行動モデルの推定

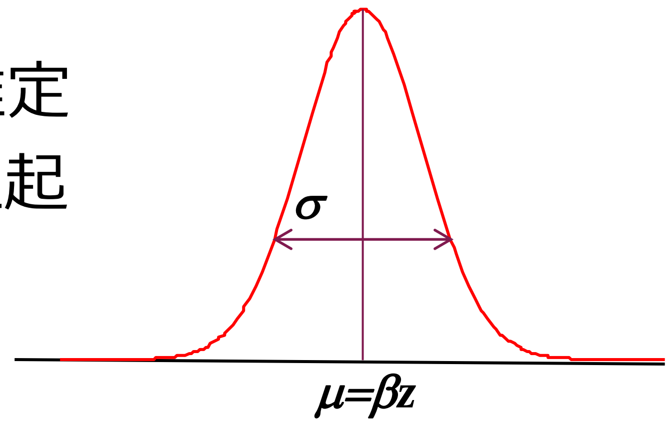
早稲田大学　佐々木邦明

君は，君の行動原理が同時に普遍的な法則となることを欲する
ことができるような行動原理だけに従って行為せよ（カント）

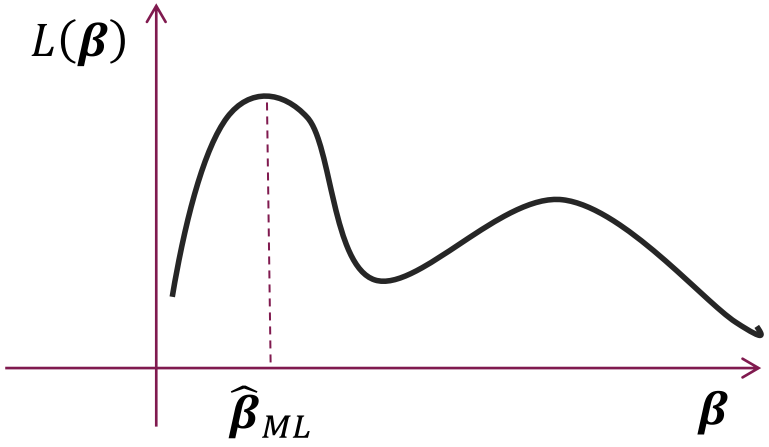
最尤推定

行動モデルの推定と最尤推定

- 有限個のパラメータで記述される確率密度関数の推定
- パラメータベクトル$\boldsymbol{\beta}$，モデル　f　による標本の生起確率を尤度とする
  - $L(\boldsymbol{\beta}) = \prod_{i=1}^{n} f(\boldsymbol{y}_i|\boldsymbol{\beta})$



- (対数)尤度関数が最大になるθ を最尤推定値とする
  - $\widehat{\boldsymbol{\beta}}_{ML} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} \log L(\boldsymbol{\beta})$



最尤推定法

- 点推定量を求める一般的な方法
- 右上の式をθの関数とみなしたものが尤度関数
- 尤度関数を最大化する θの値を最尤推定量とするのが最尤推定法

$$L_n(\boldsymbol{\beta}|x) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\beta})$$

平均値の推定を例にすると

データ(**x：3,5,4**)が得られたとき，
平均をいくつとするのがよいか？

⇒平均がいくつの分布だったら

データ(**x：3,5,4**)がもっとも得られやすいか？

ロジットモデルの最尤推定

- $L(\mu\boldsymbol{\beta}) = \prod_{i=1}^{n} f(\boldsymbol{y}_i|\mu\boldsymbol{\beta})$

選ばれた選択肢の選択確率
y_{ij}=1: if j=選択, y_{ij}= 0:それ以外

- $f(\boldsymbol{y}_i|\mu\boldsymbol{\beta}) = \prod_{j=1}^{J} \left\{ \frac{exp(\mu\boldsymbol{\beta}\boldsymbol{x}_{ij})}{\sum_{j=1}^{J} exp(\mu\boldsymbol{\beta}\boldsymbol{x}_{ij})} \right\}^{y_{ij}}$

βは未知数， x は観測値

- $V_{ij} = \mu\boldsymbol{\beta}\boldsymbol{x}_{ij} = \mu\beta_1 + \mu\beta_1 x_{1i} + \mu\beta_2 x_{2i} \cdots + \mu\beta_K x_{Ki}$

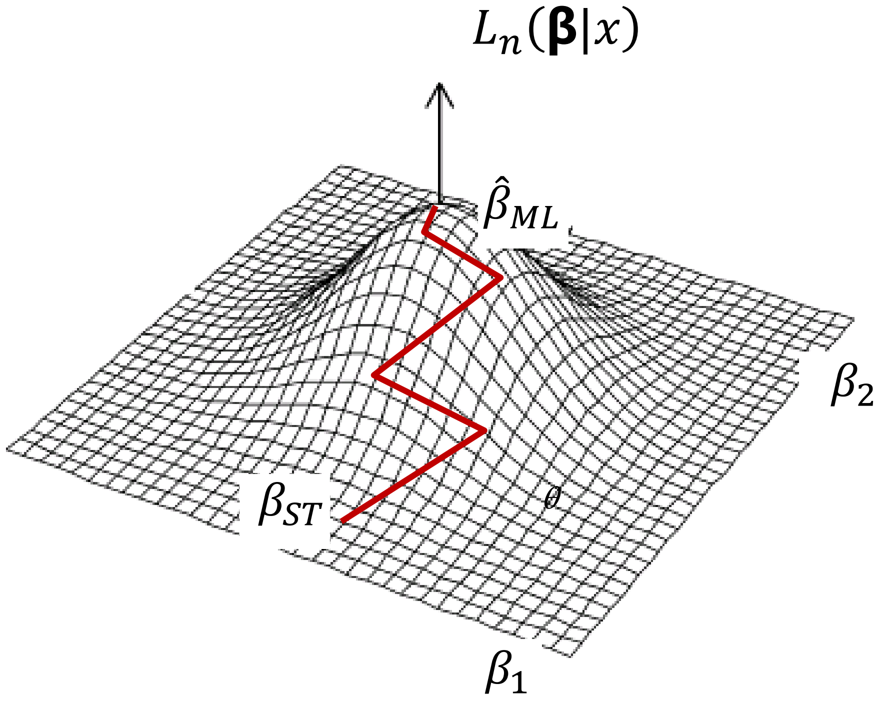
選択結果(y_i ： y_1=車, y_2=車, y_3=鉄道, y_4=鉄道, y_5=鉄道, y_6=車, ・・)が
得られたとき，$\boldsymbol{\mu\beta}$をいくつとすると，データへの適合がよいのか？

⇒μ**β**がいくつだったらデータ(**y**)が得られやすいのか？
μ**β**を色々と変えてみて一番**L**が高くなるμ**β**を探す

最大化アルゴリズムの考え方

周りがあまり見えない中で，近傍の情報から頂点を目指す

- 対数尤度関数の段階的な最大化
  - 初期値を与える
  - 初期値周りで勾配(1次微分）等を用いて次の推定値の方向を決める
  - 初期値付近で１次微分，２次微分を用いて適切に次の点を決めて推定値を得る
  - 収束基準(一次微分ベクトル)で判定し，収束していない場合は，現在の値から次の推定値に移る

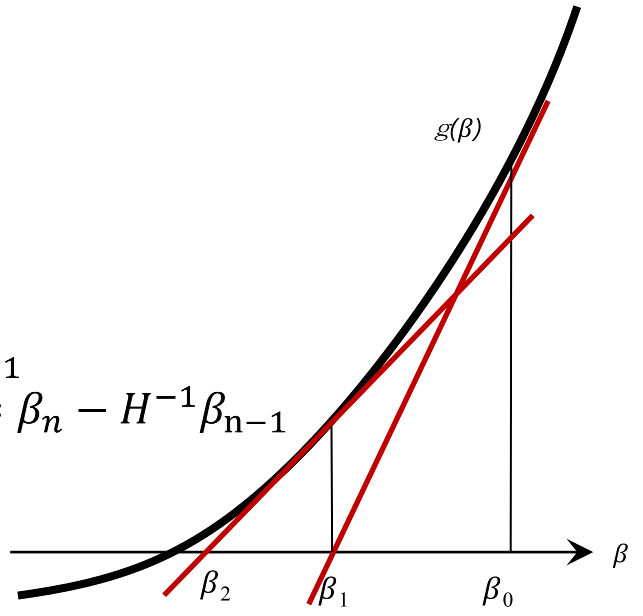


代表的な繰り返し計算法

尤度関数を最大化　尤度関数の一階微分＝０を解く

- Newton-Raphson法
  - テイラー展開の１次近似を利用して進める
- 準Newton法（BFGS，L-BFGS法）
  - ヘッセ行列を，パラメータの差分と一階微分の差分を用いて逐次近似する．
  - L-BFGSはヘッセ行列の更新式を展開して，初期値と差分の関数和で表す．

$$\beta_{n+1} = \beta_n - H^{-1}\beta_{\mathrm{n-1}}$$

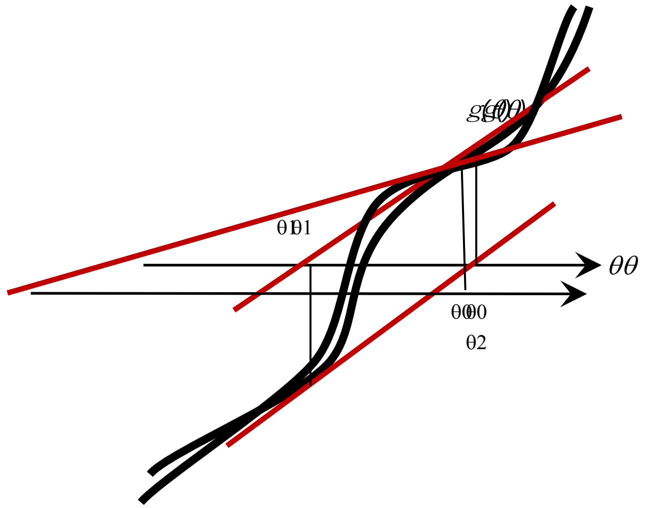


H:尤度関数の二階微分　ヘッセ行列
g: 尤度関数の一階微分

パラメータ推定がうまくいかない

- 収束するとはβ_{n+1}とβ_nが同じになる
  - g′が0になる
- 収束しない
  - 無限に繰り返す
  - β_2が計算不能
- 局所最適解
  - 見かけ上の最大化

- H^{-1}ヘッセ行列の逆行列が早々に死亡
  - **変数が完全相関**
  - **変数が効用関数に影響しないモデル**
- 関数の近似状況
  - 初期値の問題
- モデルに誤り
  - 意思決定者間で異なるが，選択肢間では異ならない変数
  - 選択肢間では異なるが，意思決定者間で異ならない変数



最尤推定法におけるモデル選択

- 真の確率密度関数を近似するものが含まれる必要がある
- 　⇒フレキシブルなモデルを選ぶ

- 最尤推定は自由度の高さ前提
- 　⇒自由度が低すぎるモデルは不適切（過適合）

$$\begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{bmatrix}$$

- 平均対数尤度の比較（KL情報量）
  - 例：共分散行列を考える
    - (非)制約モデル（A対称行列，B対角行列，C対角行列で分散同一）を考えるとCはBに含まれ，BはAに含まれるので，平均対数尤度L*は必ず
    - L*(A)≧L*(B)≧L*(C)になる．

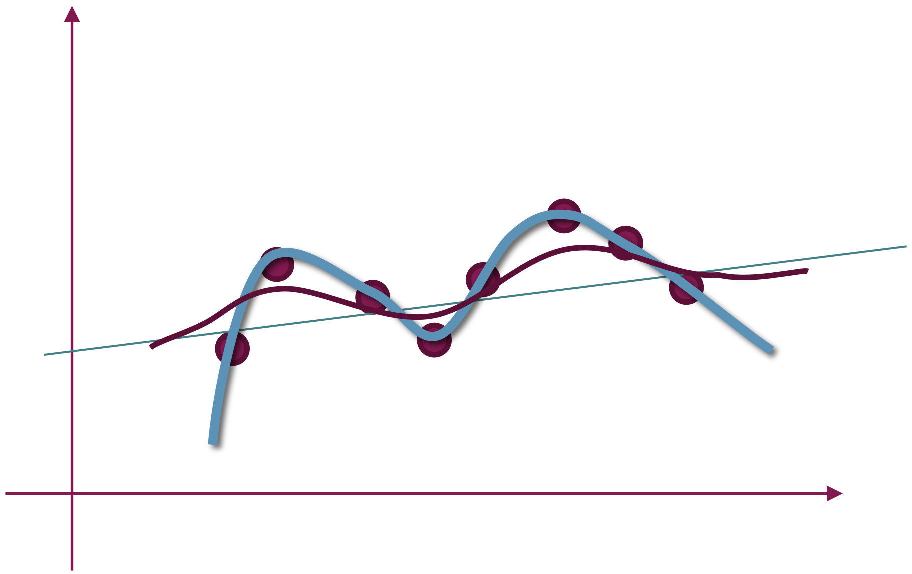
$$\mathrm{AIC} = -\sum_{i=1}^{n} \log f\left(x_i, \hat{\theta}_{ML}\right) + t$$

モデル選択

パラメータ推定と過学習

- 機械学習における学習
  - 判断の根拠となるための統計的なモデルを作る過程
- 機械学習の主な目的は「予測」
  - ある移動手段がどの程度選ばれそうか
  - ある個人が車を購入しそうか

- フィッティング



- 仮説に基づく制約をモデルとして導入せず，予測精度が上がるようにモデルを自由に作る

スパース推定と行動モデル

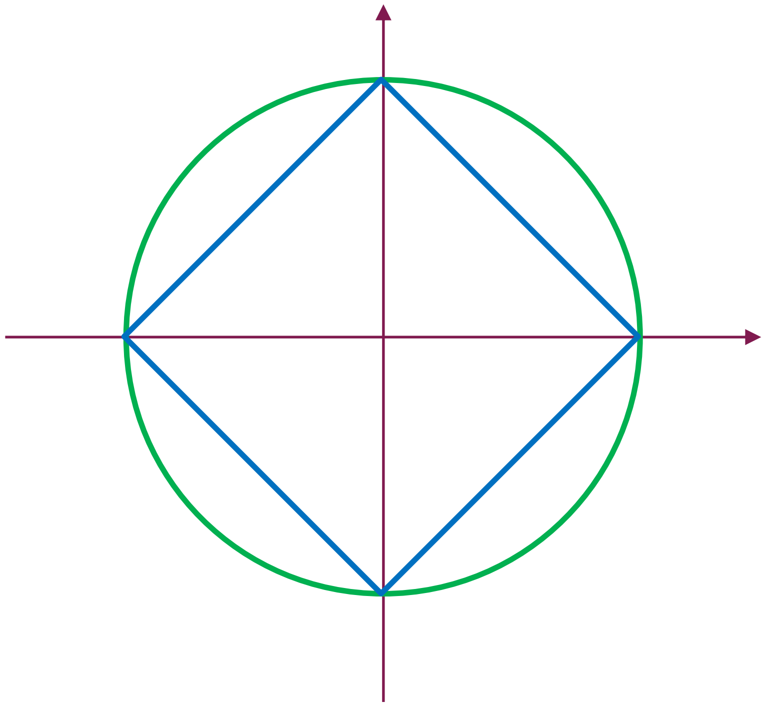
- **スパースモデリング**を用いることでデータが不足していても解析することができる
- **スパースモデリング**を利用することにより大量のデータを使わずに済み，データ解析にかかる時間を短縮できる
- **スパースモデリング**によって複雑なデータ構造をわかりやすく表現する

日高他，2017

- **行動モデル**を用いることで，データが不足していても解析することができる
- **行動モデル**を利用することにより大量のデータを使わずに済み，データ解析にかかる時間を短縮できる
- **行動モデル**によって複雑なデータ構造をわかりやすく表現する

正則化によるスパース推定

- L^2ノルム：$\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
- L^1ノルム：$|x_1| + |x_2| + \cdots + |x_n|$



P=1
P=2

- リッジ推定量
  - $\min_{\beta \in R^d} \|y - x\beta\|_2^2 + \lambda \|\beta\|_2^2$

- Lasso(Least Absolute Shrinkage and Selection Operator)
  - $\min_{\beta \in R^d} \|y - x\beta\|_2^2 + \lambda \|\beta\|_1$

政策分析の方法

行動モデルを用いた政策分析

- **行動モデルを推定し，そのパラメータを用いて，変数の変化による選択の変化を見る．**

- 政策：変数の変化
  - 例えば：**所要時間**を短縮，**駐車場の料金**を割り引く等，政策に対応した変数が必要

- 政策評価
  - 数え上げ法：個人の選択確率を予測して，積み上げる　$S(j) = \frac{1}{N}\sum_{i=1}^{N} P_i(j)$
    - 最大効用の選択肢をカウント
    - 確率の平均値を求める

シミュレーションによる政策分析

- 行動モデルを推定し，そのパラメータを用いて，変数の変化による選択の変化を見る．
- 回遊行動の分析
  - マイクロシミュレーションを用いることで，複雑なモデルの組み合わせを政策評価可能
  - シミュレーションなので，複数回実施して平均的な評価を行う

Step 1	サンプルのデータ，個人属性や発ゾーン，LOSデータなどを各段階のモデルに個人nの個人属性やLOS，各種ダミー等といった説明変数データを代入し，全ての選択肢ごとの選択確率P_{in}を求める．求めた選択確率から確率分布F_{in}を作成する．
Step 2	$[0,1]$の一様乱数γ_nを発生させ，γ_nの値が，$F_{(i-1)n} \leq \gamma_n < F_{in}$を満たす選択肢$_i$を選択するものとする．