THRESHOLD SELECTION IN THE GLR TEST FOR CHANGE DETECTION IN PARTIALLY OBSERVED DIFFUSION PROCESSES

F. Campillo, F. Le Gland
INRIA Sophia Antipolis, BP 93
06902 SOPHIA ANTIPOLIS Cédex
France

e–mail : campillo@sophia.inria.fr
legland@sophia.inria.fr

Keywords : change detection, generalized likelihood ratio test, partial observation, diffusion processes, large deviations.

1 Introduction

This paper is concerned with change detection in a partially observed diffusion model, i.e. detection of a change, occuring at some unknown change time, in the drift coefficient or the observation function. The problem is to decide, based on observations $\{Y_t\,,\,0\leq t\leq T\}$ in a finite time interval, between :

- The *null* hypothesis (H_0)

$$dX_t \;=\; b(X_t)\,dt + \varepsilon\,dW_t^\varepsilon\;,$$

$$dY_t \;=\; h(X_t)\,dt + \varepsilon\,dV_t^\varepsilon\;,$$

where $\{W_t^\varepsilon\,,\,t\geq 0\}$ and $\{V_t^\varepsilon\,,\,t\geq 0\}$ are independent Wiener processes under P_ε, and $X_0=\bar{x}$.

- The *alternate* (multiple) hypothesis (H_1) :

For some $0\leq\tau\leq T'$, with $T'<T$

$$dX_t \;=\; [b(X_t)+a(X_t)\,1_{(t\,\geq\,\tau)}]\,dt + \varepsilon\,dW_t^{\varepsilon,\tau}\;,$$

$$dY_t \;=\; [h(X_t)+g(X_t)\,1_{(t\,\geq\,\tau)}]\,dt + \varepsilon\,dV_t^{\varepsilon,\tau}\;,$$

where $\{W_t^{\varepsilon,\tau}\,,\,t\geq 0\}$ and $\{V_t^{\varepsilon,\tau}\,,\,t\geq 0\}$ are independent Wiener processes under $P_{\varepsilon,\tau}$, and $X_0=\bar{x}$.

The proposed approach is to use a generalized likelihood ratio (GLR) test, i.e. to consider the following region for rejecting the null hypothesis (H_0)

$$D_\varepsilon = \{\;\sup_{0\leq\tau\leq T'}\rho_\varepsilon(\tau) > c\;\}\;,$$

where c is a given threshold, and $\rho_\varepsilon(\tau)=\ell_\varepsilon(\tau)-\ell_\varepsilon(T)$ denotes a suitably normalized log–likelihood function for the estimation of the change time τ, based on observations $\{Y_t\,,\,0\leq t\leq T\}$.

The purpose of this paper is to prove that the threshold c can be chosen in a such way that both the probability of *false alarm* and the probability of *no detection* go to zero as $\epsilon\downarrow 0$. These probabilities are defined respectively as

$$F_\varepsilon = P_\varepsilon(D_\varepsilon) \quad\text{and}\quad M_\varepsilon = \sup_{0\leq\tau_0\leq T'} P_{\varepsilon,\tau_0}(D_\varepsilon^c)\;.$$

This result is an extension to the partially observed case, of the results obtained in Campillo, Kutoyants, LeGland [2].

2 Log–likelihood Function

It was proved in Campillo, LeGland [1], that the log–likelihood function for the estimation of the change time τ, based on observations $\{Y_t\,,\,0\leq t\leq T\}$, can be expressed as

$$\ell_\varepsilon(\tau) = \varepsilon^2\,\log\int_\Omega \exp\{\;\frac{1}{\varepsilon^2}\,F(\tau,\omega)\;\}\;dP_{\varepsilon,\tau}(\omega)$$

where $\Omega=C([0,T]\,;\,\mathbf{R}^m)$, and

$$F(\tau,\omega) = \int_0^T h^\tau(t,\omega_t)\,dY_t - \tfrac{1}{2}\int_0^T |h^\tau(t,\omega_t)|^2\,dt\;.$$

Throughout the paper, the following time–dependent coefficients, parametrized by the unknown parameter τ, are used

$$b^\tau(t,x) \;=\; b(x)+a(x)\,1_{(t\,\geq\,\tau)}\;,$$

$$h^\tau(t,x) \;=\; h(x)+g(x)\,1_{(t\,\geq\,\tau)}\;.$$

An efficient way of computing the GLR test statistics has been proposed in [1], using a pair of forward–backward SPDE's, one involving the coefficients before the change, the other involving the coefficients after the change. With this formulation, integration over the path space Ω is replaced by integration over the state space $\mathbf{R}^m$.

Indeed, consider the forward SPDE

$$d\mu_t^\varepsilon \;=\; \{\; \tfrac{1}{2}\,\varepsilon^2\,\Delta\mu_t^\varepsilon - \nabla(b\;\mu_t^\varepsilon)\;\}\,dt + \varepsilon^{-2}\;h\;\mu_t^\varepsilon\;dY_t\;,$$

with Dirac initial condition $\mu_0^\varepsilon = \delta_{\bar{x}}$, and the backward SPDE

$$dv_t^\varepsilon + \{\;\tfrac{1}{2}\,\varepsilon^2\;\Delta v_t^\varepsilon + [b+a]\;\nabla v_t^\varepsilon\;\}\,dt$$

$$+\varepsilon^{-2}\;[h+g]\;v_t^\varepsilon\;dY_t \;=\; 0\;,$$

with initial condition (at final time) $v_T^\varepsilon \equiv 1$.

Proposition 2.1 *The log–likelihood function $\rho_\varepsilon(\tau)$ can be computed as*

$$\rho_\varepsilon(\tau) = \varepsilon^2\;\log\frac{\displaystyle\int_{\mathbf{R}^m} v_\tau^\varepsilon(x)\,\mu_\tau^\varepsilon(dx)}{\displaystyle\int_{\mathbf{R}^m}\mu_T^\varepsilon(dx)}\;,$$

and the GLR test statistics as

$$\sup_{0\le\tau\le T'}\;\rho_\varepsilon(\tau) = \varepsilon^2\;\log\;\sup_{0\le\tau\le T'}\;\frac{\displaystyle\int_{\mathbf{R}^m} v_\tau^\varepsilon(x)\,\mu_\tau^\varepsilon(dx)}{\displaystyle\int_{\mathbf{R}^m}\mu_T^\varepsilon(dx)}\;.$$

3 Large Deviations

The key idea to study the asymptotic behaviour of the error probabilities F_ε and M_ε when $\varepsilon\downarrow 0$, is to study the limiting expression of the GLR test statistics $\rho_\varepsilon(\tau)$ as $\varepsilon\downarrow 0$, using large deviations arguments similar to those in James, LeGland [3]. The situation can be summarized as follows.

- Under the null hypothesis (H_0)

(i) when $\varepsilon\downarrow 0$

$$P_\varepsilon(\sup_{0\le t\le T}|Y_t - y_t| > \delta) \longrightarrow 0\;,$$

where $\{y_t\;,\;0\le t\le T\}$ is the output of the following limiting deterministic system

$$\begin{aligned}\dot{x}_t \;&=\; b(x_t)\;,\quad x_0 = \bar{x}\\ y_t \;&=\; h(x_t)\;.\end{aligned}\tag{1}$$

(ii) $\{P_\varepsilon\;,\;\varepsilon>0\}$ satisfies a large deviations principle, with rate function

$$I(\phi) = \tfrac{1}{2}\int_0^T |\dot{\phi}_t - b(\phi_t)|^2\,dt\;,$$

if $\phi\in\Omega$ is absolutely continuous, $I(\phi) = +\infty$ otherwise.

- Under the alternate hypothesis (H_1) :

For all $0\le\tau\le T'$, with $T' < T$

(i) when $\varepsilon\downarrow 0$

$$P_{\varepsilon,\tau}(\sup_{0\le t\le T}|Y_t - y_t^\tau| > \delta) \longrightarrow 0\;,$$

where $\{y_t^\tau\;,\;0\le t\le T\}$ is the output of the following limiting deterministic system

$$\begin{aligned}\dot{x}_t^\tau \;&=\; b(x_t^\tau) + a(x_t^\tau)\,1_{(t\,\ge\,\tau)}\;,\quad x_0^\tau = \bar{x}\\ y_t^\tau \;&=\; h(x_t^\tau) + g(x_t^\tau)\,1_{(t\,\ge\,\tau)}\;.\end{aligned}\tag{2}$$

(ii) $\{P_{\varepsilon,\tau}\;,\;\varepsilon>0\}$ satisfies a large deviations principle, with rate function

$$I(\tau,\phi) = \tfrac{1}{2}\int_0^T |\dot{\phi}_t - b^\tau(t,\phi_t)|^2\,dt\;,$$

if $\phi\in\Omega$ is absolutely continuous, $I(\tau,\phi) = +\infty$ otherwise.

These large deviations results are used in the next two sections, to obtain the limiting behaviour, under either the null hypothesis (H_0) or the alternate hypothesis (H_1), of the log–likelihood function $\rho_\varepsilon(\tau)$ as $\varepsilon\downarrow 0$. In both cases, the convergence result is stated in the form of a Varadhan theorem in probability, uniform w.r.t. the parameter τ (see Proposition 4.1 and Proposition 5.1 below).

4 Probability of False Alarm

If no change has occurred, the following limiting expression holds for the statistics $\rho_\varepsilon(\tau)$ under P_ε as $\varepsilon\downarrow 0$

$$\begin{aligned}\rho_0(\tau) \;=\; -\inf_{\phi\,:\,\phi_0=\bar{x}}\;[\;\tfrac{1}{2}\int_0^T |\dot{\phi}_t - b^\tau(t,\phi_t)|^2\,dt\\ +\tfrac{1}{2}\int_0^T |y_t - h^\tau(t,\phi_t)|^2\,dt\;]\\ +\inf_{\phi\,:\,\phi_0=\bar{x}}\;[\;\tfrac{1}{2}\int_0^T |\dot{\phi}_t - b(\phi_t)|^2\,dt\end{aligned}$$

$$+ \tfrac{1}{2} \int_0^T |y_t - h(\phi_t)|^2 \, dt \,]$$

$$= - \inf_{\phi \, : \, \phi_0 = \bar{x}} [\tfrac{1}{2} \int_0^T |\dot{\phi}_t - b^\tau(t, \phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_0^T |y_t - h^\tau(t, \phi_t)|^2 \, dt \,] \leq 0 \, ,$$

where $\{y_t \, , \, t \geq 0\}$ is the output of the limiting deterministic system (1). Notice that the following expression is also available for $\rho_0(\tau)$

$$\rho_0(\tau) \; = \; - \inf_{\phi \, : \, \phi_0 = \bar{x}} [\tfrac{1}{2} \int_0^\tau |\dot{\phi}_t - b(\phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_0^\tau |y_t - h(\phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_\tau^T |\dot{\phi}_t - [b+a](\phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_\tau^T |y_t - [h+g](\phi_t)|^2 \, dt \,] \; .$$

Proposition 4.1 *The following large deviations result holds*

$$P_\varepsilon(\sup_{0 \leq \tau \leq T'} |\rho_\varepsilon(\tau) - \rho_0(\tau)| > \delta) \longrightarrow 0 \; ,$$

when $\varepsilon \downarrow 0$.

Define

$$\rho_0^* = \sup_{0 \leq \tau \leq T'} \rho_0(\tau) \leq 0 \; .$$

Proposition 4.2 *The probability of false alarm* F_ε *goes to zero as* $\varepsilon \downarrow 0$*, provided the threshold* c *satisfies*

$$c > c_{\min} = \rho_0^* \; . \qquad (3)$$

The lower bound $c_{\min}$ can be computed exactly, using a pair of forward–backward dynamic programming equations, one involving the coefficients before the change, the other involving the coefficients after the change. With this formulation, optimization over the path space Ω is repaced by optimization over the state space $\mathbf{R}^m$.

Indeed, consider the forward dynamic programming equation

$$\frac{\partial S_t}{\partial t} = -\tfrac{1}{2} \, |\nabla S_t|^2 - b \, \nabla S_t + \tfrac{1}{2} \, |y_t - h|^2 \; ,$$

with initial condition $S_0(\bar{x}) = 0$, and $S_0(x) = \bar{S}$ (with $\bar{S}$ large enough) for $x \neq \bar{x}$, and the backward dynamic programming equation

$$\frac{\partial V_t}{\partial t} - \tfrac{1}{2} \, |\nabla V_t|^2 + [b+a] \, \nabla V_t + \tfrac{1}{2} \, |y_t - [h+g]|^2 = 0 \; ,$$

with initial condition (at final time) $V_T \equiv 0$.

Proposition 4.3 *The limiting statistics* $\rho_0(\tau)$ *can be computed as*

$$\rho_0(\tau) = - \inf_{x \in \mathbf{R}^m} \{ \, S_\tau(x) + V_\tau(x) \, \} \; ,$$

and the threshold lower bound $c_{\min}$ *as*

$$c_{\min} = - \inf_{0 \leq \tau \leq T'} \inf_{x \in \mathbf{R}^m} \{ \, S_\tau(x) + V_\tau(x) \, \} \; .$$

As an alternative to the exact computation of the threshold lower bound $c_{\min}$, it is interesting to see whether it would be possible to obtain estimates that would be easier to compute, and therefore could be used in practice.

□ Lower bound for $c_{\min}$

From the particular choice $\phi \equiv x^\tau$, it follows that

$$\rho_0(\tau) \geq -\tfrac{1}{2} \int_\tau^T |y_t - y_t^\tau|^2 \, dt \; .$$

From the particular choice $\phi \equiv x$, it follows that

$$\rho_0(\tau) \geq - [\, \tfrac{1}{2} \int_\tau^T |a(x_t)|^2 \, dt + \tfrac{1}{2} \int_\tau^T |g(x_t)|^2 \, dt \,] \; ,$$

and therefore

$$c_{\min} \; \geq \; c''_{\min}$$

$$= \; - [\, \tfrac{1}{2} \int_{T'}^T |a(x_t)|^2 \, dt + \tfrac{1}{2} \int_{T'}^T |g(x_t)|^2 \, dt \,] \; ,$$

where $c''_{\min}$ would be exactly the threshold lower bound in the case of complete observation, see [2].

□ Upper bound for $c_{\min}$

From the definition of $\rho_0(\tau)$, it follows that

$$\rho_0(\tau) \; \leq \; - \inf_\phi [\, \tfrac{1}{2} \int_\tau^T |\dot{\phi}_t - [b+a](\phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_\tau^T |y_t - [h+g](\phi_t)|^2 \, dt \,] \; ,$$

and therefore

$$c_{\min} \; \leq \; c'_{\min}$$

$$= \; - \inf_\phi [\, \tfrac{1}{2} \int_{T'}^T |\dot{\phi}_t - [b+a](\phi_t)|^2 \, dt$$

$$+ \tfrac{1}{2} \int_{T'}^T |y_t - [h+g](\phi_t)|^2 \, dt \,] \leq 0 \; .$$

5 Probability of No Detection

If a change has occurred at time τ_0, the following limiting expression holds for the statistics $\rho_\varepsilon(\tau)$ under P_{ε,τ_0} as $\varepsilon \downarrow 0$

$$\begin{aligned}\rho(\tau_0,\tau) \;=\; &- \inf_{\phi\,:\,\phi_0=\bar{x}} \Big[\, \tfrac{1}{2} \int_0^T |\dot{\phi}_t - b^\tau(t,\phi_t)|^2\, dt \\ &\qquad + \tfrac{1}{2} \int_0^T |y_t^{\tau_0} - h^\tau(t,\phi_t)|^2\, dt \,\Big] \\ &+ \inf_{\phi\,:\,\phi_0=\bar{x}} \Big[\, \tfrac{1}{2} \int_0^T |\dot{\phi}_t - b(\phi_t)|^2\, dt \\ &\qquad + \tfrac{1}{2} \int_0^T |y_t^{\tau_0} - h(\phi_t)|^2\, dt \,\Big] \;,\end{aligned}$$

where $\{y_t^{\tau_0}\,,\, t \geq 0\}$ is the output of the limiting deterministic system (2).

Proposition 5.1 *The following large deviations result holds*

$$P_{\varepsilon,\tau_0}(\sup_{0\leq\tau\leq T'} |\rho_\varepsilon(\tau) - \rho(\tau_0,\tau)| > \delta) \longrightarrow 0 \;,$$

when $\varepsilon \downarrow 0$.

It is clear that the mapping $\tau \mapsto \rho(\tau_0,\tau)$ achieves its maximum for $\tau = \tau_0$, which implies that the MLE for the estimation of the change time is consistent. Define

$$\rho^*(\tau_0) = \max_{0\leq\tau_0\leq T'} \rho(\tau_0,\tau) = \rho(\tau_0,\tau_0) \;,$$

i.e.

$$\begin{aligned}\rho^*(\tau_0) \;=\; &\inf_{\phi\,:\,\phi_0=\bar{x}} \Big[\, \tfrac{1}{2} \int_0^T |\dot{\phi}_t - b(\phi_t)|^2\, dt \\ &\qquad + \tfrac{1}{2} \int_0^T |y_t^{\tau_0} - h(\phi_t)|^2\, dt \,\Big] \geq 0 \;,\end{aligned}$$

where $\{y_t^{\tau_0}\,,\, t \geq 0\}$ is the output of the limiting deterministic system (2).

Proposition 5.2 *The probability of no detection* M_ε *goes to zero as* $\varepsilon \downarrow 0$*, provided the threshold* c *satisfies*

$$c < c_{\max} = \inf_{0\leq\tau_0\leq T'} \rho^*(\tau_0) \;. \qquad (4)$$

The upper bound $c_{\max}$ can be computed exactly, using a family, parametrized by the *true* change time τ_0, of dynamic programming equations, all of them involving the coefficients before the change, but using different outputs $\{y_t^{\tau_0}\,,\, 0 \leq t \leq T\}$. With this formulation, optimization over the path space Ω is repaced by optimization over the state space $\mathbf{R}^m$.

Indeed, consider the forward dynamic programming equation

$$\frac{\partial S_t^{\tau_0}}{\partial t} = -\tfrac{1}{2}\, |\nabla S_t^{\tau_0}|^2 - b\, \nabla S_t^{\tau_0} + \tfrac{1}{2}\, |y_t^{\tau_0} - h|^2 \;,$$

with initial condition $S_0^{\tau_0}(\bar{x}) = 0$, and $S_0^{\tau_0}(x) = \bar{S}$ (with $\bar{S}$ large enough) for $x \neq \bar{x}$.

Proposition 5.3 *The limiting statistics* $\rho^*(\tau_0)$ *can be computed as*

$$\rho^*(\tau_0) = \inf_{x\in\mathbf{R}^m} S_T^{\tau_0}(x) \;,$$

and the threshold upper bound $c_{\max}$ *as*

$$c_{\max} = \inf_{0\leq\tau_0\leq T'} \inf_{x\in\mathbf{R}^m} S_T^{\tau_0}(x) \;.$$

Here again, as an alternative to the exact computation of the threshold upper bound $c_{\max}$, it is interesting to see whether it would be possible to obtain estimates that would be easier to compute, and therefore could be used in practice.

□ Upper bound for $c_{\max}$

From the particular choice $\phi \equiv x$, it follows that

$$\rho^*(\tau_0) \leq \tfrac{1}{2} \int_{\tau_0}^T |y_t - y_t^{\tau_0}|^2\, dt \;.$$

From the particular choice $\phi \equiv x^{\tau_0}$, it follows that

$$\rho^*(\tau_0) \leq \tfrac{1}{2} \int_{\tau_0}^T |a(x_t^{\tau_0})|^2\, dt + \tfrac{1}{2} \int_{\tau_0}^T |g(x_t^{\tau_0})|^2\, dt \;,$$

and therefore

$$\begin{aligned}c_{\max} \;&\leq\; c''_{\max} \\ &= \inf_{0\leq\tau_0\leq T'} \Big[\, \tfrac{1}{2} \int_{\tau_0}^T |a(x_t^{\tau_0})|^2\, dt \\ &\qquad + \tfrac{1}{2} \int_{\tau_0}^T |g(x_t^{\tau_0})|^2\, dt \,\Big] \;,\end{aligned}$$

where $c''_{\max}$ would be exactly the threshold upper bound in the case of complete observation, see [2].

□ Lower bound for $c_{\max}$

From the definition of $\rho^*(\tau_0)$, it follows that

$$\begin{aligned}\rho^*(\tau_0) \;\geq\; &\inf_{\phi} \Big[\, \tfrac{1}{2} \int_{\tau_0}^T |\dot{\phi}_t - b(\phi_t)|^2\, dt \\ &\qquad + \tfrac{1}{2} \int_{\tau_0}^T |y_t^{\tau_0} - h(\phi_t)|^2\, dt \,\Big] \;,\end{aligned}$$

and therefore

$$c_{\max} \geq c'_{\max}$$

$$= \inf_{0 \leq \tau_0 \leq T'} \inf_{\phi} \Big[\tfrac{1}{2} \int_{\tau_0}^{T} |\dot{\phi}_t - b(\phi_t)|^2 \, dt + \tfrac{1}{2} \int_{\tau_0}^{T} |y_t^{\tau_0} - h(\phi_t)|^2 \, dt \Big] \geq 0 \; .$$

6 Conclusion

From the previous discussion, it is possible to select the threshold c in such a way that both the probability of false alarm and the probability of no detection go to zero as $\varepsilon \downarrow 0$. Actually, the threshold c should satisfy simultaneously

$$c_{\min} < c < c_{\max} \; ,$$

where the threshold bounds $c_{\min} \leq 0$ and $c_{\max} \geq 0$ have been defined in (3) and (4) respectively. This is possible provided the following *detectability* assumption holds

$$c_{\min} < c_{\max} \; , \tag{5}$$

i.e. whenever one of the two above quantities is non–zero. Remark that for this assumption to hold, it is necessary that $T' < T$.

Estimates

$$c''_{\min} \leq c_{\min} \leq c'_{\min} \leq 0 \; ,$$

and

$$0 \leq c'_{\max} \leq c_{\max} \leq c''_{\max} \; ,$$

have also been obtained, where $c''_{\min}$ and $c''_{\max}$ would be the threshold bounds in the case of complete observation. Therefore, a *sufficient* condition for (5) to hold, is

$$c'_{\min} < c'_{\max} \; ,$$

whereas a *necessary* condition would be

$$c''_{\min} < c''_{\max} \; .$$

Robustness w.r.t. mis–specification of the change coefficients $a(\cdot)$ and $g(\cdot)$ can also be investigated, as in [2].

References

[1] F. CAMPILLO and F. LE GLAND. Likelihood based statistics for partially observed diffusion processes. In *1st European Control Conference, Grenoble 1991*, pages 2290–2295, Paris, 1991. Hermès.

[2] F. CAMPILLO, F. LE GLAND, and Y. KUTOYANTS. Asymptotics of the GLRT for the disorder problem in diffusion processes. Rapport de Recherche 1735, INRIA, July 1992.

[3] M.R. JAMES and F. LE GLAND. Identification of partially observed diffusions with small noise. In M.H.A. Davis and R.J. Elliott, editors, *Applied Stochastic Analysis*, volume 5 of *Stochastics Monographs*, pages 561–568. Gordon and Breach, 1991.