

# 卒論テーマ候補：混雑現象の構造推定

池上 慧

2017 年 12 月 2 日

## 1 目的

Viaoux (2007) や柳沼・福田 (2007,8)、松村他などが交通手段の選択と通勤に利用する電車の出発時間選択においてプレイヤー同士の影響を考慮したゲームのモデルを用いたパラメータの推定を行っている。またその結果においては混雑による不効用を示すパラメータは有意に負の値を取っている。主に用いられていた推定手法は Agurregabiria (2004) の PML であり、Kasahara and Shimotsu (2012?) にのっとって推定結果が収束するようにしている。

上記の研究ではプレイヤーについての情報が得られている時にプレイヤーごとに選択確率を計算することでベジアンナッシュ均衡を計算し、推定に用いている。しかし、松村他で用いられたデータはサンプルサイズが小さく、センサスデータであり、また個人の通勤時間などについては別種の統計データから推計したものを利用している。ここではこのような意思決定を行う個人に関するデータを必要とせずに混雑現象を推定できるかを考える。

例えば、松村他で扱われた満員電車に選択モデルを考えると、個人ごとにどのような人がいつどの電車に乗ったかというデータを得るのは非現実的である。そのようなデータの代わりに何時の電車に何人が乗ったという現実的に容易に得られるデータを用いて推定できるモデルを作ることは有用であると考ええる。

## 2 モデル

ラッシュ時の乗車選択行動をモデル化する。全部で  $M$  個の駅について  $T$  日分のラッシュアワーに該当する電車の乗車割合に関するデータが得られており、ある駅  $m$  の  $t$  日について考える。

各駅で毎日繰り返される乗車選択ゲームを以下のように定式化する。プレイヤーは無限人とし、その乗車選択割合だけが混雑を引き起こすとする。 $m$  駅のラッシュアワー時の電車総数を  $C_m$  として、プレイヤーは乗車する電車  $c$  を  $\{1, \dots, C_m\}$  から選ぶ。ただし、これらの中から同時に選択するのではなく、各  $c$  について乗車するかしないかの選択をする最適停止問題のようなゲームをプレイするとする。この際、毎回の選択時に今までどれだけの割合の人が乗車しているかという状態についてはわからない状況を想定し、そのようなゲームにおける完全ベジアン均衡が実現しているとする。

電車  $c$  に対する behavioral strategy を  $b_c = (\tilde{b}_c, 1 - \tilde{b}_c)$  で表記する。ここで  $\tilde{b}_c$  は電車  $c$  に乗車することに割り振る確率であり、 $1 - \tilde{b}_c$  は電車  $c$  には乗らずに次以降の電車のどれかに乗ることに割り振る確率である。belief system を  $\{\mu_c\}_{c=1}^{C_m}$  で表記する。ここで  $\mu_c$  は電車  $c$  への乗車を選択する際の社会状態への信念、つまりまだ乗車していない人の割合の確率分布である。完全ベジアン均衡において belief system は behavioral strategy と矛盾しないものでないといけないため、 $\mu_c = \prod_{c'=1}^{c-1} (1 - \tilde{b}_{c'})$  として behavioral strategy が定まれば一意に定まる。

$t$  日の電車  $c$  についての効用 ( $u_c^t$ ) はについて以下のように定式化する。選好を決定する要素は大きく 3 つ存在する。1 つ目の要素はその電車についての特徴ベクトル ( $Z_c$ ) である。すなわち発車時刻や電車の種類などは乗車の意思決定に大きく作用していることは間違いない。2 つ目の要素はその電車の乗車割合  $\mu_c \tilde{b}_c$  である。混雑は乗車中の不快感や遅延の原因となるため、プレイヤーは混む電車を避けようと行動する。すなわち今選択する対象となっている電車の乗車割合はその電車にひもづく効用を減少させる要因として無視できないものである。3 つ目の要素はまだ乗車していない人の割合  $\mu_c$  である。すなわち不確実な社会状態に効用が依存する。まだ乗車していない人の割合は以降の電車選択において期待できる混雑を予想する際に重要な要因となるため効用に作用する。以上よりそれぞれのパラメータを  $(\beta_m, \gamma_m, \delta_m)$  として以下のように定式化する。

$$\begin{cases} u_c = Z'_c \beta_m + \gamma_m \mu_c \tilde{b}_c + \delta_m \mu_c & \text{when } c \neq C_m \\ u_c = Z'_c \beta_m + \gamma_m \mu_c + \delta_m \mu_c & \text{when } c = C_m \end{cases}$$

$C_m$  への乗車を決定する回まで残ったプレイヤーは必然的に  $C_m$  を選ばなくてはならないので  $c = C_m$  か否かで場合分けする必要があることに注意する。

次に各回の behavioral strategy を二項選択で定式化する。電車  $c$  に関する乗車選択の際にプレイヤーが直面する value function は以下のように書ける。

$$\begin{cases} v_{c,1} &= u_c + \epsilon_{c,1} \\ v_{c,2} &= \tilde{b}_{c+1} u_{c+1} + (1 - \tilde{b}_{c+1}) \tilde{b}_{c+2} u_{c+2} + \cdots + \Pi_{c'=c+1}^{C_m-1} (1 - \tilde{b}_{c'}) u_{C_m} + \epsilon_{c,2} \\ &= \tilde{b}_{c+1} u_{c+1} + \sum_{c'=c+1}^{C_m-1} \left( \Pi_{c''=c+1}^{c'} (1 - \tilde{b}_{c''}) \right) \tilde{b}_{c'+1} u_{c'+1} + \epsilon_{c,2} \end{cases}$$

ここで、 $\epsilon_{c,1}, \epsilon_{c,2}$  はそれぞれ独立に第 1 種極値分布に従う私的情報であり、各電車の選択段階になった時に観測できるものであるとする。 $v_{c,1}$  は電車  $c$  への乗車を決定することの価値であり、 $v_{c,2}$  はここでは乗車を見送ることの価値である。 $v_{c,2}$  は以降の乗車選択でも均衡の behavioral strategy に従った時の期待利得と誤差項から構成されており、ここでの期待利得の計算に際しては以降の意思決定時に明らかになる  $\{\epsilon_{c'}\}_{c'=c+1}^{C_m-1}$  については考慮に入れず、確定している  $\{u_{c'}\}_{c'=c+1}^{C_m-1}$  に関しての期待値を計算していることに注意する。

$t$  日の乗車選択には研究者には観測不可能な要素だがプレイヤーにとっては観測可能である要素が関わっている可能性を考慮する。例えば遅延情報などの情報は各日ごとに変化し、確実にプレイヤーの行動を変化させているが、これらについての情報は研究者には得られない。これらの日に依存する研究者には観測不可能な変数が、選択肢である電車ではなく電車ごとの効用の差分について存在すると仮定する。この仮定は各電車の絶対的な効用水準ではなく効用の差がプレイヤーの意思決定を左右している今の状況では擁護できるものである。最後の電車である  $C_m$  とそれ以外の電車  $c$  との間に存在する研究者には観測不可能だがプレイヤーにとっては観測できる要因によってもたらされる誤差を  $\xi_c^t$  で表記する。この誤差項は  $\epsilon$  とは異なり、毎日のゲーム開始時に観測されており、behavioral strategy を決定する際にも期待値計算に使われるとする。すなわち以下のように誤差が存在する。

$$(u_{C_m} - u_c)^t = (Z_{C_m} - Z_c)' \beta + \gamma(\mu_{C_m} - \mu_c \tilde{b}_c^t) + \delta_m(\mu_{C_m} - \mu_c) + \xi_c^t$$

これより電車  $c$  の乗車選択においては  $\{\xi_{c'}\}_{c'=c}^{C_m-1}$  に依存した誤差項が存在する。通常の二項選択の議論より  $t$  日の電車  $c$  に関する behavioral strategy は以下のように書ける。

$$\tilde{b}_c^t = \frac{1}{1 + \exp \left( \tilde{b}_{c+1}^t u_{c+1} + \sum_{c'=c+1}^{C_m-1} \left( \Pi_{c''=c+1}^{c'} (1 - \tilde{b}_{c''}^t) \right) \tilde{b}_{c'+1}^t u_{c'+1} - u_c \right)}$$

ここで単純な計算により以下を得る。

$$\begin{aligned} & \tilde{b}_{c+1}^t u_{c+1} + \sum_{c'=c+1}^{C_m-1} \left( \Pi_{c''=c+1}^{c'} (1 - \tilde{b}_{c''}^t) \right) \tilde{b}_{c'+1}^t u_{c'+1} - u_c \\ &= (u_{C_m} - u_c)^t - \sum_{c'=c+1}^{C_m-1} \left( \Pi_{c''=c+1}^{c'} (1 - \tilde{b}_{c''-1}^t) \right) \tilde{b}_{c'}^t (u_{C_m} - u_{c'})^t \end{aligned}$$

これより、選択確率の式を変形することで以下の式を得る。

$$\log \frac{1 - \tilde{b}_c^t}{\tilde{b}_c^t} = (u_{C_m} - u_c)^t - \sum_{c'=c+1}^{C_m-1} \left( \Pi_{c''=c+1}^{c'} (1 - \tilde{b}_{c''-1}^t) \right) \tilde{b}_{c'}^t (u_{C_m} - u_{c'})^t \quad (1)$$

状態への信念である  $\{\mu_c\}_{c=1}^{C_m}$  と、各電車への乗車割合である  $w_c^t = \mu_c^t \tilde{b}_c^t$  を用いてこの式を変形する。まず、

$$\left\{ \begin{array}{l} \hat{w}_c^t = (\mu_c^t - w_{c+1}^t \cdots - w_{C_m-1}^t)' \\ \tilde{w}_c^t = (-w_{c+1}^t \cdots - w_{C_m-1}^t)' \\ \bar{Z}_c = (Z_{C_m} - Z_c \ Z_{C_m} - Z_{c+1} \cdots Z_{C_m} - Z_{C_m-1}) \\ \bar{\mu}_c^t = (\mu_{C_m}^t - \mu_c^t \cdots \mu_{C_m}^t - \mu_{C_m-1}^t)' \\ \bar{\nu}_c^t = (\mu_{C_m}^t - w_c^t \cdots \mu_{C_m}^t - w_{C_m-1}^t)' \end{array} \right.$$

とおく。 $d_c^t = \log \frac{1-\bar{b}_c^t}{\bar{b}_c^t}$  とすると、先の式は以下のように書ける。

$$d_c^t = \left( \bar{Z}_c' \beta_m + \gamma_m \bar{\nu}_c^t + \delta_m \bar{\mu}_c^t \right)' \left( \frac{\hat{w}_c^t}{\mu_c^t} \right) + \xi_c^{t*'} \left( \frac{\tilde{w}_c^t}{\mu_c^t} \right) + \xi_c^t \quad (2)$$

ここで、 $\xi_{C_m-1}^{t*} = 0$  であることと、 $\left( \frac{w_{C_m-1}^t}{\mu_{C_m-1}^t} \right) = 1$  であることを利用して、後ろから順に  $\xi_c^t$  を代入していくことで式を一つにまとめることができる。新しく以下の記号を定義する。

$$\left\{ \begin{array}{l} g_c^t = d_c^t - (g_{c+1}^t \cdots g_{C_m-1}^t) \left( \frac{\tilde{w}_c^t}{\mu_c^t} \right) \text{ where } c \neq C_m - 1 \\ g_{C_m-1}^t = d_{C_m-1}^t \end{array} \right.$$

これを用いて、先に述べた逐次代入より以下の 1 本の式を得る。2 本目の等式は 1 本目の選択に際してはまだ誰も電車に乗っていないので  $\mu_1^t = 1$  となることによる。

$$\begin{aligned} g_1^t &= (Z_{C_m} - Z_1)' \beta + (\mu_{C_m}^t - \mu_1^t) \gamma_m + (\mu_{C_m}^t - w_1^t) \delta_m + \xi_1^t \\ &= (Z_{C_m} - Z_1)' \beta + (\mu_{C_m}^t - 1) \gamma_m + (\mu_{C_m}^t - w_1^t) \delta_m + \xi_1^t \end{aligned}$$

以上より各駅の日毎に方程式を得た。例えば  $\xi_1^t$  はどの駅についても 1 本目の電車について述べていることは明らかであるので  $\xi_m^t$  というように表記し直す。この時以下の形で回帰式を表現することができる。

$$g_m^t = (Z_{C_m} - Z_m)' \beta + (\mu_{C_m}^t - 1) \gamma_m + (\mu_{C_m}^t - w_m^t) \delta_m + \xi_m^t$$

これより  $M$  駅  $T$  日分のデータは以下のように行列できる。

$$G = \mathbf{Z}' \beta + \mu \gamma + \nu \delta + \xi$$

ただし、

$$G = \begin{pmatrix} G_1 \\ \vdots \\ G_M \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} \tilde{Z}_1 & & \\ & \tilde{Z}_2 & \\ & & \ddots \\ & & & \tilde{Z}_M \end{pmatrix}, \mu = \begin{pmatrix} \tilde{\mu}_1 & & \\ & \tilde{\mu}_2 & \\ & & \ddots \\ & & & \tilde{\mu}_M \end{pmatrix}, \nu = \begin{pmatrix} \tilde{\nu}_1 & & \\ & \tilde{\nu}_2 & \\ & & \ddots \\ & & & \tilde{\nu}_M \end{pmatrix}, \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_M \end{pmatrix}$$

$$G_m = \begin{pmatrix} g_m^1 \\ \vdots \\ g_m^T \end{pmatrix}, \tilde{Z}_m = \mathbf{1}_T' \otimes Z_m, \tilde{\mu}_m = \begin{pmatrix} \mu_{C_m}^1 - 1 \\ \vdots \\ \mu_{C_m}^T - 1 \end{pmatrix}, \tilde{\nu}_m = \begin{pmatrix} \mu_{C_m}^1 - w_m^1 \\ \vdots \\ \mu_{C_m}^T - w_m^T \end{pmatrix}, \xi_m = \begin{pmatrix} \xi_m^1 \\ \vdots \\ \xi_m^T \end{pmatrix}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_M \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_M \end{pmatrix}, \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_M \end{pmatrix},$$

である。

$\tilde{Z}_m$  には日をまたいでの変動がないのでこのままでは  $\beta_m$  を駅ごとに推定することができない。よって推定の一段階目では  $\gamma, \delta$  を推定し、二段階目で  $\{\beta_m\}_{m=1}^M$  を推定するような二段階推定を考える。一段階目の回帰式は以下のようになる。

$$G = \alpha + \mu\gamma + \nu\delta + \xi$$

ただし、 $\alpha = (\alpha_1 \cdots \alpha_M)' \otimes \mathbf{1}_T$  とする。しかし、 $\mu, \nu$  はモデルの均衡であるため内生変数である。この係数に対して一致推定を行うために操作変数法を用いる。ここで操作変数として使えるためには誤差項とは相関しない、かつ内生的な各電車の選択確率を左右する変数を用いる必要があり、このケースでは 2 番目の以降の電車に関する属性ベクトル  $\{Z_c\}_{c=2}^{C_m}$ 、曜日、天気、同日の前駅における乗車割合などを使うことができる。ここで、 $\xi_t = (\xi_1^t \cdots \xi_M^t)'$  として  $\text{var}(\xi_t) = \Sigma$  と書くと、 $\text{var}(\xi) = \Sigma \otimes \mathbf{I}_T$  とかける。すなわち第一段階は不均一分散の構造を持つため不均一分散に対して頑健な方法で推定する必要があることに注意する。

第一段階で推定した  $(\hat{\gamma}, \hat{\delta})$  を用いて第二段階の推定を行う。逐次代入を行う前の式 (2) に戻り第一段階で推定した二つのパラメータを代入する。これにより駅  $m$  ごとに以下の  $C_m - 1$  本の式を得る。

$$d_{c,m}^t - \hat{\gamma}_m \nu_{c,m}^t \frac{w_{c,m}^t}{\mu_{c,m}^t} - \hat{\delta}_m \mu_{c,m}^t \frac{w_{c,m}^t}{\mu_{c,m}^t} = \left( Z_{c,m}^- \frac{w_{c,m}^t}{\mu_{c,m}^t} \right)' \beta_m + \eta_{c,m}^t \quad (3)$$

ただし、

$$\eta_{c,m}^t = \xi_{c,m}^t *' \frac{w_{c,m}^t}{\mu_{c,m}^t} + \xi_{c,m}^t \quad \text{where } c \neq C_m - 1$$

$$\eta_{C_m-1,m}^t = \xi_{C_m-1,m}^t$$

である。

これを各駅ごとの回帰式とすると、説明変数と誤差項に相関が存在するので一致推定が行えない。上式の左辺を  $f_c^t$  と書き、 $C_m - 1$  回目の乗車選択に関して注目すると、

$$f_{C_m-1}^t = (Z_{C_m} - Z_{C_m-1})' \beta_m + \eta_{C_m-1}^t$$

である。これを用いて各電車の乗車決定において誤差項と相関を持つ項を削除していくことができる。これにより各駅ごとに以下の回帰式を得る。

$$\sum_{c'=c}^{C_m-1} \left( \sum_{j=0}^{c'-c-1} \Pi_{i=1}^{c'-c-j} \frac{w_{c+i+j}^t}{\mu_{c+i-1}^t} \right) f_{c'}^t = (Z_{C_m} - Z_c)' \beta_m + \sum_{c'=c}^{C_m-1} \left( \sum_{j=0}^{c'-c-1} \Pi_{i=1}^{c'-c-j} \frac{w_{c+i+j}^t}{\mu_{c+i-1}^t} \right) \eta_{c'}^t$$

上の式は  $c, t$  についてスタックしたデータに対して OLS で一致推定を行うことができる。ただし、誤差項には相関が存在するので分散不均一に対して頑健な手法を用いることに注意する。以上の処理を駅ごとに行うことで  $\beta$  の一致推定量  $\hat{\beta}$  を得ることができる。

### 3 シミュレーション

上で得たパラメータの推定値  $\hat{\beta}, \hat{\gamma}, \hat{\delta}$  と  $\xi_t$  の分散共分散行列の推定値  $\hat{\Sigma}$  から時刻表変更による乗車選択割合の変化をシミュレーションにより分析する。ラッシュが起こる時間帯はどのように時刻表を変更しようと会社の始業時間によって決められるので電車を配置する時間帯を変更することは本研究では考えない。また遅延は起こらないものとし、始発である中央林間駅をどの電車が何時に発車するかのみですべての駅での発車時刻と電車の種類が決定するとする。さらに、途中の乗り換えによる混雑率の変化はデータから考えることができない。加えて、下車データを用いないため、全員渋谷を目指す時に混雑率がどのように変化するかを調べることに止める。

主に比較するのは以下の状況である。

- 特急など速い電車の配置を変える
- 電車のキャパシティを変更する

これらは比較的容易に変更できる要素である。