応用システム工学
第7回　指数型分布族と一般化線形モデル

平成21年7月17日

重回帰分析

- 重回帰モデル
  - 正規線形モデル
    - 説明変数は全て連続変数
    - デザイン行列
      - 要素が全て1の列
      - 説明変数の観測値からなる列

$$y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Nm} \end{bmatrix}$$

$$E[Y_i] = \mu_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im}$$

$$Y_i \sim N\left(\mu_i, \sigma^2\right)$$

重回帰分析

- 最小二乗による最小モデルを基準に適合度評価
  - モデル $y = X\beta + e$
  - 二乗和基準 $S = \sum_{i=1}^{N} e_i^{\,2} = e^T e = (y - Xb)^T (y - Xb)$
  - 最小二乗推定量 $b = \left(X^T X\right)^{-1} X^T y$

重回帰分析

- 最小二乗基準の最小値

$$
\begin{aligned}
\hat{S} &= (y - Xb)^T (y - Xb) \\
&= y^T (y - Xb) - (Xb)^T (y - Xb) \\
&= y^T y - b^T X^T y - y^T Xb + b^T X^T Xb \\
&= y^T y - b^T X^T y - y^T Xb + b^T X^T X \left(X^T X \right)^{-1} X^T y \\
&= y^T y - b^T X^T y - y^T Xb + b^T X^T y \\
&= y^T y - y^T Xb \\
&= y^T y - b^T X^T y
\end{aligned}
$$

ベクトルの内積 $\langle y, Xb \rangle = y^T Xb = \langle b, X^T y \rangle = b^T X^T y$

重回帰分析

- 最小モデル
  - パラメータβはμだけで構成されるスカラー
  - デザイン行列Xは全て1のN×1ベクトル

$$E[Y_i] = \mu$$

$$X^T X = N$$

$$X^T y = \sum y_i$$

$$b = \frac{1}{N} \sum y_i = \overline{y}$$

重回帰分析

- 最小モデル

$$
\begin{aligned}
\hat{S}_0 &= (y - Xb)^T (y - Xb) \\
&= (y - X\bar{y})^T (y - X\bar{y}) \\
&= y^T y - X^T \bar{y} y - y X \bar{y} + X^T X \bar{y}^2 \\
&= y^T y - N \bar{y}^2 \\
&= \sum (y_i - \bar{y})^2
\end{aligned}
$$

- 最小モデルの最小二乗基準の最小値
  - 観測値の分散に比例
    - 最小二乗基準の最悪値

$$\hat{S}_0 = \sum (y_i - \bar{y})^2$$

重回帰分析

- モデルのデータ適合度の指標
  - 決定係数 $R^2 = \dfrac{\hat{S}_0 - \hat{S}}{\hat{S}_0}$

$$= \frac{y^T y - N\bar{y}^2 - \left(y^T y - b^T X^T y\right)}{y^T y - N\bar{y}^2}$$

$$= \frac{b^T X^T y - N\bar{y}^2}{y^T y - N\bar{y}^2}$$

- データの全変動に対するモデルの割合

重回帰分析

- 決定係数について
  - モデルの適合度が最小モデルの適合度と同等の場合，最小二乗基準の最小値 $\hat{S}$ はデータの全変動 $\hat{S}_0$ とほぼ同じになる
    - 決定係数は0に近い値になる
  - 最大モデルは各観測値Yiにパラメータ μ_i が一つ対応。
    - パラメータベクトル
    - デザイン行列X:NxNの単位行列
    - 最小二乗推定値b=y⇔ $\hat{\mu}_i = y_i$ $b^T X^T y = y^T y$
    - 最小二乗推定量の最小値は0となる
    - 決定係数は1
    - Rを重相関係数

重回帰分析

- モデルの選択
  - 反応変数を説明できる説明変数の必要最小限の部分集合を見つける
    - 逐次的に説明変数を足し引きして検討するステップワイズ回帰
  - ある説明変数が別の説明変数と強く相関している場合は，共線性または多重共線性が存在する
    - 共線性がある場合は，デザイン行列の列ベクトル間に線形従属関係を持つ
    - $X^T X$ は特異に近くなる→ $\left(X^T X\right)b = X^T b$ の解は不安定

重回帰分析

- 説明変数間の多重共線性の判定
  - 分散拡大因子

$$VIF_j = \frac{1}{1 - R^2_{(j)}}$$

  - $R^2_{(j)}$は第j説明変数を反応変数と見なし，その他の変数に関して重回帰を行った場合の決定係数
    - 第j変数と，他の全ての説明変数の間の相関が零

$$VIF_j = 1$$

正規線形モデルに対する分析

簡単な一般線形モデル

- 正規分布を持つ独立な確率変数Y1,･･･,YN

$$E[Y_i] = \mu_i, Y_i \sim N(\mu_i, \sigma^2)$$

  - 連結関数 $g(\mu_i, \sigma^2) = \mu_i$ (恒等連結関数)
  - 正規分布を持つ独立な確率変数e1,･･･,eN
    - 確率変数Yiの確率変数eiによる表現

$$y = X\beta + e$$ デザイン行列X，パラメータβ

$$y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}, X = \begin{bmatrix} x_1{}^T \\ \vdots \\ x_N{}^T \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix}, e = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

簡単な一般線形モデル

- パラメータβの最尤推定量

$$b = \left(X^T X\right)^{-1} X^T y$$

  - ただし $X^T X$ は正則
  - パラメータβの最尤推定量の期待値 $E[b] = \beta$
    - この推定量は不偏推定量
      - 母数のβを推定する時，推定量$\hat{\beta}$について $E\left[\hat{\beta}\right] = \beta$ が成り立つならば，$\hat{\beta}$ をβの不偏推定量という

簡単な一般線形モデル

- パラメータβの最尤推定量
  - 分散共分散行列

$$\sigma^2\left(X^T X\right)^{-1} = \Im^{-1}$$

    - 一般化線形モデルにおいてσ^2は局外パラメータ扱い
    - 分散の不偏推定量

$$\hat{\sigma}^2 = \frac{1}{N-p}\left(y - Xb\right)^T\left(y - Xb\right)$$

      - 分散共分散行列は推定可能

簡単な一般線形モデル

- 最小二乗推定

$$E[y] = X\beta$$

$$E\left[(y - X\beta)^T (y - X\beta)\right] = V$$

  – 評価関数 $S_w = (y - X\beta)^T V^{-1} (y - X\beta)$

$$\frac{\partial S_w}{\partial \beta} = -2X^T V^{-1}(y - X\beta) = 0$$

$$X^T V^{-1} y = X^T V^{-1} X\beta$$

簡単な一般線形モデル

$$X^T V^{-1} X\beta = X^T V^{-1} y$$

$$\tilde{\beta} = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} y$$

– Yは互いに独立かつ分散が同じ場合

$$V = kI$$

$$\tilde{\beta} = \left(X^T X\right)^{-1} X^T y$$

最尤推定量と一致

簡単な一般線形モデル

- 逸脱度

$$
\begin{aligned}
D &= \tfrac{1}{\sigma^2}\left(y - Xb\right)^T \left(y - Xb\right) \\
&= \tfrac{1}{\sigma^2}\left(y^T - \left(Xb\right)^T\right)\left(y - Xb\right) \\
&= \tfrac{1}{\sigma^2}\left(y^T y - \left(Xb\right)^T y - y^T Xb + \left(Xb\right)^T Xb\right) \\
&= \tfrac{1}{\sigma^2}\left(y^T y - b^T X^T y - b^T X^T y + b^T X^T Xb\right) \\
&= \tfrac{1}{\sigma^2}\left(y^T y - 2b^T X^T y + b^T X^T Xb\right)
\end{aligned}
$$

簡単な一般線形モデル

– パラメータの最尤推定量

$$b = \left(X^T X\right)^{-1} X^T y$$

- 逸脱度

$$D = \tfrac{1}{\sigma^2}\left(y^T y - 2b^T X^T y + b^T X^T y\right)$$
$$= \tfrac{1}{\sigma^2}\left(y^T y - b^T X^T y\right)$$

簡単な一般線形モデル

- 逸脱度を用いた仮説検定
  - 帰無仮説H0: $\beta = \beta_0 = \left[\beta_1, \cdots, \beta_q\right]$
    - デザイン行列X0,最尤推定量b0,逸脱度D0
  - 対立仮説H1: $\beta = \beta_1 = \left[\beta_1, \cdots, \beta_p\right]$
    - デザイン行列X1,最尤推定量b1,逸脱度D1
    - ただしq<p<N
  - H0のH1に対する検定に用いる統計量

$$\Delta D = D_0 - D_1 = \tfrac{1}{\sigma^2}\left[\left(y^T y - b_0{}^T X_0{}^T y\right) - \left(y^T y - b_1{}^T X_1{}^T y\right)\right]$$

$$= \tfrac{1}{\sigma^2}\left(b_1{}^T X_1{}^T y - b_0{}^T X_0{}^T y\right)$$

簡単な一般線形モデル

- 対立仮説H1は適合する

$$D_1 \sim x^2\left(N-p\right)$$

- 帰無仮説H0が適合しない場合

$$D_0 \sim x^2\left(N-q,v\right)$$

  – 非心パラメータv

$$\Delta D = D_0 - D_1 \sim x^2\left(p-q,v\right)$$

簡単な一般線形モデル

- F分布

$$F = \frac{D_0 - D_1}{p-q} \Bigg/ \frac{D_1}{N-q}$$

$$= \frac{\left(b_1{}^T X_1{}^T y - b_0{}^T X_0{}^T y\right)}{p-q} \Bigg/ \frac{\left(y^T y - b_1{}^T X_1{}^T y\right)}{N-q}$$

  – 帰無仮説H0が正しい場合中心F分布に従う
  – 帰無仮説H0が正しくない場合非心F分布に従う
    - FがF(p-q,N-p)に対して大きい場合，仮説H0は棄却される

中心F分布

- 中心F分布
  - 独立な中心カイ二乗確率変数
    - X_1^2:自由度n $X_1^2 \sim x^2(n)$
    - X_2^2:自由度m $X_2^2 \sim x^2(m)$
  - 確率変数各々の自由度割ったものの比

$$F = \frac{\frac{X_1^2}{n}}{\frac{X_2^2}{m}} \sim F(n, m)$$

非心F分布

- 二つの独立な確率変数を各々の自由度で割ったものの比
  - 非心カイ二乗分布の確率変数
  - 中心カイ二乗分布の確率変数

$$X_1^{\ 2} \sim x^2(n, \lambda)$$
$$X_2^{\ 2} \sim x^2(m)$$

$$F = \frac{\frac{X_1^{\ 2}}{n}}{\frac{X_2^{\ 2}}{m}}$$

  - 非心F分布の平均は，同じ自由度の中心F分布の平均よりも大きい

簡単な一般線形モデル

- パラメータ間の関係(独立性)
- デザイン行列の直交性
  - 部分行列 $X_1, \cdots, X_m$
    - 互いに直交 $X_j^T X_k = 0, j \neq k$
  - デザイン行列Xが互いに直交する要素で構成される場合，次式のように分割できる

$$X = [X_1, \cdots, X_m], m \leq p$$

  - 分割したパラメータβによる確率変数yの表現
    - Xjは各共変量に対応
      - 共変量とは目的変数に影響を与える定量的な変数

$$E[y] = X\beta = X_1\beta_1 + X_2\beta_2 + \cdots X_m\beta_m$$

簡単な一般線形モデル

- 直交分割されたデザイン行列Xの性質

$$X^T X = \begin{bmatrix} X_1{}^T X_1 & & 0 \\ & \ddots & \\ 0 & & X_m{}^T X_m \end{bmatrix} \qquad X^T y = \begin{bmatrix} X_1{}^T y \\ \vdots \\ X_m{}^T y \end{bmatrix}$$

  – 最尤推定量 $b_j = \left(X_j{}^T X_j\right)^{-1} X_j{}^T y$
    - 独立 $bX^T y = b_1{}^T X_1{}^T y + \cdots + b_m{}^T X_m{}^T y$
    - 仮説も独立に検定可能

$$H_1 : \beta_1 = 0, H_2 : \beta_2 = 0, \cdots, H_m : \beta_m = 0$$

簡単な一般線形モデル

- モデルの残差
  - モデル　$y_i = x_i^{\ T} b + e_i$
  - 当てはめ値　$\hat{\mu}_i$
  - 残差　$\hat{e}_i = y_i - x_i^{\ T} b = y_i - \hat{\mu}_i$
  - 残差ベクトル $\hat{e}$ の分散共分散行列

$$
\begin{aligned}
E\left[\hat{e}\hat{e}^T\right] &= E\left[(y - Xb)(y - Xb)^T\right] \\
&= E\left[yy^T\right] - XE\left[bb^T\right]X^T \\
&= \sigma^2\left[I - X\left(X^T X\right)^{-1} X^T\right]
\end{aligned}
$$

簡単な一般線形モデル

- 射影行列(ハット行列)

$$H = I - X\left(X^T X\right)^{-1} X^T$$

  – 第i対角成分　h_{ii}

- 分散 σ^2 の推定値 $\hat{\sigma}^2$

$$\mathrm{var}(e) = \sigma^2$$

$$\mathrm{var}(\hat{e}) = \sigma^2 (I - H)$$

- 標準化残差

$$r_i = \frac{\hat{e}_i}{\sigma\sqrt{1 - h_{ii}}}$$