応用システム工学
第三回　モデルベース状態推定

平成27年5月8日

E6-111

確率の基本

- A, B:標本区間で定義される事象に対して
  - 同時確率$P(A,B)$ $P(A \cap B)$:AとBが同時に起こる確率
  - 周辺確率$P(A)$, $P(B)$:個別の事象の起こる確率
  - 条件付確率$P(A|B)$:Bが起きた条件下でAが起きる確率
  - 独立: $P(A,B) = P(A)P(B)$が成り立つ
    - $P(A|B) = P(A)$, $P(B|A) = P(B)$
  - 排反: $P(A,B) = 0$が成り立つ
    - $P(A|B) = P(B|A) = 0$

同時確率と周辺確率

- m個の事象$A_1, A_2, \cdots, A_m$とn個の事象$B_1, B_2, \cdots, B_n$に対する同時確率$P\left(A_i, B_j\right)$と周辺確率$P(A_i), P\left(B_j\right)$の関係
  - $P(A_i) = \sum_{j=1}^{n} P\left(A_i, B_j\right)$
  - $P\left(B_j\right) = \sum_{i=1}^{m} P\left(A_i, B_j\right)$
  - $\sum_{i=1}^{m} P(A_i) = 1$
  - $\sum_{j=1}^{n} P\left(B_j\right) = 1$

乗法定理

- $P(A,B)$:AとBの同時確率
  - $P(A,B) = P(A|B)P(B)$
    - Bが起きた後，Bが起きた条件の下でAが起きる条件付確率との積
  - $P(A,B) = P(B|A)P(A)$
    - Aが起きた後，Aが起きた条件の下でBが起きる条件付確率との積
  - $P(A|B) = \frac{P(A,B)}{P(B)}$
  - $P(B|A) = \frac{P(A,B)}{P(A)}$

ベイズの定理

- $P(A, B) = P(A|B)P(B) = P(A|B)P(B)$
  - $P(A|B) = P(A)\frac{P(B|A)}{P(B)}$
    - $P(A)$:仮説に対する信頼の確率
    - $\frac{P(A|B)}{P(B)}$:事象B対する修正項
    - 独立な事象A,B
      - $P(B|A) = P(B)$なので修正項は1（何も修正されない）

統計モデル

- 観測データ$\boldsymbol{y} = \{y(1), y(2), \cdots, y(n)\}$に対する統計モデル
  - $S = \{p(\boldsymbol{y}, \boldsymbol{\theta})\}$
    - S :システムの候補集合
    - $\boldsymbol{\theta}$:未知パラメータ
    - $p(\boldsymbol{y}, \boldsymbol{\theta})$:観測データの確率分布

ベイズ統計

- $p(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{\theta})\frac{p(\boldsymbol{y}|\boldsymbol{\theta})}{p(\boldsymbol{y})}$
  - $p(\boldsymbol{\theta}|\boldsymbol{y})$:yが観測された後の事後分布
  - $p(\boldsymbol{\theta})$:事前分布
  - $p(\boldsymbol{y}|\boldsymbol{\theta})$:観測値から求まる尤度
  - $\frac{p(\boldsymbol{y}|\boldsymbol{\theta})}{p(\boldsymbol{y})}$:観測データによる修正項
  - ベイズ統計は，事前分布を観測データにより修正して事後分布として求める
    - 簡略式: $p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{\theta})p(\boldsymbol{y}|\boldsymbol{\theta})$

ベイズ統計

- 確率変数の期待値
  - $\bar{x} = \mathrm{E}[x] = \int_{-\infty}^{\infty} xp(x)dx$
    - $p(x)$: xの確率密度関数

$p(g(\boldsymbol{\theta})|\boldsymbol{y}) = p\big(g(\boldsymbol{\theta})\big)\frac{p\left(\boldsymbol{y}|g(\boldsymbol{\theta})\right)}{p(\boldsymbol{y})}$より

- パラメータの関数$g(\boldsymbol{\theta})$の事後分布の期待値
  - $\mathrm{E}[g(\boldsymbol{\theta})|\boldsymbol{y}] = \int_{-\infty}^{\infty} g(\boldsymbol{\theta})p(g(\boldsymbol{\theta})|\boldsymbol{y})d\boldsymbol{\theta}$

$$= \int_{-\infty}^{\infty} g(\boldsymbol{\theta})\frac{p\big(g(\boldsymbol{\theta})\big)p\big(\boldsymbol{y}|g(\boldsymbol{\theta})\big)}{p(\boldsymbol{y})}d\boldsymbol{\theta}$$

推定則

- 観測データ$\boldsymbol{y}$からパラメータ$\boldsymbol{\theta}$の推定値$\widehat{\boldsymbol{\theta}}$を求める推定則
  - $\widehat{\boldsymbol{\theta}} = g(\boldsymbol{y})$
- 推定誤差ベクトル
  - $\boldsymbol{e} = \boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}$
- 評価関数(損失関数)
  - $L_d(\boldsymbol{e}) = \|\boldsymbol{e}\|^d, \qquad d = 1,2,\cdots,\infty$
    - $d = 1$:絶対誤差
    - $d = 2$:二乗誤差
    - $d = \infty$:H_∞同定

$$L_\Delta(\boldsymbol{e}) = \begin{cases} 0 & \|\boldsymbol{e}\| \le \frac{\Delta}{2} \\ 1 & \|\boldsymbol{e}\| > \frac{\Delta}{2} \end{cases}$$:一様誤差

ベイズ推定

- ベイズリスク
  - 取得データに対する評価関数の事後分布期待値
    - $R_d\hat{\boldsymbol{\theta}} = \mathrm{E}[L_d(\boldsymbol{e})|\boldsymbol{y}] = \int_{-\infty}^{\infty} L_d(\boldsymbol{e})p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}$
- ベイズ推定値
  - $\hat{\boldsymbol{\theta}}^* = \arg\min_{\hat{\boldsymbol{\theta}}} R_d\hat{\boldsymbol{\theta}}$

arg min →関数値が最小となる定義域の元の集合

    - $d = 1$(絶対誤差): $\hat{\boldsymbol{\theta}}^* = \boldsymbol{\theta}_{\mathrm{median}}$ 事後分布のメジアン(中央値)
      - $\int_{\boldsymbol{\theta}_{\mathrm{median}}}^{\infty} p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta} = \int_{-\infty}^{\boldsymbol{\theta}_{\mathrm{median}}} p(\boldsymbol{\theta}|\boldsymbol{y})d\boldsymbol{\theta}$
    - $d = 2$(二乗誤差): $\hat{\boldsymbol{\theta}}^* = \mathrm{E}[\boldsymbol{\theta}|\boldsymbol{y}]$ 事後分布の条件付期待値
    - 一様誤差: $\hat{\boldsymbol{\theta}}^* = \boldsymbol{\theta}_{\mathrm{mode}} = \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\boldsymbol{y})$ 事後分布のモード(最頻値)

最尤推定（正規分布）

- 確率変数dの推定
- 観測方程式: $y = cx + w$
  - xとwは無相関→観測値yも確率変数となる
- 信号x，雑音w，観測値yの平均値と分散
  - $\mathrm{E}[x] = \bar{x}$, $\mathrm{E}[(x - \bar{x})^2] = {\sigma_x}^2$
  - $\mathrm{E}[w] = \bar{w}$, $\mathrm{E}[(w - \bar{w})^2] = {\sigma_w}^2$
  - $\mathrm{E}[y] = \bar{y} = c\bar{x} + \bar{w}$
  - $\mathrm{E}[(y - \bar{y})^2] = {\sigma_y}^2 = \mathrm{E}[\{c(x - \bar{x}) + (w - \bar{w})\}^2]$
    $= c^2{\sigma_x}^2 + {\sigma_w}^2$

確率変数と確率密度関数

- 二つの独立な確率変数x_1, x_2の和z
  - $z = x_1 + x_2$
- 確率密度関数
  - $p_3(z) = \int_{-\infty}^{\infty} p_1(x_1) p_2(z - x_1) dx_1$
- 正規分布の確率密度関数
  - $p_1(x) = \frac{1}{\sqrt{2\pi {\sigma_x}^2}} e^{-\frac{(x-\overline{x})^2}{2{\sigma_x}^2}}$

$p_2(w) = \frac{1}{\sqrt{2\pi {\sigma_w}^2}} e^{-\frac{(w-\overline{w})^2}{2{\sigma_w}^2}}$

最尤推定(正規分布)

- 累積分布関数:$P(x \le \xi)$, $P(w \le \eta)$
- 確率密度関数
  - $p_1(\xi) = \frac{d}{d\xi}P(x \le \xi)$,$p_2(\eta) = \frac{d}{d\eta}P(w \le \eta)$
  - $p_3(\theta) = \frac{d}{d\theta}P(y \le \theta)$
    $$= \int_{-\infty}^{\infty} p_1(\xi)p_2(\theta - c\xi)d\xi$$
    - $p_3(y) = \int_{-\infty}^{\infty} p_1(x)p_2(y - cx)dx$

最尤推定(正規分布)

- 出力yの確率密度関数
  - $p_3(y) = \frac{1}{\sqrt{2\pi{\sigma_y}^2}} e^{-\frac{(y-\bar{y})^2}{2{\sigma_y}^2}}$

    $= \frac{1}{\sqrt{2\pi(c^2{\sigma_x}^2 + {\sigma_w}^2)}} e^{-\frac{\{y-(c\bar{x}+\bar{w})\}^2}{2(c^2{\sigma_x}^2+{\sigma_w}^2)}}$

    - 線形変換では正規性が保たれる

最尤推定

- 事後確率密度関数$p(x|y)$を最大にする$x = \hat{x}$を最尤推定値とする
  - $p(x|y)$はベイズの定理を用いて求める
    - $p(x|y) = \frac{p_1(x)p_2(y|x)}{p_3(y)}$
  - 信号xに対する観測値yの条件付確率密度関数
    - $p_2(y|x) = \frac{1}{\sqrt{2\pi{\sigma_w}^2}} e^{-\frac{(y-cx-\bar{w})^2}{2{\sigma_w}^2}}$
  - $\hat{x}$を求める

事後確率密度関数

- $$p(x|y) = \frac{\frac{1}{\sqrt{2\pi{\sigma_x}^2}} e^{-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}} \frac{1}{\sqrt{2\pi{\sigma_w}^2}} e^{-\frac{(y-cx-\bar{w})^2}{2{\sigma_w}^2}}}{\frac{1}{\sqrt{2\pi\left(c^2{\sigma_x}^2+{\sigma_w}^2\right)}} e^{-\frac{(y-\bar{y})^2}{2{\sigma_y}^2}}}$$

$$= \sqrt{\frac{c^2{\sigma_x}^2 + {\sigma_w}^2}{2\pi{\sigma_x}^2{\sigma_w}^2}} \frac{e^{-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}-\frac{(y-cx-\bar{w})^2}{2{\sigma_w}^2}}}{e^{-\frac{(y-\bar{y})^2}{2{\sigma_y}^2}}}$$

事後確率密度関数

- $-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}-\frac{(y-cx-\bar{w})^2}{2{\sigma_w}^2}$

$$=-\frac{{\sigma_w}^2(x-\bar{x})^2+{\sigma_x}^2(y-cx-\bar{w})^2}{2{\sigma_x}^2{\sigma_w}^2}$$

$$=-\frac{{\sigma_w}^2(x-\bar{x})^2+{\sigma_x}^2\big(y-cx-(\bar{y}-c\bar{x})\big)^2}{2{\sigma_x}^2{\sigma_w}^2}$$

$$=-\frac{{\sigma_w}^2(x-\bar{x})^2+{\sigma_x}^2\big(y-\bar{y}-c(x-\bar{x})\big)^2}{2{\sigma_x}^2{\sigma_w}^2}$$

事後確率密度関数

- $\sigma_w{}^2(x-\bar{x})^2+\sigma_x{}^2\left(y-\bar{y}-c(x-\bar{x})\right)^2$
  $=(c^2\sigma_x{}^2+\sigma_w{}^2)(x-\bar{x})^2-2c\sigma_x{}^2(x-\bar{x})(y-\bar{y})$
  $+\sigma_x{}^2(y-\bar{y})^2$
  $=(c^2\sigma_x{}^2+\sigma_w{}^2)\left\{(x-\bar{x})^2-\dfrac{2c\sigma_x{}^2}{c^2\sigma_x{}^2+\sigma_w{}^2}(x-\bar{x})(y-\bar{y})\right.$
  $\left.+\left[\dfrac{c\sigma_x{}^2}{c^2\sigma_x{}^2+\sigma_w{}^2}\right]^2(y-\bar{y})^2\right\}+\dfrac{\sigma_x{}^2\sigma_w{}^2}{c^2\sigma_x{}^2+\sigma_w{}^2}(y-\bar{y})^2$
  $=(c^2\sigma_x{}^2+\sigma_w{}^2)\left\{(x-\bar{x})-\dfrac{c\sigma_x{}^2}{c^2\sigma_x{}^2+\sigma_w{}^2}(y-\bar{y})\right\}^2$
  $+\dfrac{\sigma_x{}^2\sigma_w{}^2}{c^2\sigma_x{}^2+\sigma_w{}^2}(y-\bar{y})^2$

事後確率密度関数

- $(x-\bar{x})-\frac{c{\sigma_x}^2}{c^2{\sigma_x}^2+{\sigma_w}^2}(y-\bar{y})=0$において極値をとる
  - $\hat{x}=\bar{x}+\frac{c{\sigma_x}^2}{c^2{\sigma_x}^2+{\sigma_w}^2}(y-\bar{y})$
  - $\sigma^2=\frac{{\sigma_x}^2{\sigma_w}^2}{c^2{\sigma_x}^2+{\sigma_w}^2}$として
  - $\hat{x}=\bar{x}+c\sigma^2(y-\bar{y})$
  - 極値は$\sigma^2(y-\bar{y})^2$
  - $${\sigma_w}^2(x-\bar{x})^2+{\sigma_x}^2\left(y-\bar{y}-c(x-\bar{x})\right)^2$$ $$=\left(c^2{\sigma_x}^2+{\sigma_w}^2\right)(x-\hat{x})^2+\sigma^2(y-\bar{y})^2$$
  - $-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}-\frac{(y-cx-\bar{w})^2}{2{\sigma_w}^2}==-\frac{\left(c^2{\sigma_x}^2+{\sigma_w}^2\right)\left(x-\hat{x}\right)^2+\sigma^2\left(y-\bar{y}\right)^2}{2{\sigma_x}^2{\sigma_w}^2}$
- $\frac{c^2{\sigma_x}^2+{\sigma_w}^2}{2\pi{\sigma_x}^2{\sigma_w}^2}=\frac{1}{2\pi\sigma^2}$

事後確率密度関数

- 事後確率密度関数

  – $p(x|y) = \sqrt{\frac{c^2\sigma_x{}^2+\sigma_w{}^2}{2\pi\sigma_x{}^2\sigma_w{}^2}} \frac{e^{-\frac{(x-\bar{x})^2}{2\sigma_x{}^2}-\frac{(y-cx-\bar{w})^2}{2\sigma_w{}^2}}}{e^{-\frac{(y-\bar{y})^2}{2\sigma_y{}^2}}}$

$$= \sqrt{\frac{1}{2\pi\sigma^2}} \frac{e^{-\frac{\left(c^2\sigma_x{}^2+\sigma_w{}^2\right)(x-\hat{x})^2+\sigma^2(y-\bar{y})^2}{2\sigma_x{}^2\sigma_w{}^2}}}{e^{-\frac{(y-\bar{y})^2}{2(c^2\sigma_x{}^2+\sigma_w{}^2)}}}$$

$$= \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{(x-\hat{x})^2}{2\sigma^2}}$$

事後確率密度関数

- $\sigma^2 = \frac{{\sigma_x}^2{\sigma_w}^2}{c^2{\sigma_x}^2+{\sigma_w}^2} = \frac{1}{1+c^2\frac{{\sigma_x}^2}{{\sigma_w}^2}}{\sigma_x}^2 < {\sigma_x}^2$
  – 信号の分散より推定誤差の分散のほうが小さい
- 観測雑音の分散が0
  – $\lim_{{\sigma_w}^2\to 0}\frac{\sigma^2}{{\sigma_w}^2} = \frac{1}{c^2}$
  – $\hat{x} \to \bar{x} + \frac{1}{c}\{y - (c\bar{x} - \bar{w})\} = \frac{y-\bar{w}}{c}$
  – オフセット外乱の場合は${\sigma_w}^2 = 0,\ y = cx - \bar{w}$
    - 信号は推定値に一致$x = \frac{y-\bar{w}}{c}$