APPLICATIONES MATHEMATICAE
22,1 (1993), pp. 117–122

M. MĘCZARSKI (Warszawa)

STABILITY AND CONDITIONAL Γ-MINIMAXITY IN BAYESIAN INFERENCE

Abstract. Two concepts of optimality corresponding to Bayesian robust analysis are considered: conditional Γ-minimaxity and stability. Conditions for coincidence of optimal decisions of both kinds are stated.

1. In Bayesian statistical inference arbitrariness of a unique prior distribution is a permanent question. Robust Bayesian inference deals with the problem of expressing uncertainty of the prior information and of quantitative consequences of this uncertainty. A natural measure is width (oscillation, diameter) of the range of a posterior quantity while the prior distribution π runs over a class Γ of probability distributions. If the oscillation of the posterior quantity is small, then the presence of robustness with respect to the prior inexactness can be assured.

A natural goal of research are optimal decisions under a specified loss function and a class Γ of prior distributions, with an idea of optimality related to the robustness problem. The concept of conditional Γ-minimax actions was considered in DasGupta and Studden [3] and Betrò and Ruggeri [1] and it was exhaustively substantiated therein. The idea of stability in Bayesian robust analysis was developed in Męczarski and Zieliński [5], with some additional results in Męczarski [4] and in Boratyńska and Męczarski [2]. The more formal description of the problem is given below.

Let $(\mathcal{X},\mathcal{F},\{P_\theta\}_{\theta\in(\Theta,\mathcal{B})})$ be a statistical space, $\Theta\subset\mathbb{R}$. Let Γ be a class of probability distributions on $(\Theta,\mathcal{B})$, i.e. of prior distributions. It reflects the uncertainty of the prior. Let $x\in\mathbb{R}$ be a given observation and $a\in\mathcal{A}$ a decision (an action) about θ based on x, with the action space $\mathcal{A}\subset\mathbb{R}$ being a compact interval. We consider a loss function $L(\theta,a)$ convex in a

1991 *Mathematics Subject Classification*: 62C10, 62F15, 62F35.
Key words and phrases: robust Bayesian analysis, conditional Γ-minimax action, stability of a statistical procedure.

and the corresponding expected posterior loss (the posterior risk) $\varrho(\pi, a)$ of the action a under the prior π (it depends on x).

The concept of conditional Γ-minimaxity is as follows: construct an action $a^* \in \mathcal{A}$ such that

$$\sup_{\pi\in\Gamma} \varrho(\pi, a^*) = \inf_{a\in\mathcal{A}} \sup_{\pi\in\Gamma} \varrho(\pi, a).$$

Such an action a^* is termed a *conditional Γ-minimax* (CGM) *action.*

The concept of stability is as follows: construct an action $a^{\#} \in \mathcal{A}$ such that

$$\sup_{\pi\in\Gamma} \varrho(\pi, a^{\#}) - \inf_{\pi\in\Gamma} \varrho(\pi, a^{\#}) = \inf_{a\in\mathcal{A}} \{\sup_{\pi\in\Gamma} \varrho(\pi, a) - \inf_{\pi\in\Gamma} \varrho(\pi, a)\},$$

i.e. an action $a^{\#}$ is said to be *stable* if it minimizes the oscillation of $\varrho(\pi, a)$ on Γ with respect to $a \in \mathcal{A}$.

Solutions of some particular estimation problems show that CGM and stable actions may coincide. As seen in Boratyńska and Męczarski [2] a stable solution may lead to large losses of the posterior risk. For that reason such coincidence is desirable and also the stability of CGM actions seems to be a favourable property.

2. The following theorem characterizes CGM actions.

THEOREM 1 (Betrò and Ruggeri [1]). *Assume that $\varrho(\pi, a)$ is a strictly convex function of a for each $\pi\in\Gamma$. Let $\Pi_a = \{\pi^a : \varrho(\pi^a, a) = \sup_{\pi\in\Gamma} \varrho(\pi, a)\}$ be the set of least favourable priors for a decision a. Let a_π^B denote the Bayes action under π. If at $\widehat{a} \in \mathcal{A}$ there exist π_1 and π_2 in $\Pi_{\hat{a}}$ such that $a_{\pi_1}^B \le \widehat{a} \le a_{\pi_2}^B$, then $\widehat{a}$ is a CGM action.* ■

Therefore if the stable action satisfies the conditions of Theorem 1 then it is CGM. The problem is: when is the CGM action stable?

We assume hereafter that $\Gamma = \{\pi_\alpha : \alpha \in [\alpha_1, \alpha_2]\}$.

THEOREM 2. *Let*

$$(\forall a \in \mathcal{A})(\forall \pi_\alpha \in \Gamma) \quad \varrho(\pi_\alpha, a) = r(\alpha, a) = (A\alpha + B - a)^2 + C\alpha + D,$$

with $A \ne 0$, B, C, D real constants. If $\Pi_{\hat{a}} = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$ for an action $\widehat{a} \in \mathcal{A}$, then $\widehat{a}$ is stable; conversely, $\Pi_{a^{\#}} = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$.

Proof. Let $\overline{\alpha} = (\alpha_1 + \alpha_2)/2$ and $\alpha_{\min}(a)$ be the minimum point of the function $r(\cdot, a)$. The elementary geometry of the quadratic curve implies that the oscillation of $r(\cdot, a)$ over $[\alpha_1, \alpha_2]$ is the least iff $\alpha_{\min}(a) = \overline{\alpha}$ or, equivalently, $r(\alpha_1, a) = r(\alpha_2, a)$, which is equivalent to $\Pi_a = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$ because of the form of r. This yields the value $a^{\#}$ which is the unique solution. Since

$$\sup_{\pi\in\Gamma}\varrho(\pi,a)=\sup_{\alpha\in[\alpha_1,\alpha_2]}r(\alpha,a)=\begin{cases}r(\alpha_1,a) & \text{if } \alpha_{\min}(a)\ge\overline{\alpha}\,,\\ r(\alpha_2,a) & \text{if } \alpha_{\min}(a)\le\overline{\alpha}\,,\end{cases}$$

the conditions $\Pi_{\hat{a}}=\{\pi_{\alpha_1},\pi_{\alpha_2}\}$ and $r(\alpha_1,\widehat{a})=r(\alpha_2,\widehat{a})$ are equivalent, which ends the proof. ■

COROLLARY 1. *Assume the conditions of Theorem* 2. *If the CGM action satisfies the conditions of Theorem* 1, *then it is stable. If the stable action is in the interval with endpoints* $a^B_{\pi_{\alpha_1}}$ *and* $a^B_{\pi_{\alpha_2}}$, *then it is CGM.* ■

The following examples show that the situation from Theorem 2 is realistic.

EXAMPLE 1 (Męczarski and Zieliński [5], Betrò and Ruggeri [1]). The problem is to estimate the mean λ in the Poisson distribution $\mathcal{P}(\lambda)$, given an observation x, under the prior gamma $\mathcal{G}(\alpha,\beta)$, $\alpha\in[\alpha_1,\alpha_2]$, β fixed and under the quadratic loss function. Thus

$$\Gamma=\{\mathcal{G}(\alpha,\beta):\alpha\in[\alpha_1,\alpha_2],\beta\text{ is fixed}\}\,,$$

$$\varrho(\pi_\alpha,a)=r(\alpha,a)=\left(\frac{\alpha+x}{\beta+1}-a\right)^2+\frac{\alpha+x}{(\beta+1)^2}\,,$$

$$a^B_{\pi_\alpha}=\frac{\alpha+x}{\beta+1}\,.$$

We have $\Pi_{\hat{a}}=\{\pi_{\alpha_1},\pi_{\alpha_2}\}$ only for $\widehat{a}=(\overline{\alpha}+1/2+x)/(\beta+1)$ and this is a stable action; if $a^B_{\pi_{\alpha_1}}\le\widehat{a}\le a^B_{\pi_{\alpha_2}}$, then $\widehat{a}$ is a CGM action. ■

EXAMPLE 2 (DasGupta and Studden [3], Boratyńska and Męczarski [2]). The problem is to estimate the mean θ in the normal distribution $N(\theta,b^2)$ with known $b>0$, under the normal prior $N(\mu,\sigma^2)$, where $\sigma\in[\sigma_1,\sigma_2]\subset\mathbb{R}_+$ and under the quadratic loss. For $\alpha=(b^{-2}+\sigma^{-2})^{-1}$ we can write $\Gamma=\{\pi_\alpha:\alpha\in[\alpha_1,\alpha_2]\}$. Then

$$\varrho(\pi_\alpha,a)=r(\alpha,a)=\left(\alpha\frac{x-\mu}{b^2}+\mu-a\right)^2+\alpha$$

and

$$a^B_{\pi_\alpha}=\alpha\frac{x-\mu}{b^2}+\mu\,.$$

We have $\Pi_{\hat{a}}=\{\pi_{\alpha_1},\pi_{\alpha_2}\}$ only for

$$\widehat{a}=\overline{\alpha}\frac{x-\mu}{b^2}+\frac{1}{2}\frac{b^2}{x-\mu}$$

and this is a stable action; if $\widehat{a}$ is in the closed interval with endpoints $a^B_{\pi_{\alpha_1}}$ and $a^B_{\pi_{\alpha_2}}$, then $\widehat{a}$ is a CGM action. The last conclusions are valid for $x\ne\mu$;

otherwise any a is stable and $a = \mu$ is CGM; therefore this CGM action is stable as well. ■

3. In Theorem 2 and Examples 1 and 2 the class Γ is defined by a real parameter α from a compact interval. The posterior risk is a quadratic function of α. By minimizing the oscillation of that function we arrive at the condition $\varrho(\pi_{\alpha_1}, a) = \varrho(\pi_{\alpha_2}, a)$, which defines the stable action. From the shape of the function we obtain the conditions of Theorem 1 and the stability of the CGM action.

Let us consider the following example.

Example 3 (Męczarski [4]). The problem is to estimate the parameter $\theta = e^{-\lambda}$ in the Poisson distribution $\mathcal{P}(\lambda)$. The class of priors and the loss function are as in Example 1. Then

$$\varrho(\pi_\alpha, a) = r(\alpha, a) = (e^{\beta_1(\alpha+x)} - a)^2 - e^{2\beta_1(\alpha+x)} + e^{\beta_2(\alpha+x)},$$

$$\beta_i = \log\frac{\beta+1}{\beta+1+i}, \quad i = 1, 2 \quad \text{and} \quad a^B_{\pi_\alpha} = e^{\beta_1(\alpha+x)}.$$

By looking for $\inf_{a\in\mathcal{A}} \sup_{\pi\in\Gamma} \varrho(\pi, a)$ we obtain $\Pi_{\hat{a}} = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$ only for the $\widehat{a}$ which is the unique solution of $r(\alpha_1, \widehat{a}) = r(\alpha_2, \widehat{a})$. If $a^B_{\pi_{\alpha_2}} \le \widehat{a} \le a^B_{\pi_{\alpha_1}}$, then $\widehat{a}$ is a CGM action. By investigating the monotonicity of the oscillation of $r(\cdot, a)$ over $[\alpha_1, \alpha_2]$ we conclude that $a^{\#}$ is its unique minimum point (a stable action) if and only if it satisfies $r(\alpha_1, a^{\#}) = r(\alpha_2, a^{\#})$. ■

Example 3 shows that the function $r(\cdot, a)$ need not be quadratic to obtain the considered connection between stability and conditional Γ-minimaxity. As before, the condition $r(\alpha_1, a) = r(\alpha_2, a)$ is a tool to construct an action with both properties.

Theorem 3. *Define a function r by $r(\alpha, a) = \varrho(\pi_\alpha, a)$. Assume that it satisfies the following conditions*:

(a) $r(\alpha, \cdot)$ *is strictly convex for any* α;

(b) *for any a the minimum point $\alpha_{\min}(a)$ of $r(\cdot, a)$ is unique and $\alpha_{\min}$ is a strictly monotone function of a*;

(c) *for any $\widetilde{\alpha}$ and $\widetilde{a}$ such that $\alpha_{\min}(\widetilde{a}) = \widetilde{\alpha}$ we have*

$$(\forall a' < a'' \le \widetilde{a}) \quad \frac{r(\widetilde{\alpha}, a'') - r(\widetilde{\alpha}, a')}{a'' - a'} < \frac{r(\alpha_{\min}(a''), a'') - r(\alpha_{\min}(a'), a')}{a'' - a'}$$

and

$$(\forall a'' > a' \ge \widetilde{a}) \quad \frac{r(\widetilde{\alpha}, a'') - r(\widetilde{\alpha}, a')}{a'' - a'} > \frac{r(\alpha_{\min}(a''), a'') - r(\alpha_{\min}(a'), a')}{a'' - a'};$$

(d) *the function $r(\alpha_1, a) - r(\alpha_2, a)$ is monotone in a.*

Then the conclusion of Theorem 2 holds.

P r o o f. Denote the oscillation of $r(\cdot, a)$ over $[\alpha_1, \alpha_2]$ by $Q(a)$. Then

$$Q(a) = \begin{cases} r(\alpha_2, a) - r(\alpha_1, a) & \text{if } \alpha_{\min}(a) \le \alpha_1, \\ r(\alpha_2, a) - r(\alpha_{\min}(a), a) & \text{if } \alpha_1 < \alpha_{\min}(a) < \alpha_2 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \le 0, \\ r(\alpha_1, a) - r(\alpha_{\min}(a), a) & \text{if } \alpha_1 < \alpha_{\min}(a) < \alpha_2 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \ge 0, \\ r(\alpha_1, a) - r(\alpha_2, a) & \text{if } \alpha_{\min}(a) \ge \alpha_2. \end{cases}$$

Let a_i be a solution of $\alpha_{\min}(a) = \alpha_i$, $i = 1, 2$. The type of monotonicity of the function $r(\alpha_1, a) - r(\alpha_2, a)$ agrees with that of $\alpha_{\min}$. Then for $\alpha_{\min}$ increasing we have

$$Q(a) = \begin{cases} r(\alpha_2, a) - r(\alpha_1, a) & \text{if } a \le a_1, \\ r(\alpha_2, a) - r(\alpha_{\min}(a), a) & \text{if } a_1 < a < a_2 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \le 0, \\ r(\alpha_1, a) - r(\alpha_{\min}(a), a) & \text{if } a_1 < a < a_2 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \ge 0, \\ r(\alpha_1, a) - r(\alpha_2, a) & \text{if } a \ge a_2 \end{cases}$$

and for $\alpha_{\min}$ decreasing we have

$$Q(a) = \begin{cases} r(\alpha_1, a) - r(\alpha_2, a) & \text{if } a \le a_2, \\ r(\alpha_1, a) - r(\alpha_{\min}(a), a) & \text{if } a_2 < a < a_1 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \ge 0, \\ r(\alpha_2, a) - r(\alpha_{\min}(a), a) & \text{if } a_2 < a < a_1 \\ & \text{and } r(\alpha_1, a) - r(\alpha_2, a) \le 0, \\ r(\alpha_2, a) - r(\alpha_1, a) & \text{if } a \ge a_1. \end{cases}$$

Observe that the functions $r(\alpha_i, a) - r(\alpha_{\min}(a), a)$ have their unique minimum points at a_i and they are decreasing for $a < a_i$ and increasing for $a > a_i$, $i = 1, 2$, respectively. This implies that Q has its unique minimum point at $a^{\#}$ which is defined by the condition $r(\alpha_1, a^{\#}) - r(\alpha_2, a^{\#}) = 0$.

Now,

$$\sup_{\alpha \in [\alpha_1, \alpha_2]} r(\alpha, a) = \begin{cases} r(\alpha_1, a) & \text{if } r(\alpha_1, a) \ge r(\alpha_2, a), \\ r(\alpha_2, a) & \text{if } r(\alpha_1, a) \le r(\alpha_2, a) \end{cases}$$

and therefore $\Pi_{a^{\#}} = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$. Conversely, if $\Pi_{\hat{a}} = \{\pi_{\alpha_1}, \pi_{\alpha_2}\}$, then $\widehat{a} = a^{\#}$. This ends the proof. ∎

COROLLARY 2. *Assume the conditions of Theorem* 3. *Let the function* $f(\alpha) = a^B_{\pi_\alpha}$ *be strictly monotone and* $f([\alpha_1, \alpha_2]) = \mathcal{A}$. *Then the conclusion of Corollary* 1 *holds.* ∎

Now we can see that the conditions of Theorem 3 and Corollary 2 hold in Examples 1, 2 and 3.

References

[1] B. Betrò and F. Ruggeri, *Conditional Γ-minimax actions under convex losses*, Comm. Statist. Theory Methods 21 (1992), 1051–1066.
[2] A. Boratyńska and M. Męczarski, *Robust Bayesian estimation in the one-dimensional normal model*, submitted.
[3] A. DasGupta and W. J. Studden, *Frequentist behavior of robust Bayes estimates of normal means*, Statist. Decisions 7 (1989), 333–361.
[4] M. Męczarski, *On stable Bayesian estimation in the Poisson model*, Sci. Bull. Łódź Technical Univ. No. 687, Matematyka, z. 25 (1993), 83–91.
[5] M. Męczarski and R. Zieliński, *Stability of the Bayesian estimator of the Poisson mean under the inexactly specified gamma prior*, Statist. Probab. Lett. 12 (1991), 329–333.

MAREK MĘCZARSKI
INSTITUTE OF ECONOMETRICS
WARSAW SCHOOL OF ECONOMICS
AL. NIEPODLEGŁOŚCI 162
02-554 WARSZAWA, POLAND

Received on 25.3.1993