事前情報と回帰分析

メタデータ	言語: Japanese 出版者: 公開日: 2017-10-03 キーワード (Ja): キーワード (En): 作成者: 平館, 道子 メールアドレス: 所属:
URL	http://hdl.handle.net/2297/24101

事 前 情 報 と 回 帰 分 析

平　館　道　子

1．はじめに

ベイジアンの観点からは，統計分析は標本データからの情報だけでなく，それ以外の事前の情報に条件づけられる。標本情報は一定であるとしても，さまざまな事前の情報のあり方に従って分析結果は異る。完全な分析はデータに関する確率モデルとパラメータに関する事前確率分布が詳細に規定された時に可能である。しかしそれほど詳細な事前分布を設定できる程度に事前情報が存在することはなかなかないものである。従って情報が限られている時にどの程度まで分析を進めることが可能であるか、明らかにすることは意味があるであろう。以下では，典型的な場合について考察する。

2．ベイジアン回帰分析

線形回帰モデルの下では，回帰係数 β，誤差分散 σ^2 の事後分布は，従属変数 y, 独立変数 X が与えられた時, 事前分布 $P(\beta, \sigma^2)$ と標本分布 $P(y,x \mid \beta,\sigma^2)$ とから

$$P(\beta,\sigma^2 \mid x,y)=\frac{P(\beta,\sigma^2)P(x,y \mid \beta,\sigma^2)}{P(x,y)}$$

と導かれるが，もしXが β と σ^2 とに独立な分布に従うならば，

$$P(\beta,\sigma^2 \mid x,y)=\frac{P(\beta,\sigma^2)P(y \mid \beta,\sigma^2,x)P(x)}{P(x)P(y \mid x)}$$

$$=\frac{P(\beta,\sigma^2)P(y\mid\beta,\sigma^2,x)}{P(y\mid x)}$$

となる。従ってXが β, σ^2に依存しなければ, Xが確率的に変動するかどうかは分析に関係しないことがわかる。従って上式の分子の第2の確率分布 P（y | β, σ^2, X）に関して次の様に線形回帰モデルを設定しよう。

従属変数yのT個の観測値が独立変数 x を通じて,

$$y_i=x'_i\beta+u_i,\quad i=1,\cdots\cdots,T$$

のように定まり, u_i は x が与えられた時互いに独立に確率分布 $P(u\mid\sigma^2,x)$ に従う確率変量である。ここに x_i はk×1ベクトル, βはk×1の回帰係数ベクトルである。さらに $y'=(y_1,\cdots\cdots,y_T)$, $x=(x'_1,x'_2,\cdots\cdots x'_T)$ であるとする。もしu | x が正規分布N (0, σ^2) に従うならば, T個の観測値による尤度関数は次の通りである。

$$P(y\mid x,\beta,\sigma^2)\propto(\frac{1}{\sigma})^T\exp\{-\frac{1}{2\sigma^2}(y-x\beta)'(y-x\beta)\}$$
$$=(\frac{1}{\sigma})^T\exp\{-\frac{1}{2\sigma^2}[ESS+(\beta-b)'x'x(\beta-b)].$$

bは最小2乗推定量 $(x'x)^{-1}x'y$, $ESS=(y-xb)'(y-xb)$ である。次に β の事前分布がN (b_0, V_0), $H_0=V_0^{-1}$ とすると, σ^2 が与えられた時の β の事後確率分布は

$$P(\beta\mid y,x,\sigma^2)\propto\exp\{-\frac{1}{2}[(\beta-b_1)'(H_0+H)(\beta-b_1)]\}$$

となる。ここに $H=\sigma^{-2}(x'x)$, $b_1=(H_0+H)^{-1}(H_0b_0+Hb)$.

従って σ^2 が所与である時の β の事後期待値は $b_1=(H+H_0)^{-1}(Hb+H_0b_0)$ であり, 最小2乗推定値bと事前期待値 b_0 の行列加重平均となっている。

σ^2 の値が知られていない場合（通常はそうであるが）には,（β, σ^2）に事前分布が付与されなければならない。事前分布としては, Raiffa-Schlaifer（7）が提唱している自然共役分布がしばしば用いられるが, それはこの場合正規ーガンマ分布であり, またその特別な場合である曖昧事前分布（vague prior）もよく用いられる。更に正規分布のもつ特殊性の検討にもとづいたt-ガンマ分布, g-prior（Zellner（8））が線形回帰モデルで用いられる主な事前分布である。

1.1　正規ーガンマ事前分布

事前分布は

$P(\beta,\sigma^{-2})=P_N(\beta\mid b_0,\sigma^2H_0^{-1})P_\gamma(\sigma^{-2}\mid\nu_0,s_0^2)$

である。注意すべき点は

$P(\beta,\sigma^{-2})=P(\beta\mid\sigma^{-2})P(\sigma^{-2})$

のように，β の分布が σ^2 に依存していることである。$\sigma^{-2}=h$ とおくと，この事前分布は,

$$P(\beta,\sigma^{-2})\propto h^{\frac{k}{2}}\exp\{-\frac{h}{2}(\beta-b_0)'H_0(\beta-b_0)\}h^{\frac{\nu_0}{2}-1}\exp\{-\frac{1}{2}\nu_0s_0^2h\}$$

である。事後分布は従って,

$$P(\beta,h\mid x,y)\propto h^{\frac{k+\nu_1}{2}-1}\exp\{-\frac{h}{2}(\beta-b_1)'(H_0+x'x)(\beta-b_1)\}\exp\{-\frac{1}{2}\nu_1s_1^2h\}$$

となり，この分布のパラメータは,

$b_1=(H_0+x'x)^{-1}(H_0b_0+x'xb)$

$H_1=H_0+x'x$

$\nu_1=\nu_0+T$

$s_1^2=\nu_1^{-1}[\nu_0s_0^2+(y-xb)'(y-xb)+(b-b_0)'H_0(H_0+x'x)^{-1}x'x(b-b_0)]$

である。

いま $H_0=0$，$\nu_0=0$ とすると（0 はゼロ行列をあらわす）

$P(\beta,h)\propto h^{\frac{k}{2}}$

となって，これは曖昧事前分布と呼ばれる。あるいは全く情報がない無知の状態を示すものと云われることもある（Jeffreys（3））。しかし無知であることを定義することは困難で，この事前分布をめぐってさまざまな議論が行われて来た。この曖昧事前分布の場合事後分布のパラメータは,

$b_1=b$

$H_1=x'x$

$\nu_1=T$

$s_1^2=\frac{ESS}{T}$

である。

上の事後分布から β の周辺事後分布を求めると k 次元 t 分布が得られる。

すなわち，

$$P(\beta\mid x,y)\propto(\nu_1 s_1^2+(\beta-b_1)'H_1(\beta-b_1))^{-\frac{\nu_1+k}{2}}.$$

曖昧事前分布（$H_0=0$，$\nu_0=0$）の場合，β の事後期待値は b であるが，ν_1 は T であって，T-k ではないことに注意すべきである。古典的な結果が包含されていないことになるからである。

1.2　t-ガンマ事前分布1

自然共役分布は（β，σ^2）の事前分布がすでに述べたようにP（β ｜ σ^2）・P（σ^2）という形で付与されており，β の分布は σ^2 に依存しているが，これは数学的な便利さは別として，あまり自然なこととは云えない。そこで β は σ^2 とは独立に多変量 t 分布に従い，$h=\sigma^{-2}$ がガンマ分布に従うと付与される場合を考えよう。すなわち事前分布は次のようであるとする。

$$P(\beta,h)\propto\{\nu^*+(\beta-b_0)'H^*(\beta-b_0)\}^{-\frac{\nu^*+k}{2}}h^{\frac{\nu_0}{2}-1}\exp\{-\frac{1}{2}\nu_0 s_0^2 h\}.$$

従って事後分布は

$$\begin{aligned}P(\beta,h\mid x,y)\propto&\exp\{-\frac{h}{2}(\beta-b_0)'x'x(\beta-b_0)+ESS\}h^{\frac{T}{2}}\\&\times\{\nu^*+(\beta-b_0)'H^*(\beta-b_0)\}^{-\frac{\nu^*+k}{2}}\\&\times h^{\frac{\nu_0}{2}-1}\exp\{-\frac{1}{2}\nu_0 s_0^2 h\}.\end{aligned}$$

β の周辺事後分布は次の様な2つの t 分布の積になる。

$$P(\beta\mid x,y)\propto\{\nu^*+(\beta-b_0)'H^*(\beta-b_0)\}^{-\frac{\nu^*+k}{2}}\{\nu_1+(\beta-b)'H_1(\beta-b)\}^{-\frac{\nu_1+k}{2}}$$

$$H_1=s_1^{-2}x'x,\qquad s_1^2=\nu_1^{-1}(\nu_0 s_0^2+ESS),$$

$$\nu_1=\nu_0+T-k.$$

$H^*=0$，$\nu_0=0$ の場合は，

$$P(\beta\mid x,y)\propto\{\nu_1+(\beta-b)'H_1(\beta-b)\}^{-\frac{\nu_1+k}{2}}$$

で，β の事後期待値は b，自由度 $\nu_1=T-k$, $s_1^2=\dfrac{ESS}{T-k}$ である。自然共役分布の結果と比較すると，自然共役分布が自由度の減少を許さないために，古典的な結果を再現しないのに対し，t-ガンマの場合は古典的な結果を含んでいる。すなわち個々の β について，

$t=\frac{\beta_i-b}{s_1\sqrt{[x'x]_{ii}}}$ は自由度 T-k の t 分布に従う。

1.3　若干のコメント

β が σ^2 と独立でないということは, β に関する判断は y を発生させるプロセスの精度に依存することを意味する。正規ーガンマ型ではプロセスがより不確実な場合 β の不確実性もより大きくなることになろうが，それが事前確率付与の実際を反映するかどうかは一概には云えないであろう。t-ガンマ型では β は σ^2 と独立であるから, β に関する判断はプロセスとは関係なく下せる。β に関しては，プロセス以外にさまざまな情報があり，そのような情報を反映することが恐らく重要であろう。また自然共役事前分布では，β の周辺事後分布は単峰型であり，ベイズ推定値である事後期待値は最小２乗値と事前期待値の行列加重平均である。このことは事前情報とデータ情報が著しく乖離している場合にも、事前情報を同じプロセスからの新たなデータとして組み入れてしまうことを意味すると云えよう。これに対して t-ガンマ型の場合は２種類の情報を異質なもの，すなわち，精度の異なるプロセスからのデータとして扱う。情報が異質であると判断される場合それらをプールするとすれば，結果が２つの t 分布になるのは自然なことのように思われる。

1.4　t-ガンマ事前分布２

t 分布の積の形の分布は poly-t 分布と云われるが，一般にはその特性はまだ知られていない。２つの t について次のように接近することにしよう。これは J. Dickey（１）の方法である。

β が v が与えられている時 N(b_0, v^2V) に従うとして, 更に v^{-2} がパラメータ ν', s′ の χ^2 分布に従うとすると, β の周辺分布は t 分布である。すなわち,

$$P(\beta)\propto\int\left(\frac{1}{v^2}\right)^{\frac{k}{2}}\exp\left\{-\frac{1}{2v^2}(\beta-b_0)'V^{-1}(\beta-b_0)\right\}\cdot(v^{-2})^{\frac{\nu'}{2}-1}\exp\left\{-\frac{1}{2}\nu's'^2v^{-2}\right\}dv^{-2}$$
$$\propto[\nu'+(\beta-b_0)'H(\beta-b_0)]^{-\frac{\nu'+k}{2}}$$
$$H=s'^{-2}V^{-1}.$$

従ってこの β の周辺分布は, 多変量正規分布の一般化であり, β は逆 χ^2 分布に従う v^2 を未知の分散スケールとする多変量正規分布という超母集団に由来

するものと解釈することができる。

この時線形回帰モデルの分散σ^2が前述したガンマ分布に従うとして，これをも合せて，(β, v^2, σ^2) は，$b_0, H_0, \nu_0, s_0, \nu', s'$ をパラメーターとする次の分布に従う。

$$
\begin{aligned}
P(\beta, v^{-2}, \sigma^{-2}) \propto & \exp\{-\frac{1}{2}v^{-2}(\beta-b_0)'H_0(\beta-b_0)\} \\
& \times v^{-2\frac{\nu'+k}{2}-1}\exp\{-\frac{1}{2}\nu' s'^2 v^{-2}\} \\
& \times \sigma^{-2\frac{\nu_0}{2}-1}\exp\{-\frac{1}{2}\nu_0 s_0^2 \sigma^{-2}\}.
\end{aligned}
$$

更に事後分布は次のように求められる。

$$
\begin{aligned}
P(\beta, v^{-2}, \sigma^{-2} \mid x, y) \propto & \exp\{-\frac{1}{2}[\sigma^{-2}(\beta-b)'x'x(\beta-b) \\
& + v^{-2}(\beta-b_0)'H_0(\beta-b_0)]\} \times v^{-2\frac{\nu'+k}{2}-1}\exp\{-\frac{1}{2}v^{-2}\nu' s'^2\} \\
& \times \sigma^{-2\frac{\nu_1+k}{2}-1}\exp\{-\frac{1}{2}\nu_1 s_1^2 \sigma^{-2}\}.
\end{aligned}
$$

ここに　$\nu_1 = T - k + \nu_0$

$\nu_1 s_1^2 = \nu s^2 + \nu_0 s_0^2$

$\nu s^2 = ESS.$

ここで $\omega = \dfrac{\sigma^{-2}}{\sigma^{-2}+v^{-2}} = \dfrac{v^2}{\sigma^2+v^2}$ とおくと，$x'x = H'$ とすれば

$$
\begin{aligned}
& \exp\{-\frac{1}{2}[\sigma^{-2}(\beta-b)'H'(\beta-b) + v^{-2}(\beta-b_0)'H_0(\beta-b_0)]\} \\
& = \exp\{-\frac{1}{2}[\sigma^{-2}\omega^{-2}(\beta-b_\omega)'H_\omega(\beta-b_\omega) \\
& + \sigma^{-2}(1-\omega)(b-b_0)'H'H_\omega^{-1}H_0(b-b_0)]\}
\end{aligned}
$$

となる。ここに

$H_\omega = \omega H' + (1-\omega)H_0$

$b_\omega = H_\omega^{-1}[\omega H' b + (1-\omega)H_0 b_0] = H_\omega^{-1}\omega H'(b-b_0) + b_0.$

したがって v から ω への変換におけるヤコビアンは$J(\dfrac{\partial v}{\partial \omega}) = \dfrac{1}{2}\sigma\omega^{-\frac{1}{2}}(1-\omega)^{-\frac{1}{2}}$ となるから，

(β, σ, ω) の事後分布は次の様になる。

$$
\begin{aligned}
P(\beta, \sigma, \omega \mid x, y) \propto & \exp[-\frac{1}{2}\{\sigma^{-2}\omega^{-1}(\beta-b_\omega)'H_\omega(\beta-b_\omega) \\
& + \sigma^{-2}(1-\omega)(b-b_0)'H'H_\omega^{-1}H_0(b-b_0)\}] \\
& \times \sigma^{-(\nu_1+k+1)}\exp[-\frac{1}{2}\nu_1 s_1^2 \sigma^{-2}](\frac{1-\omega}{\omega}\sigma^2)^{\frac{\nu'+\gamma+1}{2}} \times \frac{1}{2}\sigma\omega^{-\frac{1}{2}}(1-\omega)^{-\frac{1}{2}}.
\end{aligned}
$$

次にωが与えられた時のβの周辺事後分布は，$(b_\omega, s_\omega^2 H_\omega^{-1})$ をパラメータとするt分布（自由度ν^*）である。ここに，$\nu^*=\nu_1+\nu'+k$,

$\nu^* s_\omega^2=\omega\nu_1 s_1{}^2+(1-\omega)\nu'^2 s'^2+\omega(1-\omega)(b-b_0)'H'H_\omega^{-1}H_0(b-b_0)$

である。

σの同様な事後分布は自由度ν^*のガンマ分布$\Gamma(\omega^{-\frac{1}{2}}s_\omega)$である。さらに$\omega$の事後分布は

$$P(\omega\,|\,x,y)\propto|H_\omega|^{-\frac{1}{2}}\omega^{\frac{1}{2}(\nu_1+k)-1}(1-\omega)^{\frac{1}{2}(\nu'+k)-1}(\nu^* s_\omega{}^2)^{-\frac{1}{2}\nu^*}$$

となる。ωが所与の時の（β, σ^{-2}）は独立でなく，同時事後分布はパラメータが$(b_\omega, \omega^{-1}H_\omega, \nu^*, \omega^{-1}s_\omega^2)$ のk次元正規－ガンマ分布である。

次に事後のモーメントについては次の様になる。

$$\begin{aligned}E(\beta\,|\,x,y)&=E[E(\beta\,|\,\omega,x,y)]=E(b_\omega\,|\,x,y)\\&=E(\omega H_\omega^{-1}\,|\,x,y)H'(b-b_0)+b_0\end{aligned}$$

$$\begin{aligned}V(\beta\,|\,x,y)&=E(V(\beta\,|\,\omega,x,y))+V(b_\omega\,|\,x,y)\\&=V(b_\omega\,|\,x,y)+\frac{1}{\nu^*-2}E[\nu^* s_\omega{}^2 H_\omega^{-1}]\end{aligned}$$

$$E(\sigma^2\,|\,x,y)=\frac{1}{\nu^*-2}E(\nu^* s_\omega{}^2/\omega).$$

2．不完全な事前分布による分析

正規線形回帰モデルにおいて，完全な事前分布を付与できない場合，完全なベイジアン統計分析を行うことはできない。前節では単一の事前分布が付与される場合について考察したが，事前の情報はさまざまな形や程度で存在し，しかも一つの分布に集約できるほどではないということが，実際にはしばしばある。そのような時何をすることができるかを考察することは，有益であろう。

正規線形回帰モデルにおいて自然共役事前分布の場合，事後期待値は

$$b_1=(H+H_0)^{-1}(Hb+H_0b_0)$$

のように書くことができる。ここに $H=\sigma^{-2}x'x, b=(x'x)^{-1}x'y$ である。いま自然共役分布を分析のための事前分布として用いることを決定したとしても，分布のパラメータをどこまで厳密に付与できるかという問題がある。と

くに共分散行列を付与することは容易ではない。そこで，情報が次のような場合について，どのような分析が可能か検討してみよう。

1）b_0, b のみが与えられている

2）b_0, b, x′x が与えられている。

3）b_0, b, x′x と共に H_0 の範囲が与えられている。

1）b_0, b のみが与えられている場合，事後期待値は b_0 と b を結ぶ直線の b から b_0 に至る部分以外の点を除外するすべてのところに存在する可能性がある。それは次の結果から導かれる(Leamer (4), Leamer-Chamberlain (5))。

b_0, b および $B'HB=\rho I$, $B'H_0B=D$ となるような共通共役軸の逆転可能な行列 B が与えられる時（ρ は任意の正数，D は正の対角行列），

$$a_1=B^{-1}b_1=B^{-1}(H+H_0)^{-1}(Hb+H_0b_0)$$

は直方体

$$\left|a_{i1}-\frac{a_i+a_{i0}}{2}\right|<\frac{|a_i-a_{i0}|}{2}$$

の中のすべてのところに存在し得る。ここに $a=B^{-1}b$, $a_0=B^{-1}b_0$。逆にこの限界の中にある b_1 に対しては一意的な対角行列 D が存在し，これに対して b_1 は事後期待値である。

この直方体をもとの座標軸にもとせば，b と b_0 を頂点とし，辺が共通共役軸に平行な平行体となる。この結果を用いると，事後期待値はほとんどすべてのところに存在し得ることを次の様に示すことができる。この限界内にある任意の b_1 を選び，b_1, b_0, b 平面に，b_1 を含み，b_0, b を頂点とする平行四辺形をつくる。この平行四辺形の辺を共通共役軸の２つの軸とし，更に k－2 個の線形独立なベクトルを選んで共役軸の行列をつくれば，b_1 が事後期待値であるような行列Dを求めることができる。b, b_0を通る直線上で，(b, b_0) の部分以外のところの b_1 に対してはこのような平行四辺形をつくることができないから，そのような b_1 は事後期待値ではない。このように加重平均といっても，行列による加重平均は，スカラーウエイトの場合とちがって存在範囲が狭く限定されず，b と b_0 の間にくるというようなことではなく，ほとんどどこへでも動く可能性を持っていることがわかる。

2) これがもっともしばしば通常の例であると思うが，1) の場合に加えて $x'x$ が与えられる場合は，事後期待値はある楕円の内部にあることが示される。それは次の結果から得られる。

b_0 と b が与えられ，更に H_0 と H とが交換可能，すなわち $H_0H=HH_0$ であるならば，事後期待値は，$[b, b_0]$ を直径とする超球の中に存在する。逆にこの超球内の任意の点は，$H_0H=HH_0$ となるある H_0, H に対して事後期待値である。超球は，$c=\frac{b+b_0}{2}$ とすると，

$$(b_1-c)'(b_1-c)<\frac{(b-b_0)'(b-b_0)}{4}.$$

この結果を用いると事後期待値は b と b_0 を結ぶ直線の中点を中心とし，b と b_0 を境界とする標本の楕円，$x'x$ をメトリックとする楕円の内部に存在することがわかる。それは次の通りである。標本楕円を同心球に変換する座標系を求める。すなわち $B'HB=I$ となる B を求める。そこで，

$B^{-1}b_1=(I+A)^{-1}(a+Aa_0)=a_1$

とする。ここに $A=B'H_0B$, $a=B^{-1}b$, $a_0=B^{-1}b_0$ である。

I と A とは交換可能であるから，$g=\frac{a+a_0}{2}$ とすれば上の結果から

$$(a_1-g)'(a_1-g)<\frac{1}{4}(a-a_0)'(a-a_0).$$

これをもとの座標系にもとせば，$B'HB=I$ であるから

$$(B^{-1}b_1-B^{-1}c)'(B^{-1}b_1-B^{-1}c)=(b_1-c)'B^{-1}B'HBB^{-1}(b_1-c)$$
$$=(b_1-c)'H(b_1-c)$$
$$(a-a_0)'(a-a_0)=(B^{-1}b-B^{-1}b_0)'(B^{-1}b-B^{-1}b_0)$$
$$=(b-b_0)'B^{-1}B'HBB^{-1}(b-b_0)=(b-b_0)'H(b-b_0)$$

である。したがって，

$$\left(b_1-\frac{b+b_0}{2}\right)'H\left(b_1-\frac{b+b_0}{2}\right)<\frac{1}{4}(b-b_0)'H(b-b_0).$$

b, b_0 に加えて標本積率を知ると，1) ではほとんど限定されなかった b_1 の領域が1つの楕円にまで限定されることがわかった。

3) 事前共分散行列 H_0^{-1} について，これがある範囲に限定される場合について考察しよう (Polasek (6))。いま V_L を共分散行列の既知の下限として，

$$b_L=(H+(V_L+B)^{-1})^{-1}Hb$$

のような行列加重平均を考えると（b_0 は一般性を失うことなく０とする），$V_L^{-1}=H_L$ とすれば

$$b_L=(H+H_L)^{-1}Hb+(H+H_L)^{-1}H_L^{-1}(H_L(H+H_L)^{-1}H+B^{-1})^{-1}$$
$$\times H_L(H+H_L)Hb$$

のように書くことができるので，これを

$b_L=a+Az$ とおく。$a=(H+H_L)^{-1}Hb$, $A=(H+H_L)^{-1}H_L^{-1}$, $z=(H_L(H+H_L)^{-1}H+B^{-1})^{-1}H_L(H+H_L)Hb$ である。そうすると行列加重平均 z は，B が変化するに従って２）の場合の結果から，次の楕円

$$(z-\frac{1}{2}b)'(V_L+H^{-1})^{-1}(z-\frac{1}{2}b)\leqq c_z$$

の中を動くことがわかる。c_z は定数。したがって b_L に対する楕円は z に対する楕円の線形変換によって求められる。すなわち，

$$(b_L-(a+\frac{1}{2}Ab))'A'^{-1}(V_L+H^{-1})^{-1}A^{-1}(b_L-(a+\frac{1}{2}Ab))\leqq c_z.$$

これを整理すると，

$$(b_L-c_L)'(HV_LH+H)(b_L-c_L)\leqq\frac{1}{4}b'(V_L+H^{-1})^{-1}b$$

となる。ここに $c_L=\frac{1}{2}[b+V_L(V_L+H^{-1})^{-1}b]$ である。

b_0 が０でない場合は，

$$b_L=(H+(V_L+B)^{-1})^{-1}(Hb+(V_L+B)^{-1}b_0)$$

とすれば，

$$c_L=\frac{1}{2}[(b-b_0)+V_L(V_L+H^{-1})^{-1}(b-b_0)]$$

とすると，

$$(b_L-c_L)'(HV_LH+H)(b_L-c_L)\leqq\frac{1}{4}(b-b_0)'(V_L+H^{-1})^{-1}(b-b_0).$$

ここで $H_0=(V_L+B)^{-1}$ とすれば，b_L は事後の期待値である。

次に上限が与えられるとすると，いま

$$b_U=(H+V_U^{-1}+B)^{-1}Hb$$

のような形の行列加重平均を考えると，V_U は共分散行列 V の上限，$V\leqq V_U$ と解釈することができる。ここで $V^{-1}\geqq V_U^{-1}$ である。

$$b_U=(H+V_U^{-1}+B)^{-1}(H+V_U^{-1})(H+V_U^{-1})^{-1}Hb$$

と書けるから

$$b^*=(H+V_U^{-1})^{-1}Hb$$

とおくと b^* は既知の共分散行列の上限 V_U に対する事後期待値と解釈することができる。そこで同様に2）の結果を b_U に対して適用すると，b_U は次の様な楕円の中に存在する。

$$(b_U-c_U)'(V_U^{-1}+H)^{-1}(b_U-c_U)\leqq\frac{1}{4}b'H(V_U^{-1}+H)^{-1}Hb.$$

$H_0=(V_U^{-1}+B)$ とすれば b_U は事後の期待値である。

b_0 が0でない場合は，

$$b_U=(H+V_U^{-1}+B)^{-1}(Hb+(V_U^{-1}+B)b_0)$$
$$=(H+H_0+V_U+B)^{-1}(H(b-b_0)+(H+H+V_U^{-1}+B)b_0)$$
$$=b_0+(H+H+V_U+B)^{-1}H(b-b_0)$$

と書けるので，

$$c_U=b_0+\frac{1}{2}(V_U^{-1}+H)^{-1}H(b-b_0)$$

とすると，楕円

$$(b_U-c_U)'(V_U^{-1}+H)^{-1}(b_U-c_U)\leqq\frac{1}{4}(b-b^*)'H(V_U^{-1}+H)^{-1}H(b-b^*)$$

の中にある。

事前共分散行列 V_0 が $V_L\leqq V_0\leqq V_U$ のように上下から限定されている場合，事後期待値 $b_M=(H+V_0^{-1})^{-1}Hb$ は $V_L^{-1}-V_U^{-1}\geqq B$ なる B について，

$$b_M=(H+V_U^{-1}+B)^{-1}Hb$$
$$=(H+V_U^{-1}+B)^{-1}(H+V_U^{-1})(H+V_U^{-1})^{-1}Hb$$

と書いて，下限が設定されている場合の結果を適用すると，次の楕円の中にあることが導かれる。

$$(b_M-c)'(H+V_U^{-1})[(V_L^{-1}-V_U^{-1})^{-1}(H+V_U^{-1})+I](b^M-c)$$
$$\leqq\frac{1}{4}b'H(H+V_U^{-1})^{-1}(V_L^{-1}-V_U^{-1})(H+V_U^{-1})Hb.$$

ここに，

$$c=\frac{1}{2}(H+V_L^{-1})(H+V_L^{-1}+H+V_U^{-1})(H+V_U^{-1})^{-1}Hb.$$

以上に述べた4つの楕円の関係は，2）の b_1 に関する楕円の内部に b_L と b_U についての楕円が内接し，さらに b_L と b_U の楕円の交わる部分に b_M についての楕円が内接する，ということになる。

3．curve décolletage.

事前分布に関して，等密度面は付与できるが，それがどれだけの数値をとるかは付与できないという場合，何をすることが出来るであろうか。

いまyとβの密度関数が次の様に書けるとしよう。

$$P(y\,|\,x,\beta)\propto f_u[(y-x\beta)'(y-x\beta)],$$

$$P(\beta)\quad\propto\quad f_\beta[(\beta-b_0)'H_0(\beta-b_0)].$$

fは等高楕円に数値を与える関数である。このとき事後分布は次の様になる。

$$P(\beta\,|\,x,y)\propto P(y\,|\,x,\beta)P(\beta)$$

$$\propto f_u[(\beta-b)'x'x(\beta-b)+ESS]f_\beta[(\beta-b_0)'H_0(\beta-b_0)].$$

この事後分布のモードを求めると，

$$\frac{\partial \log f}{\partial\beta}=f_u^{-1}\cdot f_u'2x'x(\beta_m-b)+f_\beta^{-1}f_\beta'2H_0(\beta_m-b_0)=0,$$

$$(f_u^{-1}f_u'x'x+f_\beta^{-1}f_\beta'H_0)\beta_m=f_u^{-1}f_u'x'xb+f_\beta^{-1}f_\beta'H_0b_0.$$

$$\beta_m=(f_u^{-1}f_u'x'x+f_\beta^{-1}f_\beta'H_0)^{-1}(f_u^{-1}f_u'x'xb+f_\beta^{-1}f_\beta'H_0b_0),$$

あるいは、

$$\beta_m=(x'x+\lambda H_0)^{-1}(x'xb+\lambda H_0b_0),\qquad \lambda=\frac{f_\beta^{-1}f_\beta'}{f_u^{-1}f_u'}.$$

λはf_U, f_βに依存するが，0から無限大にまで変化するとβ_mの軌跡は曲線になる。これは，Dicky（2）がcurve décolletage, Leamer（4）がinformation contract curveと呼んでいる曲線であり，一端が最小2乗推定値，他端が事前期待値になっている。

幾何的にはこの曲線は標本密度の等高楕円と事前密度の等高楕円との接点の軌跡である。何故なら，$(\beta-b_0)'H_0(\beta-b_0)$を一定に保ち$(\beta-b)'x'x(\beta-b)$を最大にする問題を解くために，ラグランジュ関数法を用いると，まさに同じ結果が得られるからである。従ってcurve décolletageは，この曲線上にない点a′に対して，曲線上に次のような意味でより良い点aが存在するという性質をもつ。すなわち，aはa′より尤度も事前密度も低くなく，かつ少くともどちらか一方がより高い。

一般に分布が

$$P(x)=f(h(x))$$

のようにに書けるとしよう。f は微分可能で単調減少，かつ $\int f(h(x))dx=1$ をみたすものとする。尤度関数と事前分布が各々 $f_u(h_u(\beta))$, $f_\beta(h_\beta(\beta))$ と書けるならば，事後分布は

$$P(\beta|y)\propto f_u(h_u(\beta))f_\beta(h_\beta(\beta))$$

のように書ける。この分布のモードは次の方程式を満足する。

$$f_u^{-1}f_u'\frac{\partial h_u}{\partial\beta}+f_\beta^{-1}f_\beta'\frac{\partial h_\beta}{\partial\beta}=0$$

$$\frac{\partial h_u}{\partial\beta}+\frac{f_\beta^{-1}f_\beta'}{f_u^{-1}f_u'}\frac{\partial h_\beta}{\partial\beta}=0$$

$\lambda=\frac{f_u f_\beta'}{f_\beta f_u'}$ は β の関数で，$f'_U, f_\beta'<0$ であるから λ は正である。これは一般の場合の curve décolletage を表している。curve décolletage は k 次元空間の中の曲線であり，等密度面を示す関数 h_U, h_β のみに依存している。モードを求める時にはさらに f_U, f_β が必要である。従って情報が等密度面までしか付与できない程度に限られている場合には，ベイジアンは curve décolletage までは析出できることになるが，これはかなりなことである。等密度面を付与することは，β の各点を確からしさの点で順序づけることであり，k 次元の空間を 1 つの曲線にまで限定することができれば，それは分析の段階としてかなり評価できることであろう。

正規分布の場合，

$$f(z^2|\sigma^2)=e^{-\frac{1}{2\sigma^2}z^2}$$

であるから，

$$f^{-1}f'=-\frac{1}{2}\sigma^{-2}$$

となる。事前分布，標本分布が $f_\beta(z^2|\sigma^{*2})$, $f_u(z^2|\sigma^2)$ であれば，$\lambda=\frac{\sigma^{*2}}{\sigma^2}$ である。

次に f がステューデント分布である場合，

$$f_S(z^2|a,\nu)=(a+z^2)^{-\frac{\nu}{2}}$$

であるから，

$$f_S'f_S^{-1}=\frac{\nu}{2(a+z^2)}$$

となる。$f_\beta(z^2)=f_S(z^2|a^*,\nu^*)$, $f_u(z^2)=f_s(z^2|a,\nu)$ の場合，

$$\lambda=\frac{\nu^*}{\nu}\frac{a+ESS+(\beta-b)'x'x(\beta-b)}{a^*+(\beta-b_0)'H_0(\beta-b_0)}$$

である。2節で求めた b_{**} は $\lambda=\frac{1-\omega}{\omega}$ の curve décolletage となっていることがわかる。従って β の事後期待値は curve décolletage 上の点の ω に関する平均であると云うことができる。そしてモードは

$$\lambda=\frac{\nu^{*}(a+ESS+(b_{**}-b)'x'x(b_{**}-b))]}{\nu(a^{*}+(b_{**}-b_{0})'H_{0}(b_{**}-b_{0}))}$$

である時に求められる。

文献

1. Dickey, J.M. (1968) Three Multidimensional -Integral identities with Bayesian Applications. Ann. Math. Stat. 39
2. Dickey, J.M. (1975) Bayesian alternatives to the F －test and least squares estimates in the normal linear model. North－Holland
3. Jeffreys, H. (1961) Theory of probability, (3rd ed.) Oxford Univ. Press.
4. Leamer, E.E. (1978) Specification Searches, John Wiley
5. Leamer, E.E. Chamberlain, G. (1976) Matrix weighted Averages and Posterior Bounds. Journal Roy. Statist. Soc. B 38.
6. Polasek, W. (1984) Multivariate Regression Systems. Elsevier.
7. Raiffa, H. Schlaifer, R. (1961) Applied Statistical Decision Theory. Harvard Univ. Press.
8. Zellner, A. (1971) An Introduction to Bayesian Inference in Econometrics, John Wiley