SCRS/2019/059 Collect. Vol. Sci. Pap. ICCAT, 76(5): 72-77 (2019)

AN ANALYTICAL APPROACH FOR DIAGNOSTICS OF PARAMETERS ESTIMATION IN MANAGEMENT STRATEGY EVALUATION

Andrey Mikhaylov[1]

SUMMARY

The problems of diagnostics of parameter estimation within the framework of management strategy evaluation are discussed. The observations are generated by stochastic differential equation. The dependence of the error distribution on the model parameters, noise level, and duration of observation are investigated in general case. An example of an exactly solvable model is built up.

RÉSUMÉ

Le présent document aborde les problèmes de diagnostic de l'estimation des paramètres dans le cadre de l'évaluation de la stratégie de gestion. Les observations sont générées par une équation différentielle stochastique. La dépendance de la distribution d'erreur dans les paramètres du modèle, le niveau de bruit et la durée de l'observation sont étudiés *dans le cas général. Un exemple de modèle parfaitement résoluble est construit.*

RESUMEN

Se discuten los problemas de diagnósticos de estimación de los parámetros en el marco de la evaluación de estrategias de ordenación. Las observaciones se generan mediante una ecuación diferencial estocástica. La dependencia de la distribución de error en los parámetros del modelo, el nivel de ruido y la duración de la observación se investigan en los casos generales. Se construye un ejemplo de un modelo solucionable con precisión.

KEYWORDS

Mathematical models, Random processes, Simulation, Stochastic processes, Parameters estimation, Stochastic models, Stock assessment

[1] Federal Research Institute of Fishery and Oceanography, V.krasnoselskaya st. 17, Moscow, Russian Federation, mikhailov1984@gmail.com

One of the most important problems in management strategy evaluation (MSE) is the diagnostics of parameters estimation quality (Maunder, 2014). How much the parameter estimates deviate from their true values is a question. In general case, the true values of the parameters are determined through the operating model (OM), which is used as a simulation of reality. Parameter estimation and the following management procedure are evaluated using a simpler assessment model (AM). Diagnostic results of assessment are based on computer simulations (Kell et al, 2007). In this paper, a mathematical theory will be investigated, which can predict some analytical properties of simulations results. The foundation of this approach is that the model explicitly defines the way to estimate its parameters. The model hereby is understood as a set of dynamic relations and hypotheses about the distribution of random variables included in the model. Then, the estimates of the parameters will be, in its turn, a random variable, and their distributions will be determined by the true values of the parameters and the noise distribution of the model. This distribution will contain the information necessary to predict the results of computer simulation.

Consider the dynamic system described by the stochastic differential equation (1)

$$dy = f(t, y, x; \theta)dt + \sigma dw \tag{1}$$

where $y=(y_1,\dots,y_n)$ is dependent dynamic variable of process,
f – a smooth function,
$x=(x_1,\dots,x_m)$ independent variable,
$\theta=(\theta_1,\dots,\theta_d)$ vector of parameters,
$\sigma=\mathrm{diag}(\sigma_1,\dots,\sigma_n)$ diagonal matrix of covariations (standard deviation of noises),
$dw=(dw_1,\dots,dw_n)$ vector of differentials of standard Brownian motions (Oksendal.2000).

Any stochastic differential equation can be reduced to the form (1) by means of the corresponding change of variables. In this case, the noises of each components are independent, and their dispersions are constant.

$$\left\langle dw_\alpha(t_1)dw_\beta(t_2)\right\rangle = \delta_{\alpha\beta}\delta(t_1 - t_2) \tag{2}$$

Where $\langle g \rangle$ is the mathematical expectation of a variable or a function g.

Suppose that the observations are generated by equation (1). We are interested in how much the parameters estimated from these observations will differ from the true values. Since the noise is Gaussian in the construction of the model, the parameters estimation $\hat{\theta}$ is a minimum time-average square deviation

$$\hat{\theta} = \arg\min_{\theta} \overline{\left(dy - f(t, y, x; \theta)dt\right)^2} \tag{3}$$

Where $\overline{g}$ is time-average of a variable or a function g.

Equation (3) defines a random variable whose moments can be expressed in terms of the parameters of equation (1). To do this, substitute (1) in (3)

$$\hat{\theta} = \arg\min_{\hat{\theta}} \overline{\left(\left(f(t, y_\sigma(\theta), x; \theta) - f(t, y_\sigma(\theta), x; \hat{\theta})\right)dt + \sigma dw\right)^2} \tag{4}$$

Where y_σ is the solution of (1) with standard deviations of noises σ.

An explicit equation for $\hat{\theta}$ can be obtained by differentiating (4)

$$\overline{\left(\left(f(t, y_\sigma(\theta), x; \hat{\theta}) - f(t, y_\sigma(\theta), x; \theta)\right)\frac{\partial}{\partial\hat{\theta}} f(t, y_\sigma(\theta), x; \hat{\theta})\right)} = \sigma\overline{\left(\left(\frac{\partial}{\partial\hat{\theta}} f(t, y_\sigma(\theta), x; \hat{\theta})\right)dw\right)} \tag{5}$$

If $\sigma=0$, the estimate is accurate, and $\hat{\theta}=\theta$. We will calculate the parameters estimations in the framework of perturbation theory by σ.

$$\hat{\theta} = \theta + \sum \frac{1}{k!}\frac{\partial^k \hat{\theta}}{\partial \sigma^k}\sigma^k \tag{6}$$

Substituting (6) in (5), we find the expression for the first term of the series

$$\frac{\partial\hat{\theta}}{\partial\sigma} = \overline{\left(\frac{\partial f(t, y_0, x; \theta)}{\partial\theta}\frac{\partial f(t, y_0, x; \theta)}{\partial\theta}\right)}^{-1}\overline{\left(\left(\frac{\partial f(t, y_0, x; \theta)}{\partial\theta}\right)dw\right)} \tag{7}$$

The expression (7) is linear by noises, and all coefficients depend on only deterministic solution of (1). Mathematical expectation of first term is zero. Therefore, in first order by σ, the estimate is unbiased

$$\left\langle\frac{\partial\hat{\theta}}{\partial\sigma}\right\rangle = 0 \tag{8}$$

Now calculate the covariance matrix:

$$\left\langle\frac{\partial\hat{\theta}_i}{\partial\sigma_\alpha}\frac{\partial\hat{\theta}_j}{\partial\sigma_\beta}\right\rangle = \frac{1}{T}\overline{\left(\frac{\partial f}{\partial\theta}\frac{\partial f}{\partial\theta}\right)}^{-1}_{ik}\overline{\left(\frac{\partial f}{\partial\theta}\frac{\partial f}{\partial\theta}\right)}^{-1}_{jl}\overline{\left(\left(\frac{\partial f}{\partial\theta_k}\right)\left(\frac{\partial f}{\partial\theta_l}\right)\right)}\delta_{\alpha\beta} \tag{9}$$

The covariance matrix is inversely proportional to the observation time and is completely determined by the derivative of the function *f* by the parameters/

To calculate (9) we use (10) as the result of (2)

$$\langle a(t)dw_\alpha(t)b(\tau)dw_\beta(\tau)\rangle = a(t)b(t)\delta_{\alpha\beta} \Rightarrow \langle\overline{adw_\alpha}\,\overline{bdw_\beta}\rangle = \frac{1}{T}\overline{ab}\delta_{\alpha\beta} \tag{10}$$

Where *a* and *b* are some integrable functions.

Using (5) and (6), we derive an equation for the second order of perturbation theory:

$$\frac{1}{2}\overline{\frac{\partial f}{\partial\theta_i}\frac{\partial f}{\partial\theta_j}}\frac{\partial^2\hat{\theta}_j}{\partial\sigma^2} + \overline{\frac{\partial f}{\partial\theta_i}\frac{\partial^2 f}{\partial\theta_j\partial y}\frac{\partial y_\sigma}{\partial\sigma}}\frac{\partial\hat{\theta}_j}{\partial\sigma} + \frac{1}{2}\overline{\frac{\partial f}{\partial\theta_i}\frac{\partial^2 f}{\partial\theta_j\partial\theta_k}}\frac{\partial\hat{\theta}_j}{\partial\sigma}\frac{\partial\hat{\theta}_k}{\partial\sigma} + \\ + \overline{\frac{\partial^2 f}{\partial\theta_i\partial y}\frac{\partial y_\sigma}{\partial\sigma}\frac{\partial f}{\partial\theta_j}}\frac{\partial\hat{\theta}_j}{\partial\sigma} = \overline{\left(\frac{\partial^2 f}{\partial\theta_i\partial y}\frac{\partial y_\sigma}{\partial\sigma}dw\right)} \tag{11}$$

The second order terms are expressed nonlinearly through the first order terms and the derivative y_σ by σ. The time dependence of this derivative can be found by using the equations in variations (Pontryagin, Lohwater, 1962)

$$\frac{\partial y}{\partial\sigma}(t) = \int_0^t \exp\left(\int_{t_1}^t \frac{\partial f(\tau, y(\tau), x(\tau); \theta)}{\partial y}d\tau\right)dw(t_1) \tag{12}$$

To calculate correlations, we need another consequence of (2)

$$\left\langle a(t)dw(t)\int_{t_1}^{t_2} b(\tau)dw(\tau)\right\rangle = a(t)b(t)\Theta(t_2 - t)\Theta(t - t_1) \tag{13}$$

Where Θ is Heaviside’s function.

Substituting (12) in (13), we find

$$\left\langle \frac{\partial y}{\partial \sigma}(t_2) dw(t_1) \right\rangle = \exp\left(\int_{t_1}^{t_2} \frac{\partial f(\tau, y(\tau), x(\tau); \theta)}{\partial y} d\tau \right) \Theta(t_2 - t_1) \tag{14}$$

Also, substituting (7) in (11), we find the expression for second term perturbation theory for parameters estimations

$$\frac{\partial^2 \hat{\theta}_i}{\partial \sigma^2} = 2\left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{ij} \left[\overline{\left(\frac{\partial^2 f}{\partial \theta_j \partial y} \frac{\partial y_\sigma}{\partial \sigma} dw \right)} - \overline{\frac{\partial f}{\partial \theta_{(j}} \frac{\partial^2 f}{\partial \theta_{k)} \partial y} \frac{\partial y_\sigma}{\partial \sigma}} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{kl} \overline{\frac{\partial f}{\partial \theta_l} dw} \right] - \\ - \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{ij} \overline{\frac{\partial f}{\partial \theta_j} \frac{\partial^2 f}{\partial \theta_k \partial \theta_m}} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{kl} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{mn} \overline{\frac{\partial f}{\partial \theta_l} dw} \, \overline{\frac{\partial f}{\partial \theta_m} dw} \tag{15}$$

Using (14), we find the mathematical expectation of second term perturbation theory for parameters estimations

$$\left\langle \frac{\partial^2 \hat{\theta}_i}{\partial \sigma^2} \right\rangle = 2\left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{ij} \left[\overline{\left(\frac{\partial^2 f}{\partial \theta_j \partial y} \right)} - \frac{1}{T} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{kl} \overline{\frac{\partial f}{\partial \theta_{(j}} \frac{\partial^2 f}{\partial \theta_{k)} \partial y} \int_0^t \exp\left(\int_\tau^t \frac{\partial f}{\partial y} \right) \left(\frac{\partial f}{\partial \theta_l} \right) d\tau} \right] - \\ - \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{ij} \overline{\frac{\partial f}{\partial \theta_j} \frac{\partial^2 f}{\partial \theta_k \partial \theta_m}} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{kl} \left(\overline{\frac{\partial f}{\partial \theta}\frac{\partial f}{\partial \theta}} \right)^{-1}_{mn} \overline{\left(\left(\frac{\partial f}{\partial \theta_m} \right) \left(\frac{\partial f}{\partial \theta_l} \right) \right)} \tag{16}$$

where round brackets denote summarization over the indices.

The second correction of the parameter's estimation is not so trivial. The structure of this term is completely determined by the behavior of the deterministic solution near the true value of the parameters. In other words, the operating model determines the distribution of parameter errors. The use of assessment models is nothing more than optimizing for a certain subset of parameters. Thus, the estimation error consists of the error of the operating model and the deviation of the best estimation model from the true operating model. That is why the result of MSE is predetermined rather by the properties of the models themselves and not so much by the data.

As an example, consider the simplest model of population dynamics – the Fox model (Fox, 1970)

$$dB = -rB\ln(B/K) - qEB + \sigma B dw \tag{17}$$

Where B is the biomass of stock,
r the population growth rate,
K the environment capacity,
q – the catchability,
E the fishing effort

The equation (17) can be reduced to the following formula (18)

$$dy = -(\theta_y y + \theta_x x) + \sigma dw \tag{18}$$

Where $y = \ln CPUE - \overline{\ln CPUE}$
$x = E - \overline{E}$
$\theta_y = r$
$\theta_x = q$
and

$$\ln K = \overline{\ln CPUE} - \ln q + \frac{1}{r}\left(q\overline{E} + \frac{\ln CPUE(T) - \ln CPUE(0)}{T} \right) \tag{19}$$

The equation (18) has an exact solution (20)

$$y(T)=y(0)\exp(-\theta_y t)-\theta_x\int_0^T x(t)\exp(-\theta_y(T-t))dt+\sigma\int_0^T\exp(-\theta_y(T-t))dw\equiv y_0+\sigma y_1 \tag{20}$$

The equation (5) for the error of the estimated parameters gets a simple form (21)

$$\begin{pmatrix}\overline{y^2} & \overline{yx}\\ \overline{yx} & \overline{x^2}\end{pmatrix}\begin{pmatrix}\delta\theta_y\\ \delta\theta_x\end{pmatrix}=\sigma\begin{pmatrix}\overline{ydw}\\ \overline{xdw}\end{pmatrix} \tag{21}$$

Then the error of the estimated parameters is given by the expression (22)

$$\begin{pmatrix}\delta\theta_y\\ \delta\theta_x\end{pmatrix}=\sigma\frac{1}{\overline{y_\sigma^2}\,\overline{x^2}-\left(\overline{y_\sigma x}\right)^2}\begin{pmatrix}\overline{x^2} & -\overline{y_\sigma x}\\ -\overline{y_\sigma x} & \overline{y^2}\end{pmatrix}\begin{pmatrix}\overline{ydw}\\ \overline{xdw}\end{pmatrix}= \tag{22}$$

$$=\frac{\sigma}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2+2\sigma\left(\overline{y_0y_1}\,\overline{x^2}-\overline{y_0x}\,\overline{y_1x}\right)+\sigma^2\left(\overline{y_1^2}\,\overline{x^2}-\left(\overline{y_1x}\right)^2\right)}\bullet$$

$$\left(\begin{pmatrix}\overline{x^2} & -\overline{y_0x}\\ -\overline{y_0x} & \overline{y_0^2}\end{pmatrix}+\sigma\begin{pmatrix}0 & -\overline{y_1x}\\ -\overline{y_1x} & 2\overline{y_0y_1}\end{pmatrix}+\sigma^2\begin{pmatrix}0 & 0\\ 0 & 2\overline{y_1^2}\end{pmatrix}\right)\left(\begin{pmatrix}\overline{y_0dw}\\ \overline{xdw}\end{pmatrix}+\sigma\begin{pmatrix}\overline{y_1dw}\\ 0\end{pmatrix}\right)$$

The first non-trivial term in the expectation of the error will be proportional to σ^2 :

$$\begin{pmatrix}\langle\delta\theta_y\rangle\\ \langle\delta\theta_x\rangle\end{pmatrix}=\frac{\sigma^2}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2}\left(\begin{pmatrix}\overline{x^2} & -\overline{y_0x}\\ -\overline{y_0x} & \overline{y_0^2}\end{pmatrix}\begin{pmatrix}\langle\overline{y_1dw}\rangle\\ 0\end{pmatrix}+\left\langle\begin{pmatrix}0 & -\overline{y_1x}\\ -\overline{y_1x} & 2\overline{y_0y_1}\end{pmatrix}\begin{pmatrix}\overline{y_0dw}\\ \overline{xdw}\end{pmatrix}\right\rangle\right)- \tag{23}$$

$$-\frac{\sigma^2}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2}\left\langle\frac{\left(\overline{y_0y_1}\,\overline{x^2}-\overline{y_0x}\,\overline{y_1x}\right)}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2}\begin{pmatrix}\overline{x^2} & -\overline{y_0x}\\ -\overline{y_0x} & \overline{y_0^2}\end{pmatrix}\begin{pmatrix}\overline{y_0dw}\\ \overline{xdw}\end{pmatrix}\right\rangle=$$

$$=\sigma^2\begin{pmatrix}\overline{x^2}\\ -\overline{y_0x}\end{pmatrix}\langle\overline{y_1dw}\rangle+\sigma^2\begin{pmatrix}\langle\overline{xy_1}\,\overline{xdw}\rangle\\ 2\langle\overline{y_0y_1}\,\overline{xdw}\rangle-\langle\overline{y_1x}\,\overline{y_0dw}\rangle\end{pmatrix}-$$

$$-\frac{\sigma^2}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2}\begin{pmatrix}\overline{x^2} & -\overline{y_0x}\\ -\overline{y_0x} & \overline{y_0^2}\end{pmatrix}\begin{pmatrix}\langle\overline{y_0dw}\,\overline{y_0y_1}\rangle\overline{x^2}-\overline{y_0x}\langle\overline{y_0dw}\,\overline{y_1x}\rangle\\ \langle\overline{xdw}\,\overline{y_0y_1}\rangle\overline{x^2}-\overline{y_0x}\langle\overline{y_1x}\,\overline{xdw}\rangle\end{pmatrix}=$$

$$=\frac{\sigma^2}{T}\begin{pmatrix}\overline{x^2}\\ -\overline{y_0x}\end{pmatrix}+\frac{\sigma^2}{T}\begin{pmatrix}\overline{\int_0^t x(t)\exp(-\theta_y(t-\tau))x(\tau)d\tau}\\ 2\overline{\int_0^t y_0(t)\exp(-\theta_y(t-\tau))x(\tau)d\tau}-\overline{\int_0^t x(t)\exp(-\theta_y(t-\tau))y_0(\tau)d\tau}\end{pmatrix}-$$

$$-\frac{\sigma^2/T}{\overline{y_0^2}\,\overline{x^2}-\left(\overline{y_0x}\right)^2}\begin{pmatrix}\overline{x^2} & -\overline{y_0x}\\ -\overline{y_0x} & \overline{y_0^2}\end{pmatrix}\bullet$$

$$\bullet\begin{pmatrix}\overline{x^2}\,\overline{\int_0^t y_0(t)\exp(-\theta_y(t-\tau))y_0(\tau)d\tau}-\overline{y_0x}\,\overline{\int_0^t x(t)\exp(-\theta_y(t-\tau))y_0(\tau)d\tau}\\ \overline{x^2}\,\overline{\int_0^t y_0(t)\exp(-\theta_y(t-\tau))x(\tau)d\tau}-\overline{y_0x}\,\overline{\int_0^t x(t)\exp(-\theta_y(t-\tau))x(\tau)d\tau}\end{pmatrix}$$

All of the coefficients in (23) are explicit functions of the true values of the parameters.

The covariance matrix of the first order is also computed explicitly.

$$\begin{pmatrix} \langle \delta\theta_y{}^2 \rangle & \langle \delta\theta_x \delta\theta_y \rangle \\ \langle \delta\theta_x \delta\theta_y \rangle & \langle \delta\theta_x{}^2 \rangle \end{pmatrix} \approx \frac{\sigma^2}{\left(\overline{y_0{}^2}\,\overline{x^2} - \left(\overline{y_0 x}\right)^2\right)^2} \left\langle \left(\begin{pmatrix} \overline{x^2} & -\overline{y_0 x} \\ -\overline{y_0 x} & \overline{y_0{}^2} \end{pmatrix} \begin{pmatrix} \overline{y_0 dw} \\ \overline{x dw} \end{pmatrix} \right)^{\otimes 2} \right\rangle = \tag{24}$$

$$= \frac{1}{T} \frac{\sigma^2}{\left(\overline{y_0{}^2}\,\overline{x^2} - \left(\overline{y_0 x}\right)^2\right)^2} \begin{pmatrix} \overline{x^2} & -\overline{y_0 x} \\ -\overline{y_0 x} & \overline{y_0{}^2} \end{pmatrix} \begin{pmatrix} \overline{y_0{}^2} & \overline{y_0 x} \\ \overline{y_0 x} & \overline{x^2} \end{pmatrix} \begin{pmatrix} \overline{x^2} & -\overline{y_0 x} \\ -\overline{y_0 x} & \overline{y_0{}^2} \end{pmatrix} =$$

$$= \frac{1}{T} \frac{\sigma^2}{\overline{y_0{}^2}\,\overline{x^2} - \left(\overline{y_0 x}\right)^2} \begin{pmatrix} \overline{x^2} & -\overline{y_0 x} \\ -\overline{y_0 x} & \overline{y_0{}^2} \end{pmatrix}$$

The obtained formula describes the dependence of the accuracy of the estimate on the process error, the duration of observation, and the dynamics of the independent variables.

References

FLR: an open-source framework for the evaluation and development of management strategies. Kell L. T., Mosqueira I., Grosjean P., Fromentin J-M., Garcia D., Hillary R., Jardim, E., Mardle S., Pastoors M. A., Poos J. J., Scott F., Scott R. D. ICES Journal of Marine Science. 2007. 64: 640–646.

Fox W. W. An exponential yield model for optimizing exploited fish populations. Trans. Am. Fish. 1970. Soc. 99, P. 80-88.

Maunder M. N. 2014. Management strategy evaluation (MSE) implementation in stock synthesis: Application to Pacific Bluefin tuna // Inter-American Tropical Tuna Commission. Doc. SAC-05-10b. 11 p.

Oksendal B. Stochastic differential equation. An introduction and application//N.Y.Springer 2000

Pontryagin L. S., Lohwater A. J. Ordinary Differential Equations. // Elsevier Inc 1962 305p.