日本オペレーションズ・リサーチ学会
2022年 秋季研究発表会

ジニ係数を最小化する議席配分を求める２つの割当法の比較

01604870　政策研究大学院大学　*諸星穂積　MOROHOSI Hozumi

1.　はじめに

議会などの議席をできるだけ公平に配分する方法として，ジニ係数を最小化する問題に定式化し，過去の衆議院と参議院のデータを用いて，数値実験を行った．

議席配分問題 (apportionment problem) の解法としては，大きく分けて除数法 (divisor method) と割当法 (quota method) がある [1]．現在の常識では，除数法が良いとされ広く使われているが，割当を重視する姿勢も存在する．[2] では，ジニ係数を最小化する割当法を，2 次ナップサック問題として定式化している．本論では，同じ問題を線形整数計画問題として扱い，数値実験により 2 つの方法の比較を行った．

2.　定式化

議会の議院定数を都道府県に配分する問題を考える．県の集合を $\{1,\ldots,s\}$ とし，配分する議席の総数を h とする．各県の人口 $\boldsymbol{p}=(p_1,\ldots,p_s)>0$ が与えられたとき，各県への配分議席数 $\boldsymbol{a}=(a_1,\ldots,a_s)>0$ を決めたい．議席の合計について $\sum_{i=1}^s a_i=h$ が成り立つ．なお総人口を $\pi=\sum_{i=1}^s p_i$ で表す．

割当法では，各県の最低割当数として人口比率を総議席数にかけた値の整数部分，$u_i=\lfloor hp_i/\pi\rfloor$ を保証し，議席数が $u_i\le a_i\le u_i+1$ となるように配分を決める．代表的な割当法である Hamilton 法では，各県に u_i を配分した後に残った議席 $r=h-\sum_{i=1}^s u_i$ を，割当の小数部分 $hp_i/\pi-\lfloor hp_i/\pi\rfloor$ が大きい県から順に 1 議席を配分する．Hamilton 法が与える解 a_i は，$\sum_{i=1}^s(a_i-hp_i/\pi)^2$ を最小化することが知られている．

ここで [2] に倣って，ジニ係数を最小化する割当法を考える．ジニ係数はローレンツ曲線 L と対角線の間の面積の 2 倍：(少し簡略化した記法を使って) $1-2\int L$ として定義されるので，問題は $\int L$ を最大化することと等価である．ここではローレンツ曲線は，$L_0=(0,0)$, $L_k=(\frac{1}{\pi}\sum_{i=1}^k p_i,\frac{1}{h}\sum_{i=1}^k a_i)$, $k=1,\ldots,s$, を直線でつないだ線として定義する．但し，各点 L_k は，$a_1/p_1\le a_2/p_2\le\cdots\le a_s/p_s$ となるように並べ替えられているものとする．この定義では，県 i に住んでいる一人ひとりが a_i/p_i だけの分け前をもらうと考えて，全人口のジニ係数を計算していることになる．

[2] では，配分する議席を $a_i=u_i+x_i$, $x_i\in\{0,1\}$, として，追加配分の有無を表す決定変数 x_i を使って，ローレンツ曲線の（下側の）面積を x_i の 2 次関数で表現することで，制約式 $\sum_{i=1}^s x_i=r$ と合わせて，2 次ナップサック問題を導いている．

ジニ係数の定義としては，ローレンツ曲線を使う以外に，ジニ平均差という値を計算する方法が知られている．ジニ平均差は $X_i=a_i/p_i$ とおいて，$\mathsf{E}\left[|X_i-X_j|\right]$ を計算したものである．ここで期待値は，すべての X_i, X_j のペアについて平均をとることを意味する．以下の関係が成り立つ．

$$1-2\int L=\frac{\mathsf{E}\left[|X_i-X_j|\right]}{2s^2\mathsf{E}\left[X_i\right]}. \tag{1}$$

上式右辺で，期待値は住民一人あたりに関するものなので，分母は以下のように書き下せる．

$$2s^2\mathsf{E}\left[X_i\right]=2s^2\sum_{i=1}^s(a_i/p_i)\times(p_i/\pi)=2s^2h/\pi. \tag{2}$$

つまり，分母は定数になるので，分子のみ考えればよいことになって，分子を書き下すと次のようになる．

$$\begin{aligned}\mathsf{E}\left[|X_i-X_j|\right]&=\frac{1}{\pi^2}\sum_{i,j=1}^s\left|\frac{a_i}{p_i}-\frac{a_j}{p_j}\right|p_ip_j\\&=\frac{1}{\pi^2}\sum_{i,j=1}^s|p_jx_i-p_ix_j+p_ju_i-p_iu_j|.\end{aligned} \tag{3}$$

最終的に，次のような整数計画問題を解くことで

ジニ係数を最小化する．

$$\text{min.} \quad \sum_{i,j=1}^{s} |p_j x_i - p_i x_j + p_j u_i - p_i u_j| \tag{4}$$

$$\text{s. t.} \quad \sum_{i=1}^{s} x_i = r \tag{5}$$

$$x_i \in \{0, 1\}$$

目的関数の絶対値は，よく知られている方法により外すことができて，線形整数計画問題に帰着できる．

3. 数値実験

2000 年から 2015 年までの 4 回の国勢調査による都道府県別人口のデータを使い，その時点での衆議院の選挙区定数（2000, 2005, 2010 年は 300, 2015 年は 295）の配分を計算してみた．ソルバー Xpress を使い最適解を求めたところ，2 次ナップサックと線形整数計画で解の違いはないことを確認した．

次に，2 つの方法の解と，代表的な割当法である Hamilton 法の解と比較したところ，2000 年のみ，1 カ所で違う配分が行われていた．ただし，ジニ係数の差は非常に小さかった．試みに，最適解と実際の議席配分でジニ係数を計算してみた結果を図 1 に示す．参考に，実際の配分議席数から計算したジニ係数を合わせて表示した．ジニ係数を最小化する解と Hamilton 法の解がほとんど同じジニ係数の値を与えること，実際の議席配分に対するジニ係数が，最適解と比べて非常に大きいことが読み取れる．

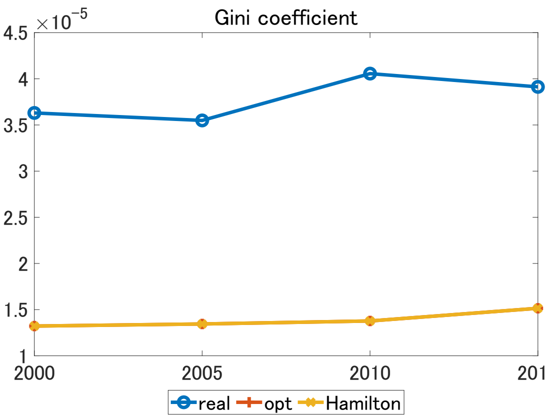

図 1: ジニ係数の値

式 (1) の右辺の分子は，ジニ平均差と呼ばれる量であり，それ自体も格差の程度を表す指標である．この値を，今度は人口で重みづけしないで，県の単位で計算してみる．すはわち，

$$\mathsf{E}\left[|X_i - X_j|\right] = \frac{1}{h^2} \sum_{i,j=1}^{s} \left| \frac{a_i}{p_i} - \frac{a_j}{p_j} \right| \tag{6}$$

を最小化する問題を解いてみた．最適解と Hamilton 法の解を比較すると，2010 年のみ異なり，他の年は同じだった．ジニ平均差を計算した結果を図 2 に示す．実際の議席配分のジニ平均差を調べると，やはり最適解に対して 2 倍以上大きいことがわかった．

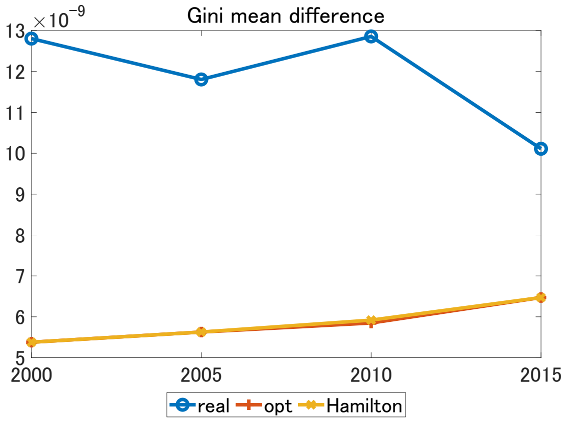

図 2: ジニ平均差の値

4. おわりに

議席配分問題という少し特殊な状況のもとではあるが，ジニ係数の最小化が 2 次ナップサック問題と線形整数計画問題とで，等価な定式化が行えるというのは興味深いと思い，いくつか実験をした結果を報告した．どちらの方法でも短時間で最適解が求まるので，もう少しいろいろな事例に対して適用してみたいと考えている．

参考文献

[1] 一森哲男, 議席配分の数理, 近代科学社, 2018.

[2] D. Pretolani, Apportionments with minimum Gini index of disproportionality: a quadratic knapsack approach, *Ann. Oper. Res*, 215(2014), 257–267.