Article ID: 1003- 6326(2000)01- 0114- 05

Admissibility of Bayes estimate with inaccurate prior in surveying adjustment ①

ZHU Jian-jun(朱建军)[1], WANG Xi-bo(王细波)[2]

1. Department of Surveying Engineering, Central South University of Technology, Changsha, 410083, P. R. China;

2. The Land Management Bureau of Huizhou, Huizhou 516000, P. R. China

Abstract: Based on the concept of admissibility in statistics, a definition of generalized admissibility of Bayes estimates has been given at first, which was with inaccurate prior for the application in surveying adjustment. Then according to the definition, the generalized admissibility of the normal linear Bayes estimate with the inaccurate prior information that contains deviations or model errors, as well as how to eliminate the effect of the model error on the Bayes estimate in surveying adjustment were studied. The results show that if the prior information is not accurate, that is, it contains model error, the generalized admissibility can explain whether the Bayes estimate can be accepted or not. For the case of linear normal Bayes estimate, the Bayes estimate can be made generally admissible by giving a less prior weight if the prior information is inaccurate. Finally an example was given.

Key words: Bayes estimate; admissibility; inaccurate prior; surveying adjustment **Document code: A**

1 INTRODUCTION

Bayes estimate has been widely used in surveying data processing[1~3]. The advantage of Bayes estimate is that it can make use of not only the information contained in the observations, but also the information from the historical data or from other areas. This means that the Bayes estimate can make comprehensive use of the knowledge in different areas, which is very important in modern science and technology[4, 5].

But there also exists a very important default in Bayes estimate in practical application, that is, in many situations the prior information is subjective. The subjectively given prior information will be vague, and will unavoidably contain deviations. These deviations will certainly affect the results of surveying adjustment. In this case, one probably questions that can the subjectively determined prior information actually improve the results? If the prior information is not reliable, that is, the prior information contains deviations or model errors, do the results of the Bayes estimate still have practical sense? In other words, are the results of Bayes estimate better than that of the estimate that does not use any prior information.

Berger in 1980 and 1984 has studied the robustness of Bayes estimate when the prior information contains deviations or model errors, his work was to restrict the deviations of prior information or prior distributions[6]. But in many cases in surveying adjustment, one usually does not know the exact deviation of the prior information. The deviation or the model errors in prior information usually can not be avoided. In this paper we will deal with the Bayes estimate in another way. Our research will focus on that in which conditions the estimate that has make use of the inaccurate prior information will be better than the estimate that does not use any prior information, and how to determine the prior parameters if the prior information has model errors or deviations.

If a Bayes estimate is better than an estimate that does not use any prior information, this Bayes estimate can be accepted, we can say this Bayes estimate is admissible.

At first the paper gives out the definition of the admissibility. Based on the definition, the paper then studies the admissibility of the normal linear Bayes estimate with the inaccurate prior information, and studies how to eliminate the affection of the model error on the Bayes estimate in surveying adjustment. Finally an example is given.

2 DEFINITION OF ADMISSIBILITY OF BAYES ESTIMATE IN SURVEYING ADJUSTMENT

2.1 Concept of classical admissibility in statistics

In surveying adjustment, the observation equations can be denoted as

$$L + V = Ax \quad (1)$$

where x denotes the unknown vector, L denotes the observations, V the residuals vector, A the coef-

① **Foundation item:** Project 49774209 supported by the National Natural Science Foundation of China

Received date: Feb. 25, 1999; **accepted date:** Oct. 12, 1999

ficient matrix. From Eqn. (1), one can get an estimate of $\boldsymbol{x}$:

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^{-1}\boldsymbol{L} \tag{2}$$

where $\boldsymbol{A}^{-1}$ denotes a generalized inverse of $\boldsymbol{A}$. In statistics, the process finding the estimate $\hat{\boldsymbol{x}}$ from Eqn. (2) is called an action. How to get the estimate usually depends on loss function. The well-known and widely used loss function is quadratic function[7]:

$$L(\boldsymbol{x}, \hat{\boldsymbol{x}}) = (\boldsymbol{x}-\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}} \tag{3}$$

Under the quadratic loss, one can get the estimate from Eqn. (1)

$$\hat{\boldsymbol{x}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{L} \tag{4}$$

Eqn. (4) is actually the least square estimate. Different observations $\boldsymbol{L}$ will bring out different $\hat{\boldsymbol{x}}$, so the values of the loss function will be different. The average loss over different values of $\hat{\boldsymbol{x}}$ is defined as the risk function:

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}) = E_{\theta}(L(\boldsymbol{x}, \hat{\boldsymbol{x}})) = \int L(\boldsymbol{x}, \hat{\boldsymbol{x}})\,\mathrm{d}P_{\theta}(\boldsymbol{L}/\theta) \tag{5}$$

For the Bayes estimate, the posterior expected loss is defined as the risk function:

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) = \int L(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})P(\boldsymbol{x}/\boldsymbol{L})\,\mathrm{d}\boldsymbol{x} \tag{6}$$

For two estimates $\hat{\boldsymbol{x}}_1$ and $\hat{\boldsymbol{x}}_2$, if there exist:

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}_1) \leqslant R(\boldsymbol{x}, \hat{\boldsymbol{x}}_2)$$

we say the estimate $\hat{\boldsymbol{x}}_1$ is R-better than $\hat{\boldsymbol{x}}_2$[8]. The classical admissibility of an estimate is defined as following:

An estimate is admissible if there exists no R-better estimate.

Under the quadratic loss function the risk function of estimate (Eqn. (4)), that is the least square estimate, is:

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}) = E((\boldsymbol{x}-\hat{\boldsymbol{x}})(\boldsymbol{x}-\hat{\boldsymbol{x}})^{\mathrm{T}}) = D(\boldsymbol{x}) \tag{7}$$

This means that under quadratic loss, the risk of the least square estimate is the variance of the estimate. The classical admissibility means that the least square estimate has least variance. Under the quadratic loss, the Bayes estimate of model (Eqn. (1)) is

$$\hat{\boldsymbol{x}}_{\mathrm{B}} = E(\boldsymbol{x}/\boldsymbol{L}) \tag{8}$$

the corresponding risk is

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) = E((\boldsymbol{x}-\hat{\boldsymbol{x}}_{\mathrm{B}})(\boldsymbol{x}-\hat{\boldsymbol{x}}_{\mathrm{B}})^{\mathrm{T}}) = D((\boldsymbol{x}/\boldsymbol{L}) < D(\boldsymbol{x}) \tag{9}$$

It means that the risk of the Bayes estimate under the quadratic loss is its posterior variance.

From above, we know that the classical admissibility of an estimate is based on the risk function. However, the loss function and the risk function are derived from an assumed model. If the assumed model is based on the practical situations, that is, the difference between the assumed model and the real situation is small, the above admissibility is reasonable and feasible. However, if the prior information is not reliable, or is determined subjectively, the assumed model perhaps deviated largely from the practical situation. In this condition, the classical admissibility is senseless for Bayes estimate. This also can be explained mathematically by the following:

Let $P_0(\boldsymbol{x})$ be the real prior distribution, $P(\boldsymbol{x})$ the assumpted prior distribution, then the classical risk will be

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) = \int L(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})P(\boldsymbol{x}/\boldsymbol{L})\,\mathrm{d}\boldsymbol{x} = \int L(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})\frac{P(\boldsymbol{L}/\boldsymbol{x})P(\boldsymbol{x})}{f(\boldsymbol{L})}\,\mathrm{d}\boldsymbol{x} \tag{10}$$

But the real risk should be

$$R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) = \int L(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})\frac{P(\boldsymbol{L}/\boldsymbol{x})P_0(\boldsymbol{x})}{f(\boldsymbol{L})}\,\mathrm{d}\boldsymbol{x} \tag{11}$$

If the prior information is not reliable or $P(\boldsymbol{x})$ is determined subjectively, it is possible that $P(\boldsymbol{x})$ contain a large model error. This model error certainly brings about a large difference between the risk function $R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$ and $R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$. This means that $R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$ can not reflect the real loss, the admissibility on $R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$ will be senseless.

2.2 Definition of generalized admissibility of Bayes estimate

In the Bayes estimate, if the prior information is subjective, one perhaps concerns mainly with whether the result of Bayes estimate is reliable. In other words, one perhaps want to know whether the results that have make use of the subjectively determined prior information will be better than the result that does not use any prior information. According to the concept of the classical admissibility, we can still use the value of the risk function to decide which estimate is better. In order to show a Bayes estimate is valuable or not, and to show the subjectively determined prior information has played a valuable role in the estimate or not, we defined an another kind of admissibility for Bayes estimate in surveying adjustment as that the Bayes estimate is generalized admissible if

$$R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) \leqslant R(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}}) \tag{12}$$

where $R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$ is found by Eqn. (11), and is determined by the real prior distribution. The value of $R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_{\mathrm{B}})$ is not affected by the model error, so it accurately reflects the expected loss of $\boldsymbol{x}_{\mathrm{B}}$. $R_0(\boldsymbol{x}, \hat{\boldsymbol{x}})$ is the risk of the estimate that do not use any prior information, but use the same observations as the Bayes estimate.

This definition means that a Bayes estimate may be determined by a kind of unreliable prior information (that is, the prior information contain model errors), only if the risk of the estimate is still smaller than that does not use any prior information. In this situation, the prior information (although it contains model error) has made the risk of the estimate smaller and played a useful role in the process of estimating.

The above definition is different from the classical admissibility. Classical admissibility of an estimate means the estimate has the least risk. Under quadratic loss, classical admissibility of an estimate means the estimate has the least variance. But above definition emphases on whether the model-error-contained prior information has played a useful role in the estimate or not. The admissible estimate only means the prior information has made the risk or the variance smaller, and the estimate can be accepted.

In order to avoid confusing with the classical admissibility in statistics, the above definition is tentatively called as generalized admissibility.

In practical computation, $P_0(\boldsymbol{x})$ is usually unknown, so the risk $R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_B)$ also is unknown. However, in many situations, the model error usually can be valued according to the practical condition, that is, the difference between $P_0(\boldsymbol{x})$ and $P(\boldsymbol{x})$ usually can be valued, so the generalized admissibility can be studied on the valued model errors.

3 ADMISSIBILITY OF LINEAR NORMAL BAYES ESTIMATE

Assume:

$$\boldsymbol{l} = \boldsymbol{Ax} + \varepsilon$$

$$\varepsilon \sim N(0, \sigma_1^2 \boldsymbol{P}_0) \tag{13}$$

where $\boldsymbol{l}$ denotes the observation, $\boldsymbol{x}$ the vector of unknown, $\boldsymbol{A}$ the coefficient matrix, ε the error vector. The least square estimate of above model is:

$$\hat{\boldsymbol{x}} = (\boldsymbol{APA})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PL}$$

The corresponding loss function will be:

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \sigma^2(\boldsymbol{APA})^{-1}$$

If we have prior information

$$\boldsymbol{x} \sim N(\boldsymbol{l}_x, \sigma^2\boldsymbol{P}_x^{-1})$$

under the quadratic loss, the Bayes estimate of Eqn. (13) will be

$$\hat{\boldsymbol{x}}_B = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA} + \boldsymbol{P}_x)^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Pl} + \boldsymbol{P}_x\boldsymbol{l}_x) \tag{14}$$

The corresponding loss will be

$$R(\boldsymbol{x}, \hat{\boldsymbol{x}}_B) = E((\boldsymbol{x} - \hat{\boldsymbol{x}}_B)(\boldsymbol{x} - \hat{\boldsymbol{x}}_B)^{\mathrm{T}})$$

For the normal prior, the model error of the prior information depends on the deviation of the prior parameters $\boldsymbol{l}_x$ and $\boldsymbol{P}_x$. For convenience of analysis, without loss generosity, we assume:

$$K\boldsymbol{I} = \boldsymbol{P}_{\hat{x}}^{-1}\boldsymbol{P}_{x_0} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA})^{-1}\boldsymbol{P}_{x_0} \tag{15}$$

$$\boldsymbol{P}_x - \boldsymbol{P}_{x_0} = \lambda_1\boldsymbol{P}_{x_0} \tag{16}$$

$$\boldsymbol{P}_{x_0} = K\boldsymbol{P}_{\hat{x}} = K(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA})$$

$$\Delta\boldsymbol{l}_x\Delta\boldsymbol{l}_x^{\mathrm{T}} = \lambda_2\boldsymbol{D}_{\hat{x}} \tag{17}$$

where $\hat{\boldsymbol{x}}$ is the least square estimate of model (Eqn. (13)), $\boldsymbol{D}_{\hat{x}}$ is the variance of $\hat{\boldsymbol{x}}$, $\boldsymbol{P}_{\hat{x}} = \boldsymbol{D}_{\hat{x}}^{-1}$ be the weight matrix of $\hat{\boldsymbol{x}}$; $\boldsymbol{P}_{x_0}$ and $\boldsymbol{l}_{x_0}$ denote the real prior parameters, $\boldsymbol{P}_x$ and $\boldsymbol{l}_x$ the prior parameters containing deviations, $\Delta\boldsymbol{l}_x$ denotes the deviations of $\boldsymbol{l}_x$. K actually denotes the ratio of the prior information to the information contained in observations, λ_1 denotes the ratio of the deviation of the prior information to real prior information, and λ_2 denotes the deviation of prior parameters $\boldsymbol{l}_x$ to the variance of the least square estimate from the observation. The ratio λ_2 can explain how large the deviation of the prior parameter be, so the model error of the prior information can be determined by λ_1 and λ_2.

Enter Eqn. (14) into Eqn. (11), we have

$$R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_B) = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA} + \boldsymbol{P}_x)^{-1}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA} + \boldsymbol{P}_x(\boldsymbol{P}_x^{-1} + \Delta\boldsymbol{x}\Delta\boldsymbol{x}^{\mathrm{T}})\boldsymbol{P}_x) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA} + \boldsymbol{P}_x)^{-1} \tag{18}$$

Considering Eqns. (15), (16) and (17), we have

$$R_0(\boldsymbol{x}, \hat{\boldsymbol{x}}_B) = \frac{1}{(1+(1+\lambda_1)K)^2} \cdot (1+(1+\lambda_1)^2K^2(\frac{1}{K}+\lambda_2) \cdot (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{PA})^{-1} = \frac{1+(1+\lambda_1)^2K(1+\lambda_2K)}{(1+(1+\lambda_1)K)^2} \cdot R(\boldsymbol{x}, \hat{\boldsymbol{x}}) \tag{19}$$

The admissibility of the Bayes estimate of model (Eqn. (13)) is therefore discussed according to three separate situations:

1) $\lambda_2 = 0$, $\lambda_1 \neq 0$, that is, the prior parameter $\boldsymbol{l}_x$ is accurate, but the weight $\boldsymbol{P}_x$ contains deviations. In this situation, if Eqn. (12) is tenable, we will have

$$\frac{1+(1+\lambda_1)^2K}{(1+(1+\lambda_1)K)^2} \leqslant 1 \tag{20}$$

After rearranging, we have

$$\frac{1+K}{1-K} > \lambda_1 > -1 \quad (K \leqslant 1) \tag{21}$$

$$\lambda_1 > \max\left\{\frac{1+K}{1-K}, -1\right\} \quad (K > 1)$$

$$\lambda_1 > -1 \quad (K > 1) \tag{22}$$

that is, if prior parameter $\boldsymbol{l}_x$ is accurate and the deviation of prior weight $\boldsymbol{P}_x$ can make Eqn. (21) or Eqn. (22) tenable, the corresponding Bayes estimate is admissible.

Eqns. (21) and (22) show that:

a. When $0 > \lambda_1 > -1$, we have $\boldsymbol{P}_x < \boldsymbol{P}_{x_0}$, this means that the used prior weight is less than the real prior weight. This case is corresponding to that the observations with higher accuracy are used as lower accurate observations. In this case, no unreasonable results will be brought out, except the observations with higher accuracy do not play completely the role in the adjustment.

b. When $K > 1$, the worst case will be $\lambda_1 = \infty$, that is, the deviation of prior parameter $P(\boldsymbol{x})$ is infinite, the results will be $\hat{\boldsymbol{x}}_B = \boldsymbol{l}_{x_0}$. This case is corresponding to the situation that only prior information is used, the observations are neglected. But $K > 1$ means $\boldsymbol{P}_{x_0} > \boldsymbol{P}_{\hat{x}}$, that is, $\boldsymbol{D}_{\hat{x}_B} = \boldsymbol{P}_{x_0}^{-1} < \boldsymbol{D}_{\hat{x}} = \boldsymbol{P}_{\hat{x}}^{-1}$, thus, the Bayes estimate is still better than the least square estimate.

c. When $K \leqslant 1$, the deviation of prior weight will be limited by Eqn. (21). For different values of K, the upper limitations can be calculated, the results are listed in Table 1.

Table 1 Upper limitation of deviation parameter λ_1

K	0.01	0.1	0.3	0.5	0.6
λ_1	1.02	1.22	1.86	3.00	4.00
K	0.7	0.8	0.9	0.95	0.99
λ_1	5.67	9.00	19.0	39.0	199

Apparently, only if $\boldsymbol{l}_x$ is accurate, that is, does not contain deviations, the limitation on the deviation of $\boldsymbol{P}_x$ is very large. Usually, the given value of the prior weight $\boldsymbol{P}_x$ is not too large, for example, $\boldsymbol{P}_x < \boldsymbol{P}_{x_0}$, the Bayes estimate always be admissible, and always can be accepted.

2) $\lambda_1 = 0$, $\lambda_2 \neq 0$, that is, the prior value $\boldsymbol{l}_x$ of $\boldsymbol{x}$ contains deviations, but prior weight $\boldsymbol{P}_x$ is accurate.

From Eqn. (19), we have

$$\frac{1+K(1+\lambda_2 K)}{(1+K)^2} < 1$$

$$1+K(1+\lambda_2 K) < (1+K)^2$$

$$1+\lambda_2 K < 2+K$$

$$\lambda_2 < \frac{1+K}{K} \tag{23}$$

that is, providing prior weight $\boldsymbol{P}_x$ is accurate, the deviations of prior value $\boldsymbol{l}_x$ of $\boldsymbol{x}$ can make Eqn. (24) tenable, the Bayes estimate is admissible.

From Eqn. (24), we know that when K take its value in the area of (0~ 1), $(1+K)/K$ will get its value in the area of (∞, 2). This means when the prior information does not play a leader role in the Bayes estimate, a large deviation of prior value of $\boldsymbol{x}$ is allowed.

3) $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, both prior parameters $\boldsymbol{l}_x$ and $\boldsymbol{P}_x$ contain deviations, from Eqn. (18), we have

$$\frac{1+(1+\lambda_1)^2 K(1+\lambda_2 K)}{(1+(1+\lambda_1)K)^2} < 1 \tag{25}$$

$$1+(1+\lambda_1)^2 K(1+\lambda_2 K) < (1+(1+\lambda_1)K)^2$$

$$(1+\lambda_1)^2 K(1+\lambda_2 K) < 2(1+\lambda_1)K + (1+\lambda_1)^2 K^2$$

$$(1+\lambda_1)K(1+\lambda_1+\lambda_2(1+\lambda_1)K - 2 - (1+\lambda_1)K) < 0$$

$$1+\lambda_1 \geqslant 0$$

$$\lambda_1+\lambda_2(1+\lambda_1)K < 1(1+\lambda_1)K$$

$$\lambda_2(1+\lambda_1)K < 1-\lambda_1+(1+\lambda_1)K$$

$$\lambda_2 < \frac{1-\lambda_1(1+\lambda_1)K}{(1+\lambda_1)K} \tag{26}$$

$$\lambda_2 < 1+\frac{1-\lambda_1}{(1+\lambda_1)K} \tag{27}$$

For different values of K and λ_1, the upper limitations of λ_2 are calculated, the results are listed in Table 2.

Table 2 Upper limitation of deviation parameters λ_2

λ_1	K_2						
	0.01	0.05	0.1	0.3	0.5	0.7	1.0
−0.9	1901	381	191	64.33	39	28.14	20
−0.5	301	61	31	11	7	5.286	4
−0.1	123.2	25.44	13.22	5.074	3.444	2.746	2.222
0	101	21	11	4.333	3	2.429	2
0.1	82.82	17.36	9.182	3.727	2.636	2.169	1.818
0.5	34.33	7.667	4.333	2.111	1.667	1.476	1.333

From Table 2, following can be concluded.

a. If K is very small and $\lambda_1 < 1$, especially $\lambda_1 < 0$, a large deviation in prior value $\boldsymbol{l}_x$ of $\boldsymbol{x}$ is allowed. This conclusion is very important and useful, for example, in many ill-solutions, there usually exists little prior information. It is very difficult to determine the prior parameters. But from above conclusion, we know that if the value of prior weight is given very small, a large deviation of prior value $\boldsymbol{l}_x$ of $\boldsymbol{x}$ is allowed. So we can give a prior value $\boldsymbol{l}_x = 0$, and a very litter prior weight $\boldsymbol{P}_x$ in the ill-solution situations.

b. For a given K, the upper limitation of λ_2 will increase when λ_1 decease. This means that when it is difficult to determine the prior parameters or the prior parameters contain deviations, the prior weight should be given a less value.

4 EXAMPLE

The following is a problem of multicollinearities which is taken from Fang[9]:

$$\boldsymbol{L}+\boldsymbol{V}=\boldsymbol{A}\boldsymbol{x}$$

where

$$\boldsymbol{A}^{\mathrm{T}}=\begin{bmatrix} 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 & 10 \\ 1.1 & 1.4 & 1.7 & 1.7 & 1.8 & 1.9 & 2.0 & 2.3 & 2.4 \\ 1.1 & 1.5 & 1.8 & 1.7 & 1.9 & 1.8 & 1.8 & 2.1 & 2.5 \end{bmatrix}$$

$$\boldsymbol{L}^{\mathrm{T}} = (16.3,\ 16.8,\ 19.2,\ 8.0,\ 19.5,\ 20.9,\ 21.1,\ 20.9,\ 20.3,\ 22.0)$$

$$\boldsymbol{D}_L = \boldsymbol{I}$$

The real values of $\boldsymbol{x}$ is $\boldsymbol{x}^{\mathrm{T}} = (1, 2, 3)$. The least square estimate of above model is

$$\boldsymbol{P}_x = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{A} = \begin{bmatrix} 1000 & 181 & 186 \\ 181 & 34.09 & 35.06 \\ 186 & 35.09 & 36.1 \end{bmatrix},$$

while

$$\boldsymbol{x} = \begin{bmatrix} 1.129 \\ 11.307 \\ -6.591 \end{bmatrix},$$

$$D_x = \begin{bmatrix} 0.026 & -0.233 & 0.092 \\ 0.092 & 27 & -25.08 \\ 0.092 & -25.08 & 23.9 \end{bmatrix}$$

In this situation, no any prior information can be used. On the principle "small as possible until it can be ignored", we can take 1% of the information contained in observations as its prior information, that is, let $P_{x_0} = 0.01P_x$. Because the prior information determined in this way is far less than that contained in observations. Its role in the adjustment can be neglected in usual situations. In practical process, 1% of the least component in the diagonal of $P_x = A^T PA$ can be taken. In this case, we take 34.09 × 1% = 0.34 as the prior, that is, let

$$P_{x_0} = 0.34I \tag{28}$$

Let the prior value of x be $l_x^T = (0, 0, 0,)$, comparing with the result of the least square estimate, we can know that there probably exists a deviation $\Delta l_{x_1} = x_1 - l_{x_1} = 1.129$ in l_x, which can make $\lambda_2 = \Delta l_{x_1}^2 / D_{x_1} = 1.129^2/0.026 = 49$. $\lambda_2 = 49$ means the deviation is very large. But if the prior weight determined by Eqn. (28) is thought to be reasonable, that is, $\lambda_1 = 0$ is assumed, enter $K = 0.01$ and $\lambda_1 = 0$ into Eqn. (26), one can get the limitation of the deviations is $101 > \lambda_2 = 49$, so in this situation the Bayes estimate is admissible and accepted.

Even if take $\lambda_1 = 0.3$, then

$$P_x = P_{x_0} + 0.3P_{x_0} = \begin{bmatrix} 0.47 & 0 & 0 \\ 0 & 0.47 & 0 \\ 0 & 0 & 0.47 \end{bmatrix}$$

one still can get the limitation of deviation be $53 > \lambda_2 = 49$, that is, the Bayes estimate still be admissible and can be accepted.

Take the prior parameters $l_x = 0$, $P_x = 0.36I$ and $l_x = 0$, $P_x = 0.47I$, the corresponding Bayes estimates are calculated, the results are listed in Table 3.

Table 3 Comparison of different results

Item	x	L_s	R_g	B_s- 1	B_s- 2
x_0	1	1.1292	1.2869	1.2612	1.2816
x_1	2	11.307	2.0992	2.2956	2.1345
x_2	3	- 6.591	1.5186	1.4670	1.5129

Note: x denote the real value of the unknown, L_s denotes the least square estimate, R_g the result of ridge estimate (the ridge trace method), B_s- 1 the result of Bayes estimate ($K = 0.36$), B_s- 2 the result of Bayes estimate ($K = 0.47$).

From Table 3, one can know:

1) The results of Bayes estimates are better than that of the classical least squares, and are close to the result of ridge estimate. This shows the conclusions in the text, that is, "when K is very small, $\lambda_1 < 1$, a large deviation in prior value l_x of x is allowed". In this case, the deviations is $\lambda_2 = 49$.

2) Even if no prior information can be used, the prior parameters can still be found on the limitations of Eqn. (27), and Bayes approach can still be used in the problem of multicollinearities, the result will be close to the ridge trace method, but the computation process will be simpler.

5 CONCLUSIONS

1) If the prior information is not accurate, that is, contains model error, the generalized admissibility can be used to show whether the Bayes estimate can be accepted or not.

2) For the linear normal Bayes estimate, if the prior parameter l_x is accurate ($\lambda_2 = 0$), the Bayes estimate is admissible in most situations. Even in the situation $K \leqslant 1$, the limitations on the deviations of prior weight P_x still are very large.

3) If the prior parameters P_x are accurate, the limitation on deviation of prior value l_x depends on $(1 + K)/K$, when K take its value in the area (0~ 1), $(1 + K)/K$ will get its value in the area (∞, 2). This means when the prior information does not play a leader role in the Bayes estimate, a large deviation of prior value of x is allowed.

4) In the situations that both parameters l_x and P_x contain deviations, if the ratio of the prior information to the information contained in observations is small, a large deviation in prior value l_x of x is still allowed. If it is difficult to determine the prior parameters, the prior weight is given a less value usually can make Bayes estimate admissible.

REFERENCES

[1] Koch K R. Bayesian Inference with Geodetic Applications [M]. New York: Springer Verlag New York Inc, 1990.

[2] ZHU Jian-jun. Sensitivity and separability of deformation models with regard to prior information [J]. Trans Nonferrous Met Soc of China, 1997, 7(4): 156~ 159.

[3] ZHU Jian-jun. Bayesian hypothesis test for deformation analysis [J]. Geomatica, 1995, 49(2): 283~ 288.

[4] Carlin B P and Louis T A. Bayes and Empirical Bayes Methods for Data Analysis [M]. London and New York: Chapman & Hall, 1996.

[5] ZHANG Jin-huai and TANG Xue-mei. Bayes Approach [M]. Changsha: National Defense University of Technology Press, 1989.

[6] Berger J O. Statistical Decision Theory [M]. New York: Springer Verlag New York Inc, 1980.

[7] CHEN Peng and CHEN Xi-ru. Parameter Estimate [M]. Shanghai: Shangshai Science and Technology Press, 1985.

[8] WANG Song-gui. The Theory of Linear Model and Its Application [M]. Hefei: Anhui Education Press, 1987.

[9] FANG Ke-tai, JIN Hui and CHEN Qin-yun. Practical Regression Analysis [M]. Beijing: Science Press, 1988.

(Edited by HE Xue-feng)