応用システム工学
第6回　指数型分布族と一般化線形モデル

平成22年7月09日

パラメータの推定

- 一般化線形モデルのパラメータの点推定・区間推定
- 推定法
  - 最尤推定法　→　これをする
  - 最小二乗法
  - その他
- 求め方
    - 数値解析
    - 収束計算(ニュートンラプソン法)

一般化線形モデルの最尤推定

- 独立なの確率変数Y1,･･･,YN
  - 期待値　$\mu_i = E[Y_i]$
  - 連結関数　$g(\mu_i) = x_i^T \beta = \eta_i$
    - 確率変数に関係するパラメータβの推定
    - xiは要素xij,j=1,2,･･･,pのベクトル
  - 各確率変数Yiの対数尤度関数

$$l_i = y_i b(\theta_i) + c(\theta_i) + d(y_i)$$

期待値・分散

$$\begin{cases} \mathrm{var}[Y_i] = \dfrac{b''(\theta_i)c'(\theta_i) - c''(\theta_i)b'(\theta_i)}{b'(\theta_i)^3} \\ E[Y_i] = \mu_i = -\dfrac{c'(\theta_i)}{b'(\theta_i)} \end{cases}$$

一般化線形モデルの最尤推定

- 全ての確率変数Yiに対する対数尤度関数

$$l = \sum_{i=1}^{N} l_i = \sum_{i=1}^{N} y_i b(\theta_i) + \sum_{i=1}^{N} c(\theta_i) + \sum_{i=1}^{N} d(y_i)$$

  – 対数尤度関数を最大化するパラメータ β_j(最尤推定値)

    - スコア関数 $U(\beta) = 0$

$$\frac{\partial l}{\partial \beta_j} = U_j = \sum_{i=1}^{N} \left[\frac{\partial l_i}{\partial \beta_j} \right] = \sum_{i=1}^{N} \left[\frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \beta_j} \right]$$

    - 非線形関数の解を求める→ニュートンラプソン法

ニュートンラプソン法

- 非線形関数yがx軸と交わる点を求める

$$y(x)=0$$

  - 値 $x^{(m-1)}$ での関数の傾き

$$\left.\frac{dy}{dx}\right|_{x=x^{(m-1)}}=y'\left(x^{(m-1)}\right)\cong\frac{y\left(x^{(m)}\right)-y\left(x^{(m-1)}\right)}{x^{(m)}-x^{(m-1)}}$$

  - $y\left(x^{(m)}\right)=0$ なら $y'\left(x^{(m-1)}\right)=\frac{0-y\left(x^{(m-1)}\right)}{x^{(m)}-x^{(m-1)}}$

    - 収束するまで計算 $x^{(m)}=x^{(m-1)}-\frac{y\left(x^{(m-1)}\right)}{y'\left(x^{(m-1)}\right)}$
      - スコア関数の微分を求める必要有

一般化線形モデルの最尤推定

- スコア関数内要素
  微分第一項を求める $\left[\frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right]$

$$\frac{\partial l_i}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\left[y_i b(\theta_i) + c(\theta_i) + d(y_i)\right] = y_i b'(\theta_i) + c'(\theta_i)$$

$$E[Y_i] = \mu_i = -\frac{c'(\theta_i)}{b'(\theta_i)}$$

  – 期待値の関係を用いる $c'(\theta_i) = -\mu_i b'(\theta_i)$

$$\frac{\partial l_i}{\partial \theta_i} = y_i b'(\theta_i) + c'(\theta_i) = y_i b'(\theta_i) - \mu_i b'(\theta_i) = b'(\theta_i)[y_i - \mu_i]$$

となる

一般化線形モデルの最尤推定

- スコア関数内要素
  微分第二項を求める $\left[\frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right]$

$$\mu_i = -\frac{c'(\theta_i)}{b'(\theta_i)} \quad \text{より}$$

$$\frac{\partial \mu_i}{\partial \theta_i} = \frac{\partial}{\partial \theta_i}\left[-\frac{c'(\theta_i)}{b'(\theta_i)}\right] = -\frac{b'(\theta_i)c''(\theta_i) - b''(\theta_i)c'(\theta_i)}{b'(\theta_i)^2}$$

  – 分散の別表現より

$$\mathrm{var}[Y_i] = \frac{b''(\theta_i)c'(\theta_i) - c''(\theta_i)b'(\theta_i)}{b'(\theta_i)^3}$$

一般化線形モデルの最尤推定

- スコア関数ない要素
  微分第二項に分散を代入

$$\left[\frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right]$$

$$\frac{\partial \mu_i}{\partial \theta_i} = b'(\theta_i)\mathrm{var}[Y_i]$$

$$\frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{\dfrac{\partial \mu_i}{\partial \theta_i}} = \frac{1}{b'(\theta_i)\mathrm{var}[Y_i]}$$

一般化線形モデルの最尤推定

- スコア関数内要素 微分第三項を求める $\left[\frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right]$

連結関数
$x_i{}^T \beta = \eta_i$

$$\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i}\frac{\partial}{\partial \beta_j}\left[x_i{}^T \beta\right] = \frac{\partial \mu_i}{\partial \eta_j} x_{ij}$$

β jだけ残る

- スコア関数

$$U_j = \sum_{i=1}^{N}\left[\frac{\partial l_i}{\partial \theta_i}\frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j}\right] = \sum_{i=1}^{N}\left[b'(\theta_i)[y_i - \mu_i]\frac{1}{b'(\theta_i)\mathrm{var}[Y_i]}\frac{\partial \mu_i}{\partial \eta_i} x_{ij}\right]$$

$$= \sum_{i=1}^{N}\left[\frac{y_i - \mu_i}{\mathrm{var}[Y_i]}\frac{\partial \mu_i}{\partial \eta_i} x_{ij}\right]$$

一般化線形モデルの最尤推定

- スコア関数の微分
  – スコア関数の微分の期待値で近似
    - 情報量を利用

$$E[U'(Y)] = -\mathrm{var}[U(Y)] = -\mathfrak{I}$$

    - スコア関数 U_j の分散共分散行列を情報行列 $\mathfrak{I}$
  – (j,k)要素

$$\mathfrak{I}_{jk} = E[U_j U_k]$$

(スコア統計量の分散は情報量 $\mathfrak{I}$)

  – 確率変数Yiは独立

$$E[(Y_j - \mu_j)(Y_k - \mu_k)] = 0,\ j \neq k$$

一般化線形モデルの最尤推定

情報行列を求める

$$\mathfrak{I}_{jk} = E\left\{\sum_{i=1}^{N}\left[\frac{y_i-\mu_i}{\mathrm{var}[Y_i]}\frac{\partial\mu_i}{\partial\eta_i}x_{ij}\right]\sum_{l=1}^{N}\left[\frac{y_l-\mu_l}{\mathrm{var}[Y_l]}\frac{\partial\mu_l}{\partial\eta_l}x_{lk}\right]\right\}$$

$$= \sum_{i=1}^{N}\left\{\frac{E\left[(y_i-\mu_i)^2\right]}{\mathrm{var}[Y_i]^2}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 x_{ij}x_{ik}\right\}$$

$E\left[(y_i-\mu_i)^2\right] = \mathrm{var}[Y_i]$ より

$\mathfrak{I}_{jk} = \sum_{i=1}^{N}\left[\frac{x_{ij}x_{ik}}{\mathrm{var}[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\right]$ となる

これを行列で表現すると $\mathfrak{I} = X^T wX$ できる

ただしwは $w_{ii} = \frac{1}{\mathrm{var}[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2$ のNxN対角行列

最尤推定のスコア関数表現

- ニュートン・ラプソン法公式 $x^{(m)} = x^{(m-1)} - \dfrac{y\left(x^{(m-1)}\right)}{y'\left(x^{(m-1)}\right)}$
- パラメータ・スコア関数表現 $\theta^{(m)} = \theta^{(m-1)} - \dfrac{U^{(m-1)}}{U'^{(m-1)}}$
  - U'を期待値E[U']で近似 $U' \approx E\left[U'(Y)\right] = -\mathfrak{I}$

$$\theta^{(m)} = \theta^{(m-1)} + \frac{U^{(m-1)}}{\mathfrak{I}^{(m-1)}}$$

一般化線形モデルの最尤推定

- ニュートン･ラプソン法への適用
  - パラメータベクトルbでの表現

$$b^{(m)} = b^{(m-1)} + \left[\mathfrak{I}^{(m-1)}\right]^{-1} U^{(m-1)}$$

- ただし $b^{(m)}$ はm回目反復時におけるパラメータ
  β1, β2,･･･, βpの推定値ベクトル
  $U^{(m-1)}$ はスコアベクトル
  $\left[\mathfrak{I}^{(m-1)}\right]^{-1}$ は情報行列の逆行列
- 逆行列を元に戻す

$$\mathfrak{I}^{(m-1)} b^{(m)} = \mathfrak{I}^{(m-1)} b^{(m-1)} + U^{(m-1)}$$

一般化線形モデルの最尤推定

- 情報行列を代入して要素を求める
  - j要素

$$\left[\mathfrak{I}^{(m-1)}b^{(m-1)}+U^{(m-1)}\right]_j=\sum_{k=1}^{P}\sum_{i=1}^{N}\left[\frac{x_{ij}x_{ik}}{\mathrm{var}[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\right]b_k^{(m-1)}+\sum_{i=1}^{N}\left[\frac{y_i-\mu_i}{\mathrm{var}[Y_i]}\frac{\partial\mu_i}{\partial\eta_j}x_{ij}\right]$$

まとめると　$X^TWXb^{(m)}=X^TWz$　できる

ただし　$z_i=\sum_{k=1}^{P}x_{ij}b_k^{(m-1)}+(y_i-\mu_i)\frac{\partial\mu_i}{\partial\eta_j}$

重み付き最小二乗の変形であり，反復重み付最小二乗法という