

The MA(1) process

Stationarity Ergodicity and Estimation

Remember that, a linear process $y = (y_t)_{t \in \mathbb{Z}}$ defined upon a strictly stationary and ergodic white noise process $\varepsilon = (\varepsilon_t)_{t \in \mathbb{Z}}$ and a sequence of coefficients $(\theta_j)_{j \in \mathbb{N}}$ which are absolutely summable, is defined by

$$y_t := \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} \quad \forall t \in \mathbb{Z}.$$

Consider now a linear process where ~~$\theta_j = \begin{cases} 1, & j=0 \\ 0, & j=1 \\ 0, & j>1 \end{cases}$~~

Given all the above, the MA(1) process is defined as the process $(y_t)_{t \in \mathbb{Z}}$ where $y_t = \theta \varepsilon_{t-1} + \varepsilon_t$, for some $\theta \in \mathbb{R}$

$\theta_j = 1, j=0$
$\theta_j = \theta, j=1$
$\theta_j = 0, j>1$

• Given the results in the paragraph Transformations and Heredity and the ones about: Causal linear process, since $(\varepsilon_t)_{t \in \mathbb{Z}}$ is strictly stationary and ergodic, $(y_t)_{t \in \mathbb{Z}}$ is also a strictly stationary and ergodic process.

• Remember that, for a linear process built on a stationary ergodic WNR (ε_0^2) we have already proved:

- $E(y_t) = 0 \quad \forall t \in \mathbb{Z}$, independent of t .

- $E(y_t y_{t-k}) = E(\varepsilon_0^2) \sum_{j=0}^{\infty} \theta_j \theta_{j+k} = \begin{cases} E(\varepsilon_0^2) (1 + \theta^2), & k=0 \\ E(\varepsilon_0^2) \theta, & k=1 \\ 0, & k>1 \end{cases}$ independent of t , $\forall t \in \mathbb{Z}, \forall k \in \mathbb{Z}$

thus, $(y_t)_{t \in \mathbb{Z}}$ is weakly stationary.

Notice that,

► The stationarity ergodicity property and the weak stationarity do not depend on θ .

► Strict stationarity and ergodicity holds if $(\varepsilon_t)_{t \in \mathbb{Z}}$ is simply stationary ergodic

► Weak stationarity property holds if $(\varepsilon_t)_{t \in \mathbb{Z}}$ is simply WNR $(E\varepsilon_0^2)$.

• GMM Estimation of the MA(1) parameter

In the case of an AR(1) process since essentially we had a linear model we did OLS, however, in the context of MA(1), $y_t = \theta_0 \epsilon_{t-1} + \epsilon_t$, the process $(\epsilon_{t-1})_{t \in \mathbb{Z}}$ is non-observable, so we can't derive the OLS estimator.

Suppose for the sake of simplicity that $(\epsilon_t)_{t \in \mathbb{Z}}$ is comprised by iid random variables such that $E(\epsilon_0^2) = 1$, $k_4 = E(\epsilon_0^4) < +\infty$.

We know that $(y_t)_{t \in \mathbb{Z}}$ satisfies $y_t = \theta_0 \epsilon_{t-1} + \epsilon_t$ for $\theta_0 \in \mathbb{R}$ and our statistical problem is the estimation of θ_0 .

Thus, the semi-parametric statistical model is:

$$\left\{ (y_t(\theta))_{t \in \mathbb{Z}}, y_t = \theta \epsilon_{t-1} + \epsilon_t, \theta \in \mathbb{R}, \epsilon_t \text{ iid}, E(\epsilon_0) = 0, E(\epsilon_0^2) = 1, k_4 = E(\epsilon_0^4) < +\infty \right\}$$

Do we know anything about a moment of y_t that includes θ_0 ?

Remember that, for $E(\epsilon_0^2) = 1$, $E(y_t y_{t-1}) = \theta_0$.

Thus, by the analogy principle we can consider $M_T(w, \theta) = \frac{1}{T} \sum_{t=1}^T (y_t y_{t-1} - \theta)$ which is well-defined and the identification condition holds.

Hence, we can define the objective function as:

$$Q_T(w, \theta) = \left(\frac{1}{T} \sum_{t=1}^T (y_t y_{t-1} - \theta) \right)^2$$

$$\text{Obviously, } \hat{\theta}_T = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} Q_T(w, \theta) = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T (y_t y_{t-1} - \theta) \right)^2 = \frac{1}{T} \sum_{t=1}^T y_t y_{t-1}$$

We have that:

$$\bullet E_{\theta_0} \hat{\theta}_T = E_{\theta_0} \left(\frac{1}{T} \sum_{t=1}^T y_t y_{t-1} \right) = \frac{1}{T} \sum_{t=1}^T E_{\theta_0} (y_t y_{t-1}) = \frac{1}{T} \sum_{t=1}^T \theta_0 = \theta_0$$

therefore, the estimator is unbiased.

$$\bullet \operatorname{Var}_{\theta_0} \hat{\theta}_T = E_{\theta_0} \left[\left(\frac{1}{T} \sum_{t=1}^T y_t y_{t-1} - \theta_0 \right)^2 \right] = E_{\theta_0} \left[\left(\frac{1}{T} \sum_{t=1}^T y_t y_{t-1} \right)^2 \right] - \theta_0^2$$

$$= \frac{1}{T^2} E_{\theta_0} \left(\sum_{t=1}^T y_t y_{t-1} \sum_{s=1}^T y_s y_{s-1} \right) - \theta_0^2 = \frac{1}{T^2} \sum_{t,s=1}^T E(y_t y_{t-1} y_s y_{s-1}) - \theta_0^2$$

We have that:

$$\begin{aligned}
 E_{\theta_0}(y_t^2 y_{t-1}^2) &= E_{\theta_0}[(\epsilon_t + \theta_0 \epsilon_{t-1})^2 (\epsilon_{t-1} + \theta_0 \epsilon_{t-2})^2] \\
 &= E_{\theta_0}[(\epsilon_t^2 + 2\theta_0 \epsilon_t \epsilon_{t-1} + \theta_0^2 \epsilon_{t-1}^2) (\epsilon_{t-1}^2 + 2\theta_0 \epsilon_{t-1} \epsilon_{t-2} + \theta_0^2 \epsilon_{t-2}^2)] \\
 &= E_{\theta_0}(\epsilon_t^2 \epsilon_{t-1}^2) + 2\theta_0 E_{\theta_0}(\epsilon_t^2 \epsilon_{t-1} \epsilon_{t-2}) + \theta_0^2 E_{\theta_0}(\epsilon_t^2 \epsilon_{t-2}^2) + 2\theta_0 E_{\theta_0}(\epsilon_t \epsilon_{t-1}^3) \\
 &\quad + 4\theta_0^2 E_{\theta_0}(\epsilon_t \epsilon_{t-1}^2 \epsilon_{t-2}) + 2\theta_0^3 E_{\theta_0}(\epsilon_t \epsilon_{t-1} \epsilon_{t-2}^2) + \theta_0^2 E_{\theta_0}(\epsilon_{t-1}^4) \\
 &\quad + 2\theta_0^3 E_{\theta_0}(\epsilon_{t-1}^3 \epsilon_{t-2}) + \theta_0^4 E_{\theta_0}(\epsilon_{t-1}^2 \epsilon_{t-2}^2) \\
 &\stackrel{\text{indep.}}{=} E(\epsilon_t^2) E(\epsilon_{t-1}^2) + 2\theta_0 E(\epsilon_t^2) E(\epsilon_{t-1}) E(\epsilon_{t-2}) + \theta_0^2 E(\epsilon_t^2) E(\epsilon_{t-2}^2) \\
 &\quad + 2\theta_0 E(\epsilon_t) E(\epsilon_{t-1}^3) + 4\theta_0^2 E(\epsilon_t) E(\epsilon_{t-1}^2) E(\epsilon_{t-2}) + 2\theta_0^3 E(\epsilon_t) E(\epsilon_{t-1}) E(\epsilon_{t-2}^2) \\
 &\quad + \theta_0^2 E(\epsilon_{t-1}^4) + 2\theta_0^3 E(\epsilon_{t-1}^3) E(\epsilon_{t-2}) + \theta_0^4 E(\epsilon_{t-1}^2) E(\epsilon_{t-2}^2) \\
 &\stackrel{\text{homog.}}{=} 1 + 0 + \theta_0^2 + 0 + 0 + 0 + \theta_0^2 k_4 + 0 + \theta_0^4 \\
 &= 1 + (1 + k_4) \theta_0^2 + \theta_0^4
 \end{aligned}$$

Consider now the computation of $E_{\theta_0}(y_t y_{t-1} y_s y_{s-1})$.

When $|t-s|=1$:

$$\begin{aligned}
 E_{\theta_0}(y_t y_{t-1} y_s y_{s-1}) &= E_{\theta_0}(y_t y_{t-1} y_{t-1} y_{t-2}) = E(y_t y_{t-1}^2 y_{t-2}) = E[(\epsilon_t + \theta_0 \epsilon_{t-1}) y_{t-1}^2 y_{t-2}] \\
 &= E_{\theta_0}(\epsilon_t y_{t-1}^2 y_{t-2}) + \theta_0 E_{\theta_0}(\epsilon_{t-1} y_{t-1}^2 y_{t-2}) \\
 &\stackrel{\text{indep.}}{=} E(\epsilon_t) E(y_{t-1}^2 y_{t-2}) + \theta_0 E[(\epsilon_{t-1} + \theta_0 \epsilon_{t-2})^2 y_{t-2}] \\
 &= 0 + \theta_0 E[\epsilon_{t-1} (\epsilon_{t-1}^2 + 2\theta_0 \epsilon_{t-1} \epsilon_{t-2} + \theta_0^2 \epsilon_{t-2}^2) y_{t-2}] \\
 &= \theta_0 E(\epsilon_{t-1}^3 y_{t-2}) + 2\theta_0^2 E(\epsilon_{t-1}^2 \epsilon_{t-2} y_{t-2}) + \theta_0^3 E(\epsilon_{t-1} \epsilon_{t-2}^2 y_{t-2}) \\
 &\stackrel{\text{indep.}}{=} \theta_0 E(\epsilon_{t-1}^3) E(y_{t-2}) + 2\theta_0^2 E(\epsilon_{t-1}^2) E(\epsilon_{t-2} y_{t-2}) + \theta_0^3 E(\epsilon_{t-1}) E(\epsilon_{t-2}^2 y_{t-2}) \\
 &\stackrel{\text{homog.}}{=} 0 + 2\theta_0^2 E[\epsilon_{t-2} (\epsilon_{t-2} + \theta_0 \epsilon_{t-3})] + 0 \\
 &= 2\theta_0^2 E(\epsilon_{t-2}^2) + 2\theta_0^3 E(\epsilon_{t-2} \epsilon_{t-3}) = 2\theta_0^2 + 0 = 2\theta_0^2
 \end{aligned}$$

When $|t-s|=2$:

$$\begin{aligned}
 E_{\theta_0}(y_t y_{t-1} y_s y_{s-1}) &= E_{\theta_0}(y_t y_{t-1} y_{t-2} y_{t-3}) = E_{\theta_0}[(\epsilon_t + \theta_0 \epsilon_{t-1}) y_{t-2} y_{t-3}] \\
 &= E_{\theta_0}(\epsilon_t y_{t-2} y_{t-3}) + \theta_0 E_{\theta_0}(\epsilon_{t-1} y_{t-2} y_{t-3}) \\
 &\stackrel{\text{indep.}}{=} E_{\theta_0}(\epsilon_t) E(y_{t-2} y_{t-3}) + \theta_0 E_{\theta_0}[\epsilon_{t-1} (\epsilon_{t-1} + \theta_0 \epsilon_{t-2}) y_{t-2} y_{t-3}] \\
 &= 0 + \theta_0 E(\epsilon_{t-1}^2 y_{t-2} y_{t-3}) + \theta_0^2 E(\epsilon_{t-1} \epsilon_{t-2} y_{t-2} y_{t-3}) \\
 &\stackrel{\text{indep.}}{=} \theta_0 E(\epsilon_{t-1}^2) E(y_{t-2} y_{t-3}) + \theta_0^2 E(\epsilon_{t-1}) E(\epsilon_{t-2} y_{t-2} y_{t-3}) \\
 &= \theta_0 \cdot \theta_0 + 0 = \theta_0^2
 \end{aligned}$$

When $|t-s| \geq 2$:

$$E_0(y_t y_{t-1} y_s y_{s-1}) = E_0(y_t y_{t-1} y_{t-k} y_{t-k-1}) \stackrel{\text{indep.}}{=} E_0(y_t y_{t-1}) \cdot E(y_{t-k} y_{t-k-1}) = \sigma_0^2$$

$$\text{Hence, } E_0(y_t y_{t-1} y_s y_{s-1}) = \begin{cases} 1 + (1+k_4)\sigma_0^2 + \sigma_0^4, & |t-s|=0 \\ 2\sigma_0^2, & |t-s|=1 \\ \sigma_0^2, & |t-s| \geq 2 \end{cases}$$

So, now we have that:

$$\begin{aligned} \text{Var}_{\sigma_0} \theta_T &= \frac{1}{T^2} \sum_{t,s=1}^T E_0(y_t y_{t-1} y_s y_{s-1}) - \sigma_0^2 \\ &= \frac{1}{T^2} \left[T \cdot E_0(y_t^2 y_{t-1}^2) + 2(T-1) E_0(y_t y_{t-1}^2 y_{t-2}) + 2(T-2) E_0(y_t y_{t-1} y_{t-2} y_{t-3}) \right. \\ &\quad \left. + \dots \right] - \sigma_0^2 \\ &= \frac{1}{T^2} \left[T [1 + (1+k_4)\sigma_0^2 + \sigma_0^4] + 2(T-1) 2\sigma_0^2 + 2 \sum_{i=2}^{T-1} (T-i) \sigma_0^2 \right] - \sigma_0^2 \\ &= \frac{1}{T^2} \left[T [1 + (1+k_4)\sigma_0^2 + \sigma_0^4] + 2(T-1)\sigma_0^2 + 2 \sum_{i=1}^{T-1} (T-i) \sigma_0^2 \right] - \sigma_0^2 \\ &= \frac{1}{T^2} \left[T [1 + (1+k_4)\sigma_0^2 + \sigma_0^4] + 2(T-1)\sigma_0^2 + T(T-1)\sigma_0^2 \right] - \sigma_0^2 \\ &= \frac{1}{T} \left[1 + (1+k_4)\sigma_0^2 + \sigma_0^4 \right] + \frac{(T-1)(T+2)}{T^2} \sigma_0^2 - \sigma_0^2 \end{aligned}$$

Asymptotic properties of the GMM estimator

$$\theta_T = \frac{1}{T} \sum_{t=1}^T y_t y_{t-1}$$

Since $(y_t)_{t \in \mathbb{Z}}$ is strictly stationary and ergodic, the process $(y_t y_{t-1})_{t \in \mathbb{Z}}$ is also strictly stationary and ergodic. Therefore, by Birkhoff's LLN given that:

$(y_t y_{t-1})_{t \in \mathbb{Z}}$ is stationary ergodic and $E|y_t y_{t-1}| < +\infty$,

$$\frac{1}{T} \sum_{t=1}^T y_t y_{t-1} \xrightarrow{\text{P a.s.}} E_0(y_t y_{t-1}) = \sigma_0 \quad \text{P a.s.}$$

i.e. the GMM estimator is a consistent estimator of σ_0 .

In order to find the asymptotic distribution of $\hat{\theta}_T$ set:

$$\begin{aligned} v_t &= y_t y_{t-1} - \theta_0 = (\epsilon_t + \theta_0 \epsilon_{t-1})(\epsilon_{t-1} + \theta_0 \epsilon_{t-2}) - \theta_0 \\ &= \epsilon_t \epsilon_{t-1} + \theta_0 \epsilon_t \epsilon_{t-2} + \theta_0 \epsilon_{t-1}^2 + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} - \theta_0 \\ &= \epsilon_t \epsilon_{t-1} + \theta_0 \epsilon_t \epsilon_{t-2} + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} \end{aligned}$$

Assume that we have in hand a CLT for stationary ergodic s.m.d. processes.

$(y_t y_{t-1} - \theta_0)_{t \in \mathbb{Z}}$ is not a m.d. process thus the aforementioned wouldn't be applicable.

Consider, $f_t := \sigma(\epsilon_{t+i}, i \geq 0)$.

$$\begin{aligned} E(v_t / f_{t-1}) &= E[\epsilon_t \epsilon_{t-1} + \theta_0 \epsilon_t \epsilon_{t-2} + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} / f_{t-1}] \\ &= E(\epsilon_t \epsilon_{t-1} / f_{t-1}) + \theta_0 E(\epsilon_t \epsilon_{t-2} / f_{t-1}) + \theta_0 E(\epsilon_{t-1}^2 - 1 / f_{t-1}) + \theta_0^2 E(\epsilon_{t-1} \epsilon_{t-2} / f_{t-1}) \\ &= \epsilon_{t-1} E(\epsilon_t / f_{t-1}) + \theta_0 \epsilon_{t-2} E(\epsilon_t / f_{t-1}) + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} \\ &\stackrel{\text{indep.}}{=} \epsilon_{t-1} E(\epsilon_t) + \theta_0 \epsilon_{t-2} E(\epsilon_t) + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} \\ &= 0 + 0 + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} \neq 0 \end{aligned}$$

i.e. (v_t) is not a m.d. process w.r.t. $(f_t)_{t \in \mathbb{Z}}$, $f_t = \sigma(\epsilon_{t+i}, i \geq 0)$, thus not a s.m.d.

Can we find another filtration w.r.t. which $(v_t)_{t \in \mathbb{Z}}$ is a s.m.d. process? (Find it)

Consider the case where we can't find any filtration w.r.t. which the process is:

We need a generalization of the CLT for s.m.d. that we already know.

Giordini's CLT: Suppose that the stochastic process (v_t) is stationary ergodic. If

1. $E(v_0^2) < +\infty$

2. $\lim_{j \rightarrow \infty} E[E(v_t / f_{t-j})^2] = 0$ for all t

3. for $k_{t,j} = E(v_t / f_{t-j}) - E(v_{t-1} / f_{t-j-1})$, $\sum_{j=0}^{\infty} (E k_{t,j}^2)^{1/2} < +\infty$.

then $E(v_0) = 0$, $\sum_{j=0}^{\infty} |k_{t,j}| < +\infty$ and $\frac{1}{\sqrt{T}} \sum_{t=1}^T v_t \xrightarrow{d} z \sim N(0, \sum_{j=-\infty}^{\infty} k_j^2)$

Since $(y_t y_{t-1})_{t \in \mathbb{Z}}$ is stationary ergodic process, $(v_t)_{t \in \mathbb{Z}}$ is also stationary ergodic.

$$\begin{aligned} 1. E_0(v_0^2) &= E_0[(y_t y_{t-1} - \theta_0)^2] = E_0(y_t^2 y_{t-1}^2 - 2\theta_0 y_t y_{t-1} + \theta_0^2) \\ &= E(y_t^2 y_{t-1}^2) - 2\theta_0 E(y_t y_{t-1}) + \theta_0^2 = 1 + (1+k_4)\theta_0^2 + \theta_0^4 - 2\theta_0\theta_0 + \theta_0^2 \\ &= 1 + k_4\theta_0^2 + \theta_0^4 < +\infty \end{aligned}$$

$$\begin{aligned}
 2. E_0(v_t / f_{t-j}) &= E_0(y_t y_{t-1} - \theta_0 / f_{t-j}) \\
 &= E_0[\epsilon_t \epsilon_{t-1} + \theta_0 \epsilon_t \epsilon_{t-2} + \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2} / f_{t-j}] \\
 &= \begin{cases} \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2}, & j=1 \\ 0, & j > 1 \end{cases}
 \end{aligned}$$

therefore, as $j \rightarrow \infty$, $E_0(v_t / f_{t-j}) \rightarrow 0$, so the second condition holds.

3. From the previous calculations we have that

$$k_{t,j} = E_0(v_t / f_{t-j}) - E_0(v_t / f_{t-j-1}) = \begin{cases} \theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2}, & j=1 \\ 0, & j > 1 \end{cases}$$

thus,

$$E(k_{t,j}^2) = \begin{cases} E[\theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2}]^2, & j=1 \\ 0, & j > 1 \end{cases}$$

$$\begin{aligned}
 E(k_{t,1}^2) &= E\left[\left(\theta_0 (\epsilon_{t-1}^2 - 1) + \theta_0^2 \epsilon_{t-1} \epsilon_{t-2}\right)^2\right] = \\
 &= E\left[\theta_0^2 (\epsilon_{t-1}^2 - 1)^2 + 2\theta_0^3 (\epsilon_{t-1}^2 - 1) \epsilon_{t-1} \epsilon_{t-2} + \theta_0^4 \epsilon_{t-1}^2 \epsilon_{t-2}^2\right] \\
 &\stackrel{\text{indep.}}{=} \theta_0^2 E(\epsilon_{t-1}^4 - 2\epsilon_{t-1}^2 + 1) + 2\theta_0^3 E((\epsilon_{t-1}^2 - 1) \epsilon_{t-1}) E(\epsilon_{t-2}) + \theta_0^4 E(\epsilon_{t-1}^2) E(\epsilon_{t-2}^2) \\
 &= \theta_0^2 (\kappa_4 - 1) + 0 + \theta_0^4 \\
 &= \theta_0^2 (\kappa_4 - 1) + \theta_0^4
 \end{aligned}$$

$$\text{Hence, } \sum_{j=1}^{\infty} \left(E_0(k_{t,j}^2)\right)^{1/2} = \left(E_0(k_{t,1}^2)\right)^{1/2} = \sqrt{\theta_0^2 (\kappa_4 - 1) + \theta_0^4}$$

Since, Gordin's conditions hold for the process $(y_t y_{t-1} - \theta_0)_{t \in \mathbb{Z}}$,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T (y_t y_{t-1} - \theta_0) \xrightarrow{d} z \sim \mathcal{N}\left(0, \sum_{j=-\infty}^{\infty} \gamma_j\right)$$

$$\text{where } \gamma_j = \text{Cov}(v_t, v_{t-j}) = E_0\left[(y_t y_{t-1} - \theta_0)(y_{t-j} y_{t-j-1} - \theta_0)\right]$$

$$= E_0(y_t y_{t+1} y_{t-j} y_{t-j-1}) - \theta_0^2 = \begin{cases} 1 + (1 + \kappa_4) \theta_0^2 + \theta_0^4 - \theta_0^2, & j=0 \\ 2\theta_0^2 - \theta_0^2, & |j|=1 \\ \theta_0^2 - \theta_0^2, & |j| > 1. \end{cases}$$

$$\gamma_j = \begin{cases} 1 + \kappa_4 \theta_0^2 + \theta_0^4, & j=0 \\ \theta_0^2, & |j|=1 \\ 0, & |j| > 1 \end{cases}$$

$$\text{and } \sum_{j=-\infty}^{\infty} \gamma_j = 1 + \kappa_4 \theta_0^2 + \theta_0^4 + 2\theta_0^2 = (1 + \theta_0^2)^2 + \kappa_4 \theta_0^2$$

Therefore,

$$\sqrt{T} (\hat{\theta}_T - \theta_0) \xrightarrow{d} z \sim \mathcal{N}\left(0, (1 + \theta_0^2)^2 + \kappa_4 \theta_0^2\right)$$