OD 表の整数推定法：実数推定に帰着する解法

中山 靖子, 米田 清, 和田 維作, 松本 敏明

(株) 東芝

交通網や通信網において，所定時間内の各地点での流入量と流出量から各地点間の traffic を推定する．この問題は，線形制約の下で与えられた行列に最も近い行列を求める matrix balancing 問題 の一つであり，また間接観測からの状態推定問題でもある．

交通の主体が離散である場合には，実数推定よりも整数推定の方が，結果の解釈が容易で分解能が高いので望ましい．本稿では実数推定をくりかえし行うことに帰着する整数推定法を提案し，評価する．提案手法は，従来の metaheuristics を用いた整数推定法と比較して，1) 制約条件を完全に満たす，2) 結果に再現性がある，3) 計算時間が短い，という点で優れている．

Integer Estimation of OD Tables : Reduction to Iterative Real Estimation

Yasuko Nakayama, Kiyoshi Yoneda, Isaku Wada, Toshiaki Matsumoto

Toshiba Corp.

We propose a new algorithm which solves the integer matrix balancing problem. The algorithm has been developed for the application in the integer estimation of *origin-destination* (OD) tables. When the traffic is known to be discrete, integer estimation is more desirable than real estimation especially when the sample size is small. The algorithm is based on the observation that an integer estimation may be constructed if we apply the maximum likelihood principle to the sequence of customer arrival. Evaluation of the algorithm is based on a real-life size numerical example.

1 はじめに

交通網や通信網において, 所定時間内に, どの地点からどの地点にどれだけの traffic があったかを調べたい. これをまとめた, 図 1 のような表を origin-destination table とか OD 表とよぶ. OD 表は網の基礎的な資料であり, これに基づいて設備計画や制御を行なう.

	到着点 $1 \cdots j \cdots C$	合計
1	$x_{11} \cdots x_{1j} \cdots x_{1C}$	$x_{1.}$
$\vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots$
出発点 i	$x_{i1} \cdots x_{ij} \cdots x_{iC}$	$x_{i.}$
$\vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots$
R	$x_{R1} \cdots x_{Rj} \cdots x_{RC}$	$x_{R.}$
合計	$x_{.1} \cdots x_{.j} \cdots x_{.C}$	$x_{..}$

図 1: OD 表

OD 表を作るには, packet とか乗客など traffic の主体に出発点で tag をつけて, 到着点で集計すればよい. たとえば鉄道網の OD 表は原理的には各乗客について出発点と到着点を自動改札でおさえられる. しかしこの調査法は高価で, 実行できない場合が多い. 鉄道でも乗降客数のみを集計し, 発駅と着駅の対に関して集計はしていない. 各地点における流入量と流出量は, 簡単に測定できる場合が多い.

そこで, 各出発点から網内への流入量と, 各到着点からの流出量から, 地点間別の traffic を推定したいという要求がある. すなわち, 図 1 の行和と列和から, 表の中身を推定したい. これは等式よりも変数の方が多い連立一次方程式になるので, 矛盾のない限り解は不定である.

その中から 1 つだけ解を選ぶには, これくらい流量があるだろうという先験的な OD 表をあらかじめ作っておいて, それになるべく近い解を選ぶ. 先験 OD 表の与え方には, さまざまな流儀がある.

OD 表の推定問題は, 線形制約の下で与えられた行列に最も近い行列を求める matrix balancing 問題 [1] の一つであり, また間接観測からの状態推定問題でもある. その他に, OD 表は推移行列 [10] とみなすこともできるし, 投入産出表あるいは産業連関表ともみなせる. 収入支出の会計上のつじつま合わせや統計結果の調整 [3] にも, 同様の問題がある. このように, OD 表の推定と類似の問題は交通や通信のみでなく, 社会学や生物学を含むさまざまな分野に現れる. また, OD 表は 2 方向の分割表 [7] とみなせるので, 他の分割表についても同様の推定問題がある.

多くの場合, OD 表の要素は整数であることがわかっている. たとえば, 交通の主体が離散である場合等である. このような問題に対しては, 実数推定より整数推定の方が, 結果の解釈が容易で分解能が高いので望ましい. 自然数値をとる推定は, 非負実数値をとる推定に加えて整数であるという情報を使っているので, 特に標本が小さい時には, より正確な推定が得られる可能性がある.

しかし, もちろん実数推定の結果を単に丸めただけではつじつまの合う整数推定はできない. 例えば, 1.8 を 2 と丸めると, 与えられた線形制約を満たさない.

OD 表の整数推定法としては，metaheuristics を用いた OD 表の整数推定法 [13, 14] がある．しかし，従来の整数推定法には (1) 制約条件が完全には満たされない，(2) 乱数を用いるので計算結果に再現性がない，(3) 計算時間が長い，といった問題点があった．

ここでは，上の欠点を解消する新しい OD 表の整数推定法 – 構成順序推定法 – を提案する．提案手法は，整数推定を実数推定を繰り返し行うことに帰着する解法である．実数推定の部分は既知の実数推定法のどれを使ってもよい．以降，第 2 節で数値例を紹介し，第 3 節で間接観測問題の定式化を行なう．第 4 節で構成順序推定法を説明し，第 5 節で実用規模の計算例を示し，評価する．最後に第 6 節でまとめと今後の課題を述べる．

2 数値例

実数推定と整数推定の例を示す．表 1は入力データで，表の中身が先験 OD 表，表の最終行が列和，最終列が行和を表している．

表 2は実数推定，表 3は整数推定の計算結果，である．

表 1: 入力データ (先験 OD 表と行和・列和)

OD	1	2	3	4	計
1	1	0	1	1	2
2	1	0	1	1	1
3	0	0	1	1	2
計	1	0	3	1	5

表 2: 実数推定の結果

0.667	0.000	1.000	0.333
0.333	0.000	0.500	0.167
0.000	0.000	1.500	0.500

表 3: 整数推定の結果

1	0	1	0
0	0	1	0
0	0	1	1

3 間接観測問題

間接観測における再構成問題は，一般的には次のように書ける．実数の (m, n) 次元の行列 A と非負の m 次元ベクトル q, それに n 次元ベクトル b が与えられる．通常は $n \leq m$ である．

$$yA = b \tag{1}$$

を満し，かつ q に最も近い非負の m 次元ベクトル y を求める．先験ベクトル q の与え方や，先験ベクトル q と求める ベクトル y との遠さの表し方には，さまざまな方法がある．

Entropy 最大化原理 [4] によって遠さを Kullback-Leibler 情報量で計るならば, 目的関数は

$$\sum y_k \log(y_k/q_k) \longrightarrow \min \tag{2}$$

である.

この場合, 反復尺度法 (iterative scaling) [2] や一般の数理計画法で非負実数の解が求められる.

目的関数に KL 情報量以外の凸関数を選んでも同様で, たとえば文献 [11] で議論されている. 既知の結果の詳細については, たとえば文献 [7, 2, 5, 6, 12, 8, 4, 9] とそれらの引用文献を参照されたい.

4　構成順序推定法

等式制約 (1) の下で, 目的関数 (2) を最小化するような, 0 を含む自然数で構成されたベクトル y を求める.

y を q に近くなるように構成するという静的な視点に加えて, 動的な視点を導入する. すなわち, y は「変数 y_k を 1 つ選んでは 選んだ y_k に 1 を加える」という操作をくりかえして構成するものと考える. 選択するのは, その時点で最大の事後確率を持つ y_k である. これは y に代えて, y の構成順序を Bayes 推定することに相当する. 手順にまとめると, 次のとおり.

構成順序推定法の算法 0:

1. $r := \sum y_k$
2. $y := 0$ とする.
3. 以下を r 回くりかえす.
   (a) 問題 (1), (2) を解き, その実数解 z の最大要素を z_h とする.
   (b) $y_h := y_h + 1$ とする.
   (c) (1) において y_h を含む等式制約に関わる b_l について
   $b_l := b_l - 1$ とする.
   (d) $q_k := z_k$ $(k = 1, ..., m$ 但し $k \neq h)$
   $q_h := (z_h - 1)^+$ とする.
   ここに $x^+ := max(0, x)$.

上記の算法で, (d) において $0 \leq z_h - 1$ の場合, 更新された q は, 更新された b を含む制約式 (1) を既に満たしている. そのため, 次に (a) を実行して得られる実数解 z は, (d) で得られた q と等しくなる. したがって以下のように近道をしても, 同じ結果が得られる.

構成順序推定法の算法 1:

1. $s := \sum y_h$
2. 問題 (1), (2) を解く. その解 z を整数に打ち切って y とする.
3. $q_k := z_k - y_k$ $(k = 1, ..., m)$ とする.
4. $r := s - \sum y_h$
5. **構成順序法の算法 0** の手順 3 以降に同じ.

5 実用規模の計算例

本節では, ある応用での実用規模の OD 表の例を示す.

表 4 は入力を示す. 表の中は正規化されていない先験的 OD 表である. 対角要素と 16 列目の 0 は交通の可能性がないことを, 他の 1 は交通の可能性を示す. Σ 欄は行和, 列和を示す.

表 5 は 反復尺度法 [2] による実数推定の結果を示す. 制約条件は満されている. 収束判定は, 各要素の変化が 10^{-10} より小さくなったことで行なった. 0 でない項目が多く, あいまいである.

表 6 は 従来の整数推定法 [13, 14] の結果を示す. 制約条件が完全には満されていない. 実行には workstation で約 10 時間かかった.

表 7 は 構成順序推定法による整数推定結果を示す. 制約条件は完全に満たされている. 解の質は, 実数推定法, 従来の整数推定法よりも明らかに高い. 0 の項目が多く, 解像度は高い. 実行は 5 時間半で, 従来の整数推定法より短い.

表 4: 20 × 20 OD 表推定の入力データの例

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ
1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	40
2	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1215
3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	33
4	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	26
5	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	5
6	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	150
7	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	71
8	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	44
9	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	28
10	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	54
11	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	166
12	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	68
13	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	185
14	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	38
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	122
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	28
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	100
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	9
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	59
Σ	5	156	54	7	4	129	192	114	174	176	449	213	303	128	121	0	6	163	24	24	2442

表 5: 反復尺度法による実数推定結果

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ
1	0.00	4.69	0.81	0.10	0.06	2.03	2.92	1.72	2.60	2.66	7.18	3.24	4.86	1.92	1.88	0.00	0.09	2.51	0.36	0.36	40.00
2	2.58	0.00	27.78	3.59	2.04	69.61	100.40	58.92	89.40	91.38	246.55	111.27	166.94	66.01	64.53	0.00	3.08	86.22	12.24	12.47	1215.00
3	0.06	3.94	0.00	0.09	0.05	1.70	2.46	1.44	2.19	2.24	6.03	2.72	4.09	1.62	1.58	0.00	0.07	2.11	0.30	0.30	33.00
4	0.05	3.05	0.53	0.00	0.04	1.32	1.90	1.12	1.69	1.73	4.67	2.11	3.16	1.25	1.22	0.00	0.06	1.63	0.23	0.24	26.00
5	0.01	0.59	0.10	0.01	0.00	0.25	0.36	0.21	0.33	0.33	0.90	0.41	0.61	0.24	0.23	0.00	0.01	0.31	0.04	0.04	5.00
6	0.30	18.50	3.19	0.41	0.23	0.00	11.53	6.76	10.26	10.49	28.31	12.77	19.16	7.58	7.41	0.00	0.35	9.90	1.41	1.43	150.00
7	0.14	8.97	1.55	0.20	0.11	3.87	0.00	3.28	4.97	5.09	13.72	6.19	9.29	3.67	3.59	0.00	0.17	4.80	0.68	0.69	71.00
8	0.09	5.38	0.93	0.12	0.07	2.33	3.35	0.00	2.99	3.05	8.24	3.72	5.58	2.21	2.16	0.00	0.10	2.88	0.41	0.42	44.00
9	0.06	3.51	0.60	0.08	0.04	1.51	2.19	1.28	0.00	1.99	5.37	2.42	3.63	1.44	1.40	0.00	0.07	1.88	0.27	0.27	28.00
10	0.11	6.77	1.17	0.15	0.09	2.93	4.22	2.48	3.76	0.00	10.36	4.68	7.02	2.77	2.71	0.00	0.13	3.62	0.51	0.52	54.00
11	0.38	23.68	4.08	0.53	0.30	10.23	14.75	8.66	13.14	13.43	0.00	16.35	24.53	9.70	9.48	0.00	0.45	12.67	1.80	1.83	166.00
12	0.14	8.66	1.49	0.19	0.11	3.74	5.40	3.17	4.81	4.91	13.26	0.00	8.97	3.55	3.47	0.00	0.17	4.63	0.66	0.67	68.00
13	0.40	24.65	4.25	0.55	0.31	10.65	15.36	9.02	13.68	13.98	37.72	17.02	0.00	10.10	9.87	0.00	0.47	13.19	1.87	1.91	185.00
14	0.07	4.67	0.81	0.10	0.06	2.02	2.91	1.71	2.59	2.65	7.15	3.23	4.84	0.00	1.87	0.00	0.09	2.50	0.35	0.36	38.00
15	0.24	14.99	2.58	0.33	0.19	6.47	9.34	5.48	8.32	8.50	22.93	10.35	15.53	6.14	0.00	0.00	0.29	8.02	1.14	1.16	122.00
16	0.00	0.12	0.02	0.00	0.00	0.05	0.07	0.04	0.07	0.07	0.18	0.08	0.12	0.05	0.05	0.00	0.00	0.06	0.01	0.01	1.00
17	0.05	3.29	0.57	0.07	0.04	1.42	2.05	1.20	1.82	1.86	5.03	2.27	3.40	1.35	1.32	0.00	0.00	1.76	0.25	0.25	28.00
18	0.20	12.49	2.15	0.28	0.16	5.40	7.78	4.57	6.93	7.08	19.11	8.63	12.94	5.12	5.00	0.00	0.24	0.00	0.95	0.97	100.00
19	0.02	1.06	0.18	0.02	0.01	0.46	0.66	0.39	0.59	0.60	1.63	0.73	1.10	0.44	0.43	0.00	0.02	0.57	0.00	0.08	9.00
20	0.11	6.97	1.20	0.15	0.09	3.01	4.34	2.55	3.87	3.95	10.67	4.81	7.22	2.86	2.79	0.00	0.13	3.73	0.53	0.00	59.00
Σ	5.00	156.00	54.00	7.00	4.00	129.00	192.00	114.00	174.00	176.00	449.00	213.00	303.00	128.00	121.00	0.00	6.00	163.00	24.00	24.00	2442.00

表 6: Metaheuristics を用いた整数推定法による推定結果

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ
1	0	5	1	0	0	2	3	2	3	3	7	3	5	2	2	0	0	2	0	0	40
2	5	0	29	6	2	69	100	60	87	89	239	109	164	67	64	0	4	85	13	14	1206
3	0	5	0	0	0	2	2	1	2	2	6	3	4	2	2	0	0	2	0	0	33
4	0	4	1	0	0	1	2	1	2	2	5	2	3	1	1	0	0	2	0	0	27
5	0	2	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	5
6	0	17	3	0	0	0	12	7	10	11	29	13	20	8	8	0	1	10	1	1	151
7	0	9	1	0	0	4	0	3	5	5	14	6	10	4	4	0	0	5	1	1	72
8	0	5	1	0	0	2	3	0	3	3	9	4	6	2	2	0	0	3	0	1	44
9	0	4	0	0	0	2	2	1	0	2	6	3	4	1	1	0	0	2	0	0	28
10	0	7	1	0	0	3	4	2	4	0	11	5	7	3	3	0	0	4	1	0	55
11	0	22	4	1	1	10	15	9	13	14	0	16	25	10	10	0	0	13	2	2	167
12	0	8	1	0	0	4	6	3	5	5	14	0	9	4	3	0	0	5	1	1	69
13	0	24	4	0	1	11	16	9	14	14	38	17	0	10	10	0	1	13	2	2	186
14	0	5	1	0	0	2	3	2	3	3	8	3	5	0	2	0	0	2	0	0	39
15	0	14	3	0	0	7	9	6	9	9	23	11	16	6	0	0	0	8	1	1	123
16	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
17	0	3	1	0	0	2	2	1	2	2	5	2	4	1	1	0	0	2	0	0	28
18	0	12	2	0	0	5	8	5	7	7	20	9	13	5	5	0	0	0	1	1	100
19	0	1	0	0	0	0	1	0	1	1	2	1	1	0	0	0	0	1	0	0	9
20	0	7	1	0	0	3	4	2	4	4	11	5	7	3	3	0	0	4	1	0	59
Σ	5	154	54	7	4	129	192	114	174	176	449	213	304	129	121	0	6	163	24	24	2442

表 7: 構成順序推定法による整数推定結果

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Σ
1	0	4	1	0	0	2	4	1	2	2	7	3	4	6	2	0	0	2	0	0	40
2	5	0	27	3	2	69	100	63	89	91	246	111	166	66	64	0	3	86	12	12	1215
3	0	7	0	0	0	3	2	1	2	2	6	2	4	1	1	0	0	2	0	0	33
4	0	3	0	0	0	1	7	1	1	1	4	2	3	1	1	0	0	1	0	0	26
5	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	5
6	0	18	3	0	0	0	11	8	10	10	28	13	19	7	7	0	0	14	1	1	150
7	0	8	1	0	1	4	0	3	11	5	13	6	9	3	3	0	0	4	0	0	71
8	0	5	6	0	0	2	3	0	3	3	8	3	5	2	2	0	0	2	0	0	44
9	0	3	0	0	0	1	2	1	0	5	5	2	3	1	1	0	0	4	0	0	28
10	0	7	1	0	0	7	4	2	4	0	10	4	7	2	3	0	0	3	0	0	54
11	0	23	4	2	1	10	14	8	13	13	0	16	24	9	9	0	2	12	3	3	166
12	0	8	1	0	0	3	5	3	4	4	13	0	16	3	3	0	0	4	1	0	68
13	0	24	4	2	0	10	15	9	13	15	37	17	0	10	10	0	1	13	4	1	185
14	0	4	1	0	0	2	3	1	2	2	7	3	4	0	7	0	0	2	0	0	38
15	0	18	2	0	0	6	9	5	8	8	25	10	15	6	0	0	0	8	1	1	122
16	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
17	0	3	0	0	0	1	2	1	3	3	5	2	3	1	1	0	0	3	0	0	28
18	0	12	2	0	0	5	7	4	6	7	19	8	12	5	5	0	0	0	2	6	100
19	0	1	0	0	0	0	0	0	0	0	1	6	1	0	0	0	0	0	0	0	9
20	0	8	1	0	0	3	4	3	3	5	10	5	7	5	2	0	0	3	0	0	59
Σ	5	156	54	7	4	129	192	114	174	176	449	213	303	128	121	0	6	163	24	24	2442

6 おわりに

状態と観測が 0 を含む自然数のベクトルで表され，状態の情報が線形変換で縮約されて観測される場合の状態推定法として，構成順序推定法を提案した．

構成順序推定法は，整数推定の問題を実数推定をくりかえし行うことに帰着する．その原理は entropy 最大化原理を，解の空間でなく，解を構成する到着順序全体の空間に適用することである．

ここで提案した方法は，従来の metaheuristics を用いた推定方法に比べて，制約条件を完全に満たし，結果に再現性があり，かつ計算時間が短いという点で優れている．

今後は，くりかえし使う実数推定を高速化し，計算時間の短縮を図りたい．また，観測結果が矛盾を含む場合の処理は実用上，大きな障害なので，その解消法を検討したい．

謝辞

Dr. S. Thomas McCormick (The University of British Columbia), 國澤 直樹 氏 (東京電力) に重要な文献を紹介していただき感謝します．

参考文献

[1] M.H. Cheneider and S.A. Zenios. A comparative study of algorithms for matrix balancing. *Operations Research*, Vol. 38, No. 3, pp. 439–455, 1990.

[2] J. N. Darroch and D. Ratcliff. Generalized iterative scaling for loglinear models. *Annals of Mathematical Statistics*, Vol. 43, No. 5, pp. 1470–1480, 1972.

[3] W. E. Deming. *Statistical Adjustiment of Data.* John Wiley & Sons, Inc, 1943.

[4] H. Gzyl. *Maximum Entropy.* World Scientific, 1995.

[5] S. J. Haberman. Log-linear models for frequency tables derived by indirect observation: Maximum likelihood equations. *Annals of Statistics*, Vol. 2, No. 5, pp. 911–924, 1974.

[6] S. J. Haberman. Iterative scaling procedures for log-linear models for frequency tables derived by indirect observation. In *Proceedings of the American Statistical Association*, pp. 45–49, 1975.

[7] C. T. Ireland and S. Kullback. Contingency tables with given marginals. *Biometrika*, Vol. 55, pp. 179–188, 1968.

[8] J. N. Kapur. *Maximum Entropy Models in Science and Engineering.* Springer, 1989.

[9] B. N. Pshenichnyj. *The linearization method for constrained optimization.* Springer, 1994.

[10] T. Sasaki. Probabilistic models for trip distribution. In W. Lewtzbach and P. Baron, editors, *Beiträge zur Theorie des Verkehrsflusses*, pp. 205–210. Bundesminister für Verkehr, 1968.

[11] I. Takahashi. A method for solving network transportation problem with quadratic cost functions. *Waseda University Institute for Productivity*, Vol. 1, pp. 25–31, 1970.

[12] A. G. Wilson. A statistical theory of spatial distribution models. *Transportation Research*, Vol. 1, pp. 253–269, 1967.

[13] K. Yoneda. An integer estimation method for origin-destination tables. ORSA Telecommunications Conference, Boca Raton, FL, 1992.

[14] K. Yoneda. Integer estimation of origin-destination tables. *Trans. Inst. Electric Engineers of Japan*, Vol. 114-C, No. 4 (April), 1994.