎用最適化パッケージを用いて直接解くことは必ずしも容易でない．そこで，最大目的関数値の上界と下界をうまく利用することを考える．右図のように，$m-1$ 番目までの順位判定のうちいずれか一つの不等号（$<$）を等号つき（$\le$）に変えることにより $m-1$ 個のレベルの上界が得られ，また，1番目を除く順位判定をそれ以降を含めて取り去る（すなわち順位判定を途中で打ち切る）ことにより $m-1$ 個のレベルの下界が得られる．さらに，m 番目の順位判定を等号なし（$<$）に変えて得られる，よりタイトな下界を用いることにより，勝敗数以外の要素に基づく $m+1$ 番目以降の判定基準を考慮する必要性の有無も判定することができる．



数値実験

B.LEAGUE（男子プロバスケットボール）のB1チャンピオンシップ・トーナメント進出のエリミネーション数について，2016–2017シーズン第15節以降の50日分の勝敗記録に対して数値実験を行った結果を以下に示す．全18クラブのエリミネーション数を計算するのに要した個々の時間を一日単位で箱ひげ図にプロット（四分位点から1.5 IQRを超える外れ値を点で表示）しているが，上下界の情報を使わずに解いた場合（左図）と比較して，これを活用した場合（右図）は，極めて時間がかかるケースを抑えることにより平均計算時間を短縮することに成功している．





参考文献　S. Ito and Y. Shinano (2018), urn:nbn:de:0297-zib-70591