ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1996

Seria: AUTOMATYKA z. 120 Nr kol. 1340

Marian BŁACHUTA

ESTYMACJA CIĄGŁYCH STOCHASTYCZNYCH SYSTEMÓW DYNAMICZNYCH NA PODSTAWIE POMIARÓW PRÓBKOWANYCH *

Streszczenie. Artykuł dotyczy estymacji parametrów stacjonarnych procesów gaussowskich o wymiernej gęstości spektralnej na podstawie dyskretnych w czasie próbek z błędem pomiarowym za pomocą dokładnej metody największej wiarygodności. Dla rozpatrywanej klasy procesów sformułowano stochastyczne równania stanu, zdefiniowano ich rozwiązanie i wyznaczono model dla czasu dyskretnego. Identyfikację modelu ciągłego i wariancji zakłócenia na podstawie tego modelu nazwano sformułowaniem naturalnym. Następnie przedstawiono kowariancyjnie równoważną realizację modelu dyskretnego, prowadzącą do modelu ARMA. Wskazano na problemy, jakie mogą wystąpić przy wyznaczaniu parametrów modelu ciągłego na podstawie modelu dyskretnego. Pokazano również, w jaki sposób model ARMA można wykorzystać do konstrukcji efektywnego numerycznie algorytmu estymacji parametrów modelu ciągłego. W końcu przedstawiono problemy estymacji parametrów ciągłych modeli wektorowych oraz podano metody ich rozwiązania.

ESTIMATION OF STOCHASTIC CONTINUOUS-TIME DYNAMICAL SYSTEMS BASED ON SAMPLED DATA

Summary. In the paper, the Maximum Likelihood parameter estimatation algorithms of continuoustime stochastic processes based on discrete-time samples are addressed. An algorithm for calculation of the likelihood function which uses a discrete-time model is presented. A covariance equivalent realization which leads to the ARMA model is then introduced and is employed within an computationally efficient method for the estimation of the continuous-time parameters. Finally, estimation problems of multivariate continuous-time models are stated and solved.

* *Niniejsza praca została wykonana w ramach Projektu Badawczego KBN nr 3 0683 91 01 w czasie pobytu autora w The University of Birmingham.*

1. Wstęp

Większość procesów spotykanych w realnym świecie ma charakter ciągły w czasie. Procesy z czasem ciągłym pojawiają się np.przy opisie ruchu samolotu w warunkach przepływu turbulentnego (Maine, Iliff 1981), ruchu anteny radarowej spowodowanej wirami powietrza (Unbehauen, Rao 1987), przechodzeniu światła przez ośrodek z losowo zmieniającym się współczynnikiem załamania (Kushner 1971) itp. Modele ciągłe mają również szerokie zastosowanie w ekonometrii (Bergstrom 1976, Sargan 1974, Wymer 1972).

Jedną z podstawowych klas ciągłych procesów stochastycznych jest klasa procesów gaussowskich, stacjonarnych z wymierną gęstością spektralną, podstawowa dla klasycznej liniowej teorii sterowania i filtracji.

Rozwój teorii sterowania automatycznego zawdzięcza wiele koncepcjom powstałym na gruncie modeli tych procesów. Jednak wraz z rozwojem techniki cyfrowej i jej zastosowaniem w automatyce, dla kompatybilności z obliczeniami cyfrowymi, problemy sterowania zaczęto dyskretyzować. Oryginalny, ciągły w czasie opis procesów sterowania zaczęto zastępować opisem dyskretnym, ważnym tylko dla dyskretnych chwil czasu. Również rozwój algorytmów sterowania związanych z modelami dyskretnymi zaczął dominować nad rozwojem algorytmów sterowania dla czasu ciągłego. W sytuacji, gdy istnieje olbrzymia literatura dotycząca identyfikacji modeli dyskretnych, literatura dotycząca identyfikacji modeli ciągłych jest znacznie skromniejsza.

Jednym z niebagatelnych powodów mniejszej popularności modeli ciągłych, zwłaszcza w środowiskach inżynierskich, może być większa złożoność matematyczna teorii stochastycznych równań różniczkowych i całkowych w porównaniu z teorią stochastycznych równań różnicowych.

Można zatem spytać: po co trudzić się z teorią ciągłych procesów dynamicznych w sytuacji, gdy implementacje algorytmów sterowania są dyskretne?

Powodów takich jest wiele. Algorytmy dyskretne dają wyniki jedynie dla dyskretnych chwil czasu ignorując zjawiska zachodzące pomiędzy nimi. Każda metoda projektowania dla czasu ciągłego, nawet w oparciu o próbkowane pomiary, wymaga modelu ciągłego. Z analogiczną sytuacją mamy do czynienia stosując metody czasu dyskretnego, pozostawiając jednak okres próbkowania jako jedną ze zmiennych podlegających doborowi. Z kolei przy wysokich częstotliwościach próbkowania zarówno modele dyskretne, jak i algorytmy dyskretne zaczynają tracić sens choćby ze względu na problemy numeryczne. Dlatego teoria ciągłych układów stochastycznych ma zasadnicze znaczenie dla dyskretno-czasowych algorytmów obliczeniowych pracujących przy wysokich częstotliwościach próbkowania. Bardzo ważne jest tu stwierdzenie, że w przypadku wzrostu częstotliwości próbkowania nie powstają problemy niemożliwe do pokonania.

Dyskretyzacja danego modelu ciągłego jest jednoznaczna. Jednakże w przypadku identyfikacji modelu dyskretnego dobór częstotliwości próbkowania nie jest sprawą błahą. W przypadku braku wiedzy a'priori o zakresie stałych czasowych, wyniki identyfikacji mogą być istotnie zależne od doboru czasu próbkowania. Ponadto odtworzenie modelu ciągłego na podstawie zidentyfikowanego modelu dyskretnego może okazać się niemożliwe.

Należy podkreślić, iż - z punktu widzenia automatyki - problemy przedstawione w tym opracowaniu dotyczą modeli zakłóceń i opierają się jedynie na obserwacji jednego sygnału. Innym problemem jest wyznaczanie modelu obiektu oparte na obserwacji dwu sygnałów: wejściowego i wyjściowego, przy czym wyjście może być zakłócane. Problem ten jest szeroko omówiony w pracy Nahorskiego (1991).

Estymacja parametrów obiektu łącznie z parametrami zakłócenia za pomocą metody największej wiarygodności jest w pełni możliwa. Bardzo ważnym elementem tej metody są wówczas wstępne oceny parametrów, które mogą być znalezione oddzielnie dla modelu obiektu i zakłócenia.

2. Modele ciągłych procesów stochastycznych

2.1. Równania procesów

Stosunkowo szeroką klasę procesów stochastycznych można przedstawić za pomocą układu równań:

$$d\boldsymbol{x}_t = A\boldsymbol{x}_t dt + \boldsymbol{g} d\xi_t, \tag{2.1}$$

$$z_t = \boldsymbol{d}'\boldsymbol{x}_t, \tag{2.2}$$

gdzie z_t jest procesem skalarnym, $\boldsymbol{x}_t$ jest n-wymiarowym wektorem stanu, A jest macierzą o stałych współczynnikach, $\boldsymbol{g}$ oraz $\boldsymbol{d}$ są stałymi wektorami, zaś ξ_t jest standardowym procesem Wienera (Gikhman, Skorokhod 1969, 1972). Symbol d oznacza różniczkę. Warunek początkowy $\boldsymbol{x}_0$ jest wektorem losowym o rozkładzie normalnym i o wartości oczekiwanej $\boldsymbol{m}_0 = E(\boldsymbol{x}_0)$ oraz kowariancji $Q_0 = E[(\boldsymbol{x}_0 - \boldsymbol{m}_0)(\boldsymbol{x}_0 - \boldsymbol{m}_0)']$.

W dalszym ciągu będziemy zakładali, że macierz A oraz wektory $\boldsymbol{g}$ i $\boldsymbol{d}$ są takie, iż są spełnione zależności:

$$\operatorname{rank}\,[\boldsymbol{g}, A\boldsymbol{g}, \ldots A^{n-1}\boldsymbol{g}] = n, \tag{2.3}$$

$$\operatorname{rank}\,[\boldsymbol{d}, A'\boldsymbol{d}, \ldots A'^{n-1}\boldsymbol{d}] = n. \tag{2.4}$$

Warunki (2.3)-(2.4) są kolejno warunkiem sterowalności i obserwowalności. Reprezentacja (2.1)-(2.2) ma kluczowe znaczenie w zagadnieniach prognozowania i estymacji.

Szczególne znaczenie, zwłaszcza w zagadnieniach estymacji, mają postacie kanoniczne konstruowane za pomocą minimalnej liczby parametrów. Przykładem może być tzw. postać kanoniczna obserwatorowa (Söderström, 1991), dla której:

$$A = \begin{bmatrix} -a_1 & 1 & . & . & . & 0 \\ -a_2 & 0 & 1 & . & . & 0 \\ . & . & . & . & . & . \\ -a_{n-1} & 0 & . & . & . & 1 \\ -a_n & 0 & . & . & . & 0 \end{bmatrix}, \; g = \begin{bmatrix} c_1 \\ c_2 \\ . \\ \vdots \\ c_n \end{bmatrix}, \; d = \begin{bmatrix} 1 \\ 0 \\ . \\ \vdots \\ 0 \end{bmatrix}. \tag{2.5}$$

Równanie (2.1), będące stochastycznym równaniem różniczkowym, należy rozumieć jako symboliczny zapis stochastycznego równania całkowego:

$$x_t = x_0 + \int_0^t A x_s ds + \int_0^t g d\xi_s, \tag{2.6}$$

przy czym druga z całek występująca we wzorze (2.6) jest całką stochastyczną w sensie Ito (Gikhman, Skorokhod 1969, 1972). Rozwiązanie równania (2.6) ma postać:

$$x_t = e^{At} x_0 + \int_0^t e^{A(t-s)} g d\xi(s). \tag{2.7}$$

2.2. Charakterystyki procesów stochastycznych

Wartość oczekiwana m_t oraz autokowariancja $Q_t(\tau)$ wyrażają się wzorami:

$$m_t = \mathrm{E}\,(x_t) = e^{At} m_0, \tag{2.8}$$

$$Q_t = \mathrm{E}\,(x_t - m_t)(x_t - m_t)' = e^{At} Q_0 e^{A't} + \int_0^t e^{A(t-s)} g g' e^{A'(t-s)} ds, \tag{2.9}$$

$$Q_t(\tau) = E(x_t - m_t)(x_{t+\tau} - m_{t+\tau})' = Q_t e^{A'\tau}. \tag{2.10}$$

Macierz kowariancji Q_t spełnia równanie różniczkowe Lapunowa:

$$\dot{Q}_t = A Q_t + Q_t A' + g g' \tag{2.11}$$

z warunkiem początkowym Q_0.

Obecnie skorzystamy z powyższych zależności w celu przedstawienia charakterystyk procesu z_t określonego równaniami (2.1)-(2.2). Wartość oczekiwana μ_t oraz funkcja autokorelacji $\rho_t(\tau)$ procesu z_t, zgodnie z (2.8)-(2.10) będą określone zależnościami:

$$\mu_t = d' m_t = d' e^{At} m_0, \tag{2.12}$$

$$\rho_t(\tau) = d' Q_t(\tau) d = d' Q_t e^{A'\tau} d. \tag{2.13}$$

2.3. Procesy stacjonarne

Ponieważ proces z_t opisany równaniami (2.1)-(2.2) jest w pełni scharakteryzowany przez dwa pierwsze momenty, warunkiem koniecznym i wystarczającym stacjonarności jest, aby wartość oczekiwana μ_t i funkcja korelacji $\rho_t(\tau)$ były niezależne od chwili bieżącej t.

Można pokazać, że proces z_t określony równaniami (2.1)-(2.2), z warunkami (2.3)-(2.4) jest stacjonarny wtedy i tylko wtedy gdy:

- macierz A jest stabilna, to znaczy, gdy dla $i = 1, 2 \ldots n$ zachodzi $\mathrm{Re}\,\lambda_i(A) < 0$,
- wartość oczekiwana warunku początkowego $m_0 = 0$,
- kowariancja Q_0 jest rozwiązaniem algebraicznego równania Lapunowa:

$$AQ + QA' = -gg'. \tag{2.14}$$

Przy spełnieniu powyższych warunków wartość oczekiwana procesu będzie zerowa, zaś funkcja korelacji wyrazi się zależnością:

$$\rho(\tau) = d'Q_0 e^{A'\tau} d. \tag{2.15}$$

Funkcję korelacji $\rho(\tau)$ można zatem wyrazić w postaci:

$$\rho(\tau) = \sum_{i=1}^{N} a_i p_i(\tau) e^{-\lambda_i \tau}, \tag{2.16}$$

gdzie $p_i(\tau)$ są wielomianami skończonego stopnia, a części rzeczywiste liczb λ_i są ściśle dodatnie.

Ponieważ gęstość widmowa mocy $\Sigma(\omega)$ oraz funkcja korelacji własnej $\rho(\tau)$ są związane zależnością (Kushner, 1971):

$$\Sigma(\omega) = \int_{-\infty}^{+\infty} \rho(t) e^{-j\omega t} dt, \tag{2.17}$$

więc na podstawie (2.14), (2.15) i (2.17)

$$\Sigma(\omega) = d'(sI - A)^{-1} gg'(-sI - A')^{-1} d|_{s=j\omega}. \tag{2.18}$$

$\Sigma(\omega)$ jest zatem rzeczywistą funkcją wymierną (Gikhman, Skorokhod, 1969) i może być przedstawiona w postaci:

$$\Sigma(\omega) = \frac{C(s)C(-s)}{A(s)A(-s)}|_{s=j\omega}, \tag{2.19}$$

gdzie:

$$A(s) = \det(sI - A), \tag{2.20}$$

$$C(s) = d'[\operatorname{adj}(sI - A)]g \tag{2.21}$$

oraz

$$A(s) = s^n + a_1 s^{n-1} + \ldots + a_n, \tag{2.22}$$

$$C(s) = \quad c_1 s^{n-1} + \ldots + c_n. \tag{2.23}$$

Jeśli wielomian $C(s)$ jest stopnia $m \leq n - 1$, wówczas część współczynników $c_1, c_2 \ldots c_{n-m-1}$ wielomianu $C(s)$ jest równa zeru. Wobec spełnienia warunków (2.3)-(2.4) wielomiany $A(s)$ i $C(s)$ są względnie pierwsze.

Z równania (2.19) widać, że przy zadanej funkcji gęstości widmowej $\Sigma(\omega)$ procesu z_t oraz przy ustalonym wielomianie $A(s)$ istnieje wiele wielomianów $C(s)$, a zatem wiele wektorów g, dla których układ (2.1)-(2.2) jest modelem procesu z_t. Pośród nich zawsze można znaleźć taki wektor g, aby wszystkie pierwiastki wielomianu $C(s)$ leżały w lewej półpłaszczyźnie. Realizację taką nazywa się *realizacją odwracalną*. Z równań (2.1)-(2.2), (2.5) i (2.19) wynika metodyka tworzenia realizacji stacjonarnego procesu stochastycznego o zadanej funkcji gęstości widmowej mocy $\Sigma(\omega)$. Problem wyznaczania stabilnego wielomianu $C(s)$ na podstawie funkcji wymiernej $\Sigma(\omega)$, będący kluczowym elementem tej metodyki, nosi nazwę *faktoryzacji spektralnej* (Åström, 1970).

3. Modele obserwacji procesów stochastycznych

Proces z_t jest na ogół obserwowany nie bezpośrednio, lecz za pomocą urządzeń pomiarowych wprowadzających własne błędy. Pomiary mogą mieć charakter ciągły w czasie lub dyskretny w czasie.

3.1. Model obserwacji ciągłej

W literaturze, np. Kushner (1971), jako równanie pomiaru przyjmuje się często:

$$dy_t = z_t dt + d\zeta_t = d' x_t dt + d\zeta_t \tag{3.1}$$

lub w wersji równania całkowego:

$$y_t = y_0 + \int_0^t d' x_s ds + \int_0^t d\zeta_s, \tag{3.2}$$

gdzie ζ_t jest procesem Wienera.

Zapis (3.2) implikuje dokonywanie operacji całkowania pewnej wielkości $d'x_t$ będącej liniową funkcją stanu z dodanym zakłóceniem. W rozwiązaniach technicznych pomiar całki sygnału wyjściowego na ogół nie jest stosowany.

W literaturze (Kwakernaak, Sivan, 1972) spotykany jest również opis wyjścia zawierający ciągły biały szum:

$$y_t = d'x_t + v_t. \tag{3.3}$$

Model (3.3) pozwala na sformułowanie równań ciągłego filtru Kalmana, jednakże jest on krytykowany (Kushner 1971), gdyż nie jest jasne, jak powinien wyglądać jego odpowiednik dyskretny w czasie.

Z obserwacjami ciągłymi związanych jest kilka metod estymacji transmitancji ciągłych układów dynamicznych (Saha, Rao 1983; Sinha, Rao 1991; Unbehauen, Rao 1987). Interesujące nas procesy stochastyczne są tam traktowane jako zakłócenia, zaś ich parametry nie są wyznaczane.

Z kolei Van Schuppen (1983), Moore (1988), Chen i Guo (1991) badają algorytmy rekurencyjne rozszerzonej metody najmniejszych kwadratów, pozwalające na estymację procesów stochastycznych. Okazuje się jednak, iż dla zbieżności konieczne są stosunkowo silne założenia odnośnie do charakterystyk procesów.

3.2. Model obserwacji dyskretnej

Sygnał ciągły, będący realizacją procesu ciągłego, jest we współczesnych rozwiązaniach technicznych na ogół próbkowany, to znaczy jest on mierzony w dyskretnych, równo odległych chwilach czasu $t_i = \Delta i$, gdzie Δ jest okresem próbkowania, zaś i jest liczbą całkowitą. Częstotliwość próbkowania może być stosunkowo wysoka.

Oznaczmy $y_i = y(t_i)$ oraz $x_i = x(t_i)$. Równanie pomiaru próbkowanego przyjmie postać:

$$y_i = d'x_i + r_i, \tag{3.4}$$

gdzie r_i jest dyskretnym białym szumem gaussowskim, tzn. $E[r_i r_j'] = \rho\delta_{ij}, E[r_i] = 0$.

Proces y_i jest procesem dyskretnym w czasie i może zostać opisany za pomocą układu dyskretnych równań stochastycznych (Kushner, 1971):

$$x_{i+1} = Fx_i + w_i, \tag{3.5}$$

$$y_i = d'x_i + r_i, \tag{3.6}$$

gdzie

$$F = e^{A\Delta}, \tag{3.7}$$

zaś w_i jest wektorowym białym szumem gaussowskim o macierzy kowariancji W:

$$W = \int_0^\Delta e^{As} gg' e^{As'} ds. \tag{3.8}$$

Procesy w_i oraz r_i są niezależne.

Niech Q oznacza macierz kowariancji stacjonarnego procesu dyskretnego x_i, spełniającą dyskretne równanie Lapunowa:

$$Q = FQF' + W. \tag{3.9}$$

Ponieważ wektory x_t oraz x_i są tymi samymi wektorami, więc ich macierze kowariancji są również równe. Z tej własności wynika sposób obliczenia całki (3.7). W tym celu należy rozwiązać równanie Lapunowa (2.14) względem Q, a następnie skorzystać z (3.9) otrzymując:

$$W = Q - FQF'. \tag{3.10}$$

4. Estymacja parametrów - sformułowanie naturalne

Punktem wyjścia do estymacji za pomocą metody największej wiarygodności jest jednokrokowa kalmanowska predykcja optymalna, której błąd, zwany procesem innowacji, jest białym szumem gausowskim.

4.1. Proces innowacji

W przypadku modelu (3.5)-(3.8) równania procesu innowacji ϵ_t wynikają z filtru Kalmana. Załóżmy, że $\rho > 0$ oraz przyjmijmy, że macierz kowariancji $W = \mathrm{cov}(w_i)$ jest proporcjonalna do ρ, czyli

$$W = W^0\rho. \tag{4.1}$$

Wówczas macierz kowariancji stanu początkowego $Q_0 = Q$ będzie również proporcjonalna do ρ:

$$Q = Q^0\rho, \tag{4.2}$$

gdzie Q^0 jest rozwiązaniem dyskretnego algebraicznego równania Lapunowa:

$$Q^0 = FQ^0F' + W^0 \tag{4.3}$$

lub równania (2.14) dla czasu ciągłego.

Oznaczmy przez $\hat{x}_{i|i-1}$ liniową ocenę stanu zapewniającą minimum błędu średniokwadratowego stanu x_i przy danych pomiarach do chwili $i-1$, zaś kowariancję błędu oceny stanu $\tilde{x}_i = x_i - \hat{x}_{i|i-1}$ przez:

$$\mathrm{cov}(\tilde{x}_i) = S_{i|i-1}\rho. \tag{4.4}$$

Wówczas jednokrokowy predyktor stanu jest określony poprzez reprezentację innowacyjną. Dla $i = 1, 2 \ldots N$ błędy predykcji jednokrokowej ϵ_i oraz ich wariancja $\operatorname{var}(\epsilon_i) = \sigma_i$ są określone zależnościami:

$$\epsilon_i = y_i - d'\hat{x}_{i|i-1}, \tag{4.5}$$

$$\hat{x}_{i+1|i} = F\hat{x}_{i|i-1} + h_i\epsilon_i, \hat{x}_{0|-1} = 0, \tag{4.6}$$

$$h_i = \frac{FS_{i|i-1}d}{1 + d'S_{i|i-1}d}, \tag{4.7}$$

$$S_{i+1|i} = W^0 + F(S_{i|i-1} - \frac{S_{i|i-1}dd'S_{i|i-1}}{1 + d'S_{i|i-1}d})F', S_{0|-1} = Q^0, \tag{4.8}$$

$$f_i = 1 + d'S_{i|i-1}d, \tag{4.9}$$

$$\sigma_i = \rho f_i. \tag{4.10}$$

Przy $i \to \infty$ zachodzi $S_{i|i-1} \to S$, gdzie S jest dodatnio określonym, symetrycznym rozwiązaniem algebraicznego równania Riccatiego:

$$S = W + F(S - \frac{Sdd'S}{1 + d'Sd})F' \tag{4.11}$$

Określając wektor h i skalar σ wzorami:

$$h = \frac{FSd}{1 + d'Sd}, \tag{4.12}$$

$$\sigma = \rho(1 + d'Sd), \tag{4.13}$$

otrzymamy postać asymptotyczną równań predyktora:

$$\hat{x}_{i+1|i} = F\hat{x}_{i|i-1} + h\epsilon_i, \tag{4.14}$$

$$y_i = d'\hat{x}_{i|i-1} + \epsilon_i. \tag{4.15}$$

Równanie (4.14) można także zapisać w postaci odwrotnej:

$$\hat{x}_{i+1|i} = F^*\hat{x}_{i|i-1} + hy_i, \tag{4.16}$$

$$\epsilon_i = y_i - d'\hat{x}_{i|i-1}, \tag{4.17}$$

przy czym

$$F^* = F - hd'. \tag{4.18}$$

Predyktor ten jest predyktorem asymptotycznie optymalnym.

4.2. Estymator ML

Ponieważ estymator największej wiarygodności jest estymatorem najefektywniejszym i posiadającym cały szereg zalet, w niniejszym rozdziale wyprowadzimy równania estymatora ML.

Zakładając, że wielomian $C(s)$ jest stopnia $m \leq n-1$, wektor θ estymowanych parametrów zawiera $n+m+1$ wielkości, na które składają się współczynniki wielomianów $A(s)$, $C(s)$ oraz wariancja szumu pomiarowego ρ:

$$\theta = [a_1, a_2 \dots a_n, c_{n-m}, \dots c_n, \rho]'. \tag{4.19}$$

Oznaczmy przez $y = \{y_i, i = 1 \dots N\}$ zaobserwowaną realizację szeregu czasowego.

Ponieważ $\{\epsilon_i, i = 1, 2, \dots N\}$ jest ciągiem niezależnych liczb losowych o rozkładzie normalnym i wariancji σ_i (Anderson, Moore, 1979), więc funkcja wiarygodności wyrazi się:

$$L(y,\theta) = (2\pi)^{-N/2} \cdot (\sigma_1 \sigma_2 \dots \sigma_N)^{-1/2} \cdot \exp\left\{-\frac{1}{2}\sum_{i=1}^{N}\frac{\epsilon_i^2}{\sigma_i}\right\}. \tag{4.20}$$

W praktycznych obliczeniach zamiast maksymalizacji funkcji wiarygodności (4.20) przeprowadza się minimalizację funkcji $l(y,\theta) = -\ln L(y,\theta)$. Z dokładnością do stałej wyraża się ona zależnością:

$$l(y,\theta) = 1/2\sum_{i=1}^{N}\ln\sigma_i + 1/2\sum_{i=1}^{N}\epsilon_i^2/\sigma_i, \tag{4.21}$$

co po uwzględnieniu (4.10) daje:

$$l(y,\theta) = N/2\ln\rho + 1/2\sum_{i=1}^{N}\ln f_i + 1/(2\rho)\sum_{i=1}^{N}\epsilon_i^2/f_i. \tag{4.22}$$

Obliczając $(\partial/\partial\rho)l(y,\theta)$ i przyrównując wynik do zera otrzymuje się:

$$\hat{\rho} = N^{-1}\sum_{=1}^{N}\epsilon_i^2/f_i. \tag{4.23}$$

Minimalizacja funkcji (4.22) względem pozostałych zmiennych jest równoważna minimalizacji funkcji:

$$L^*(y,\theta) = N^{-1}(f_1 f_2 \dots . f_N)^{1/N}\sum_{i=1}^{N}\epsilon_i^2/f_i, \tag{4.24}$$

gdzie f_i są określone wzorem (4.9).

Minimalizacji funkcji $L^*(y,\theta)$ można dokonać za pomocą metod newtonowskich korzystając z gradientu i hesjanu wyznaczonych analitycznie lub numerycznie. Należy podkreślić, że jest to problem obliczeniowo złożony oraz, że nie są znane analityczne formuły dla gradientu i hesjanu.

5. Estymacja modelu ciągłego na podstawie modelu ARMA

W niniejszym punkcie przedstawimy wzajemne związki zachodzące pomiędzy modelem ARMA a modelem próbkowanego ciągłego procesu stochastycznego oraz wskażemy na możliwość wykorzystania istniejących algorytmów obliczania funkcji wiarygodności i jej gradientu do konstrukcji efektywnych numerycznie algorytmów estymacji parametrów modelu ciągłego.

5.1. Model ARMA jako model próbkowanego procesu ciągłego

Błachuta i Polański (1987, 1990) pokazali, że modelem kowariancyjnie równoważnym modelowi (3.5)-(3.6) jest:

$$x^*_{i+1} = Fx^*_i + hv_i, \tag{5.1}$$

$$y_i = d'x^*_i + v_i, \tag{5.2}$$

gdzie:

$$\operatorname{cov}(x^*) = Q^*\sigma, \ \mathrm{E}(v_iv_j) = \sigma\delta_{ij}, \ \mathrm{E}(v_i) = 0, \tag{5.3}$$

$$Q^* = Q^0 - P, \tag{5.4}$$

$$h = \frac{FPd}{1 + d'Pd}, \tag{5.5}$$

$$\sigma = \rho(1 + d'Pd), \tag{5.6}$$

zaś P jest *dowolnym* symetrycznym rozwiązaniem algebraicznego równania Riccatiego:

$$P = W^0 + F(P - \frac{Pdd'P}{1 + d'Pd})F'. \tag{5.7}$$

Model (5.1)-(5.2) bywa w literaturze błędnie nazywany modelem innowacyjnym. Wynika to z faktu, iż w przypadku wybrania *dodatnio określonego* rozwiązania równania (5.7) reprezentacja (5.1)-(5.2) z parametrami określonymi poprzez (5.5)-(5.6) jest identyczna z ustalonym filtrem Kalmana (4.16), w którym warunek początkowy jest wektorem zdeterminowanym. Jednakże ustalony filtr Kalmana jest jedynie asymptotycznie optymalny, zaś skonstruowany na jego podstawie proces jest także jedynie asymptotycznie równoważny procesowi (3.5)-(3.6).

Model zredukowany (5.1)-(5.2) można również zapisać w postaci odwrotnej:

$$x^*_{i+1} = F^*x^*_i + hy_i, \tag{5.8}$$

$$v_i = d'x^*_i - y_i, \tag{5.9}$$

gdzie:

$$F^* = F - hd'. \tag{5.10}$$

Ważną cechą modelu zredukowanego jest to, iż implikuje on, jako model obserwacji dyskretnych, model ARMA:

$$\varphi(z^{-1})y_i = \vartheta(z^{-1})v_i, \tag{5.11}$$

przy czym

$$\varphi(z^{-1}) = 1 - \varphi_1 z^{-1} - \ldots - \varphi_n z^{-n} = \det(I - Fz^{-1}), \tag{5.12}$$

$$\vartheta(z^{-1}) = 1 - \vartheta_1 z^{-1} - \ldots - \vartheta_n z^{-n} = \det(I - F^* z^{-1}). \tag{5.13}$$

Postać kanoniczną obserwatorową układu (5.1)-(5.2) można wyrazić poprzez współczynniki wielomianów (5.12)-(5.13) (np. Shea, 1989):

$$F = \begin{bmatrix} \varphi_1 & 1 & 0 & . & . & . & 0 \\ \varphi_2 & 0 & 1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ \varphi_{n-1} & & & & & & 1 \\ \varphi_n & 0 & . & . & . & . & 0 \end{bmatrix}, \; h = \begin{bmatrix} \varphi_1 & - & \vartheta_1 \\ \varphi_2 & - & \vartheta_2 \\ & . & \\ & \vdots & \\ \varphi_n & - & \vartheta_n \end{bmatrix}, \; d = \begin{bmatrix} 1 \\ 0 \\ . \\ \vdots \\ 0 \end{bmatrix}. \tag{5.14}$$

Model zredukowany (5.1)-(5.2) związany z dodatnio określonym rozwiązaniem równania Riccatiego (5.7), dla którego wszystkie pierwiastki wielomianu $z^n\vartheta(z^{-1})$ leżą wewnątrz okręgu jednostkowego na płaszczyźnie z, jest tzw. modelem odwracalnym cechującym się tym, że równanie (5.8) lub odpowiadające mu

$$\vartheta(z^{-1})v_i = \varphi(z^{-1})y_i \tag{5.15}$$

są równaniami stabilnymi.

5.2. Estymacja modelu ARMA

Dla modelu zredukowanego (5.1)-(5.2) równania opisujące predyktor jednokrokowy upraszczają się (Gardner, Harvey, Phillips, 1980). Oznaczmy

$$\operatorname{cov}(x^*) = Q^*\sigma, \tag{5.16}$$

przy czym Q^* jest rozwiązaniem algebraicznego równania Lapunowa

$$Q^* = FQ^*F' + hh'. \tag{5.17}$$

Oznaczmy przez $\hat{x}^*_{i|i-1}$ liniową ocenę stanu zapewniającą minimum błędu średniokwadratowego stanu x^*_i przy danych pomiarach do chwili $i-1$, zaś kowariancję błędu oceny stanu $\tilde{x}^*_i = x^*_i - \hat{x}^*_{i|i-1}$ przez

$$\operatorname{cov}(\tilde{x}^*_i) = \Sigma_{i|i-1}\sigma. \tag{5.18}$$

Wówczas jednokrokowy predyktor stanu jest określony poprzez reprezentację innowacyjną:

$$\hat{x}^*_{i+1|i} = F\hat{x}^*_{i|i-1} + (h + k_i)\epsilon_i, \hat{x}^*_{0|-1} = 0, \tag{5.19}$$

$$k_i = \frac{F^* \Sigma_{i|i-1} d}{1 + d' \Sigma_{i|i-1} d}, \tag{5.20}$$

$$\Sigma_{i+1|i} = F^*(\Sigma_{i|i-1} - \frac{\Sigma_{i|i-1} dd' \Sigma_{i|i-1}}{1 + d' \Sigma_{i|i-1} d}) F^{*\prime}, \Sigma_{0|-1} Q^*. \tag{5.21}$$

Dla $i = 1, 2 \ldots n$, błędy predykcji jednokrokowej ϵ_i oraz ich wariancja var $(\epsilon_i) = sigma f_i$ są określone zależnościami:

$$\epsilon_i = z_i - d'\hat{x}^*_{i|i-1}, \tag{5.22}$$

$$f_i = 1 + d' \Sigma_{i|i-1} d. \tag{5.23}$$

W przypadku modelu odwracalnego przy $i \to \infty$ zachodzi $\Sigma_{i|i-1} \to 0$, $k_i \to 0$ i równania predyktora przyjmą postać asymptotyczną taką samą jak w (4.16).

$$\hat{x}^*_{i+1|i} = F^* \hat{x}^*_{i|i-1} + h y_i. \tag{5.24}$$

Metody identyfikacji i estymacji parametrów modelu ARMA są znakomicie rozwinięte (np. Hannan, 1988; Hannan, Kavalieris, 1983; Hannan, Rissanen, 1982). Obszerny przegląd metod identyfikacji rzędu modelu ARMA znajduje się w pracy Błachuty (1994). Dlatego też nasuwa się następująca koncepcja estymacji modelu ciągłego (Söderström 1984, 1991):

- estymacja wektora parametrów θ^* modelu ARMA,
- odtworzenie wektora parametrów θ modelu ciągłego.

Biorąc pod uwagę fakt, że model ARMA powstał z dyskretnych obserwacji systemu ciągłego musimy, zgodnie z (5.11)-(5.12), przyjąć równe stopnie wielomianów MA i AR.

Jest zatem:

$$\theta^* = [\varphi_1, \varphi_2 \ldots \varphi_n, \ \vartheta_1, \vartheta_2 \ldots \vartheta_n, \sigma]. \tag{5.25}$$

Podobnie jak uprzednio oceny największej wiarygodności wektora parametrów θ^* dostaniemy minimalizując funkcję:

$$L^*(y, \theta^*) = N^{-1}(f_1 f_2 \ldots f_N)^{1/N} \sum_{i=1}^{N} \epsilon_i^2 / f_i, \tag{5.26}$$

przy czym oceną σ jest:

$$\hat{\sigma} = N^{-1} \sum_{=1}^{N} \epsilon_i^2 / f_i. \tag{5.27}$$

Zależność (5.26) jest ogólną postacią funkcji wiarygodności dla modelu ARMA, niezależną od metody jej obliczenia. Praktyczne algorytmy korzystają na ogół z innych równań, których elementy są mniej pracochłonne od rekurencji równania (5.21), np. Mélard (1984). Istotnym problemem jest metodyka minimalizacji (5.26) i składające się na nią elementy: wyznaczanie wstępnych ocen parametrów oraz obliczanie gradientu i hesjanu funkcji wiarygodności względem estymowanych parametrów. Zagadnienia te, dobrze obecnie opracowane (np. Kohn, Ansley, 1982; Burshtein, 1993), wykraczają poza zakres niniejszego opracowania.

Osobnym problemem jest zastosowanie wysokiej częstotliwości próbkowania. Model ARMA w wersji (5.11)-(5.14) staje się wówczas praktycznie nieprzydatny. Metoda wymaga przeformułowania w kategoriach operatora δ. Zagadnienie to również wykracza poza zakres tego opracowania.

5.3. Wykorzystanie modelu ARMA do estymacji modelu ciągłego

Jeśli przez $\theta^* = \mathcal{F}(\theta)$ oznaczyć operację przekształcania parametrów opisu ciągłego na parametry opisu dyskretnego, to przez odtworzenie wektora parametrów θ modelu ciągłego należy rozumieć procedurę odwrotną: $\theta = \mathcal{F}^{-1}(\theta^*)$ lub jej aproksymację, gdy dla danego wektora θ^* procedury $\theta = \mathcal{F}^{-1}(\theta^*)$ przeprowadzić się nie da. Na operację $\mathcal{F}(\theta)$ składa się:

- utworzenie postaci kanonicznej (2.5) układu ciągłego (2.1)-(2.2),
- obliczenie macierzy $\boldsymbol{F}$ na podstawie (3.7),
- obliczenie macierzy $\boldsymbol{W}$ zgodnie z (3.8),
- wyznaczenie postaci zredukowanej (5.1) na podstawie (5.4)-(5.7),
- sprowadzenie (5.1) do postaci kanonicznej (5.14).

Część operacji odwrotnej polegająca symbolicznie na obliczeniu

$$A = \Delta^{-1} \ln F \tag{5.28}$$

jest prosta i polega w istocie na relacji między zerami α_i wielomianu $A(s)$ oraz zerami α_i^* wielomianu $z^n\varphi(z^{-1})$:

$$\alpha_i = \Delta^{-1} \ln \alpha_i^*, \ (i = 1, 2 \ldots n). \tag{5.29}$$

Jeśli α_i^* jest rzeczywiste i dodatnie, nie pojawia się żaden problem. Jeśli jest ono rzeczywiste i ujemne, problem nie ma rozwiązania. Jeśli α_i^* jest zespolone, problem jest niejednoznaczny. Można wówczas przyjąć np. rozwiązanie o najmniejszej części rzeczywistej. Przegląd algorytmów numerycznych do obliczenia (5.28) można znaleźć w pracy (Sinha, Rao 1991).

Poważniejsze problemy nastręcza obliczenie elementów wielomianu $C(s)$ oraz wariancji zakłócenia pomiarów ρ. Przegląd algorytmów numerycznych do wyznaczania operacji odwrotnej $\theta = \mathcal{F}^{-1}(\theta^*)$, dla przypadku $\rho = 0$, znajdują się w pracach Söderströma (1984, 1991).

Koncepcyjnie najprostszy z nich, ważny również dla $\rho > 0$, polega na numerycznym odwróceniu operacji $\mathcal{F}$. Wartość θ znajduje się wówczas w wyniku minimalizacji:

$$\theta = \arg\min \|\theta^* - \mathcal{F}(\theta)\|^2. \tag{5.30}$$

W przypadku, gdy $\min \|\theta^* - \mathcal{F}(\theta)\| = 0$, można powiedzieć, że dla danego wektora θ^* istnieje system ciągły, którego wersja dyskretna odpowiada dokładnie zidentyfikowanemu systemowi dyskretnemu. Ponieważ operacja $\mathcal{F}(\theta)$ nakłada na wektor θ^* pewne więzy, które nie są brane pod uwagę przy estymacji $\hat{\theta}^*$, prawdopodobieństwo istnienia takiego systemu ciągłego jest znikome. Uwaga ta dotyczy przydatności wszystkich algorytmów obliczania $\theta = \mathcal{F}^{-1}(\theta^*)$ do celów identyfikacji θ; przy tym poza algorytmem (5.30) należy liczyć się z rozwiązaniami bardzo dalekimi od poprawnych lub wręcz z brakiem rozwiązania.

Jednakże wyjątkowo dobrze opracowane i niezwykle efektywne numerycznie metody wyznaczania funkcji wiarygodności modelu ARMA skłaniają do wykorzystania ich do estymacji parametrów modelu ciągłego, nie korzystając z (5.30). Można to zrobić zauważając, że:

$$\theta = \arg\min L^*[y, f(\theta)]. \tag{5.31}$$

Gradient funkcji $L^*[y, f(\theta)]$

$$\nabla_\theta L^*[y, f(\theta)] = \frac{\partial f(\theta)}{\partial \theta} \nabla_{\theta^*} L^*[y, \theta^*] \tag{5.32}$$

można wyznaczyć znając wrażliwość $\partial f(\theta)/\partial\theta$ oraz gradient w przestrzeni parametrów modelu ARMA. Pierwszy czynnik w zależności (5.32) może być obliczony numerycznie, zaś drugi, łącznie z (5.31) za pomocą algorytmów podanych np. w pracach (Kohn, Ansley, 1982), (Burshtein, 1993). W przypadku dużego wymiaru y, co w procesie ciągłym jest oczywiste, stanowi to efektywną obliczeniowo alternatywę w stosunku do numerycznego wyznaczania gradientu $\nabla_\theta L^*[y, \theta]$, gdzie $L^*[y, \theta]$ jest określone przez (4.24).

6. Modele wektorowe

Naturalnym uogólnieniem procesów skalarnych są procesy wektorowe, które dla x_t zawierającego m zmiennych można opisać układem równań:

$$dx_t = Ax_t dt + G d\xi_t, \tag{6.1}$$

$$z_t = D'x_t. \tag{6.2}$$

Macierze G i D są obecnie $n \times m$ wymiarowe, zaś ξ_t jest wektorem m-wymiarowym, którego elementy są niezależnymi procesami Wienera.

Podobnie jak dla procesu skalarnego równanie pomiaru próbkowanego przyjmie postać:

$$y_i = D'x_i + r_i, \tag{6.3}$$

gdzie r_i jest dyskretnym białym szumem gaussowskim, tzn. $E[r_i r_j'] = \mathcal{R}\delta_{ij}$, przy czym $\mathcal{R}$ jest macierzą $m \times m$ wymiarową.

Proces y_i jest opisany za pomocą układu dyskretnych równań stochastycznych:

$$x_{i+1} = Fx_i + w_i, \tag{6.4}$$

$$y_i = D'x_i + r_i, \tag{6.5}$$

gdzie

$$F = e^{A\Delta}, \tag{6.6}$$

zaś w_i jest wektorowym białym szumem gaussowskim o macierzy kowariancji W:

$$W = \int_0^\Delta e^{As}GG'e^{A's}ds. \tag{6.7}$$

Proces innowacji jest opisany równaniami:

$$e_i = y_i - D'\hat{x}_{i|i-1}, \tag{6.8}$$

$$\hat{x}_{i+1|i} = F\hat{x}_{i|i-1} + K_i e_i,\ \hat{x}_{0|-1} = 0, \tag{6.9}$$

$$H_i = FS_{i|i-1}DS_i^{-1}, \tag{6.10}$$

$$S_{i+1|i} = W + F(S_{i|i-1} - S_{i|i-1}DS_i^{-1}D'S_{i|i-1})F',\ S_{0|-1} = Q, \tag{6.11}$$

$$S_i = \mathcal{R} + D'S_{i|i-1}D. \tag{6.12}$$

Funkcja wiarygodności będzie miała postać:

$$L(y,\theta) = (2\pi)^{-Nm/2} \cdot [\det(S_1 S_2 \ldots S_N)]^{-1/2} \cdot \exp\left\{-\frac{1}{2}\sum_{i=1}^{N} e_i S_i^{-1} e_i\right\}. \tag{6.13}$$

Ponieważ obecnie błędy pomiarowe tworzą wektor o macierzy wariancji $\mathcal{R}$, proste obliczenie jej oceny, tak jak we wzorze (5.5), nie jest możliwe i musi być ona traktowana tak samo jak inne parametry, będąc jednym z argumentów procedury minimalizacji funkcji

$l(y,\theta) = -\ln L(y,\theta)$. Jak się okazuje, problemy ze zbieżnością tych algorytmów względem macierzy $\mathcal{R}$ są poważne. Dla zilustrowania problemu Maine i Iliff (1981) podają prosty przykład układu skalarnego:

$$x_{i+1} = x_i, x_0 = 0, \tag{6.14}$$

$$y_i = x_i + r_i, \tag{6.15}$$

$$Er_i^2 = \rho = \lambda^2. \tag{6.16}$$

Minimalizowana funkcja $f(\lambda) = l(y,\lambda)$ ma postać:

$$f(\lambda) = 1/2N \ln \lambda^2 + \frac{1}{2\lambda^2}\sum_{i=1}^{N} y_i^2. \tag{6.17}$$

Oznaczając

$$r = N^{-1}\sum_{=1}^{N} y_i^2 \tag{6.18}$$

mamy:

$$\nabla_\lambda f(\lambda) = -Nr\lambda^{-3} + N\lambda^{-1}, \tag{6.19}$$

$$\nabla_\lambda^2 f(\lambda) = 3Nr\lambda^{-4} - N\lambda^{-2}. \tag{6.20}$$

Z równania (6.8) jest oczywiste, że $\hat{\rho} = \hat{\lambda}^2 = r$. Jednakże stosując do przykładu iteracje Newtona-Raphsona

$$\theta_{k+1} = \theta_k - [\nabla_\theta^2 f(\theta)]^{-1}[\nabla_\theta f(\theta)], \tag{6.21}$$

dostaje się:

$$\lambda_{k+1} = \lambda_k \frac{4r - 2\lambda_k^2}{3r - \lambda_k^2}. \tag{6.22}$$

Zbieżność do $\lambda^2 = r$, będącego punktem równowagi równania (6.22), ma miejsce dla $0 < \lambda_k^2 < 2r$, przy czym dla małych wartości λ_k jest ona bardzo wolna.

Reasumując, sformułowanie naturalne może być stosowane w trzech sytuacjach:

- gdy pomiary są dokładne,
- gdy wartość wariancji $\mathcal{R}$ szumu pomiarowego jest znana,
- gdy dysponujemy dobrymi wstępnymi ocenami $\hat{\mathcal{R}}$.

W sytuacji, gdy powyższe warunki nie są spełnione, Maine i Iliff (1981) proponują następujące rozwiązanie. Zauważmy, że przy $i \to \infty$ zachodzi $\boldsymbol{S}_{i|i-1} \to \boldsymbol{S}$, gdzie $\boldsymbol{S}$ jest dodatnio określonym, symetrycznym rozwiązaniem algebraicznego równania Riccatiego:

$$\boldsymbol{S} = \boldsymbol{W} + \boldsymbol{F}(\boldsymbol{S} - \boldsymbol{S}\boldsymbol{D}\mathcal{S}^{-1}\boldsymbol{D}'\boldsymbol{S})\boldsymbol{F}', \tag{6.23}$$

a także $\mathcal{S}_i \to \mathcal{S}$, gdzie

$$\mathcal{S} = \mathcal{R} + \boldsymbol{D}'\boldsymbol{S}\boldsymbol{D} \tag{6.24}$$

oraz $H_i \rightarrow H$, gdzie

$$H = FSDS^{-1}. \tag{6.25}$$

Równania predyktora przyjmą postać asymptotyczną:

$$\hat{x}_{i+1|i} = F\hat{x}_{i|i-1} + He_i, \tag{6.26}$$

$$y_i = D'\hat{x}_{i|i-1} + e_i. \tag{6.27}$$

Maine i Iliff nazywają (6.26)-(6.27) reprezentacją innowacyjną i przyjmują $\hat{x}_{0|-1} = 0$. Traktując e_i jako innowacje o niezmiennej w czasie macierzy kowariancji $\mathcal{S}$ - co jest asymptotycznie poprawne - funkcję wiarygodności otrzymujemy w postaci:

$$L(y, \theta) = [(2\pi)^m \cdot \det \mathcal{S}]^{-N/2} \cdot \exp\left\{-\frac{1}{2}\sum_{i=1}^{N} e_i'\mathcal{S}^{-1}e_i\right\}, \tag{6.28}$$

zaś ujemny logarytm naturalny z dokładnością do stałej:

$$l(y, \theta) = \frac{1}{2}\sum_{i=1}^{N} e_i'\mathcal{S}^{-1}e_i + N/2\ln(\det \mathcal{S}). \tag{6.29}$$

Niewątpliwą zaletą tego sformułowania jest prostota estymacji wariancji $\mathcal{S}$

$$\hat{\mathcal{S}} = N^{-1}\sum_{=1}^{N} e_i e_i'. \tag{6.30}$$

Główną zaletą sformułowania pierwotnego jest bezpośrednia estymacja parametrów układu ciągłego, wadą są problemy ze zbieżnością względem wariancji $\mathcal{R}$. Istotą sformułowania mieszanego jest dążenie do zachowania zalet dwu poprzednich sformułowań i eliminacja ich wad. W sformułowaniu mieszanym uwzględnimy zatem parametry macierzy A, G oraz wariancję $\mathcal{S}$.

Należy podkreślić, że ponieważ obecnie - zgodnie z (6.25) - H jest funkcją $\mathcal{S}$, a co za tym idzie $e_i = e_i(\mathcal{S})$, więc (6.30) nie jest precyzyjną oceną największej wiarygodności wariancji $\mathcal{S}$. Istotnie, różniczkując $l(y, \theta^*)$ względem $\mathcal{S}$ otrzymujemy:

$$\begin{aligned} 2N^{-1}\mathcal{S}\nabla_{\mathcal{S}} l(y, \theta)\mathcal{S} &= \mathcal{S} - N^{-1}\sum_{i=1}^{N} e_i e_i' \\ &+ \ 2N^{-1}\mathcal{S}\left\{\sum_{i=1}^{N} e_i'\mathcal{S}_i^{-1}(\nabla_{\mathcal{S}}\hat{y}_i)\right\}\mathcal{S}, \end{aligned} \tag{6.31}$$

przy czym $\hat{y}_i = D'\hat{x}_{i|i-1}$.

Wyrażenie w drugim wierszu jest zmienną losową o zerowej wartości oczekiwanej i wariancji zmierzającej do zera przy N zmierzającym do nieskończoności. Przy dostatecznie dużych wartościach N staje się ono pomijalnie małe. Dlatego też estymator (6.30) jest tylko asymptotycznie estymatorem ML.

Szybka zbieżność tej metody jest okupiona dalszymi przybliżeniami, tym jednak lepszymi, im krótszy jest okres próbkowania. Przepiszmy równanie (6.23) w nieco zmienionej formie:

$$F^{-1}SF'^{-1} = F^{-1}WF'^{-1} + S - SDS^{-1}D'S. \tag{6.32}$$

Przyjmijmy przybliżenia

$$F^{-1} \cong I - A\Delta, \tag{6.33}$$

$$W \cong \Delta gg'. \tag{6.34}$$

Wówczas, pomijając człony rzędu Δ^2, równanie (6.23) przekształci się do:

$$AS + SA' - SDS^{-1}D'S + gg' = 0. \tag{6.35}$$

Równanie (6.35) ma postać równania Riccatiego dla czasu ciągłego i może być rozwiązane za pomocą standardowych metod. Macierz H, przy $\Delta \to 0$, wyrazi się prostym wzorem

$$H = SDS^{-1}. \tag{6.36}$$

W dalszym ciągu zakładamy, że jako S przyjmujemy dodatnio określone rozwiązanie (6.26), tak że model dyskretny jest odwracalny. Odwracalność jest konieczna dla stabilności równań odwrotnych:

$$\hat{x}_{i+1|i} = (F - HD')\hat{x}_{i|i-1} + Hy_i, \tag{6.37}$$

$$e_i = -D'\hat{x}_{i|i-1} + y_i, \tag{6.38}$$

służących do wyznaczania e_i.

Maine i Iliff (1981) prezentują skuteczne algorytmy minimalizacji (6.29) na podstawie metod newtonowskich z analitycznie obliczonym gradientem i hesjanem. Metoda została zaimplementowana w programie o nazwie MMLE3, używanym przez NASA do badań ruchu obiektów latających w warunkach turbulencji. Wyniki tego algorytmu mogą zostać potraktowane jako dobre oceny wstępne dla dokładnego estymatora największej wiarygodności, otrzymywanego poprzez maksymalizację (6.13). Teoretycznie jest możliwe usprawnienie wyznaczania estymatora dokładnego poprzez wprowadzenie wektorowego modelu VARMA, zgodnie z metodologią przedstawioną w p.4. Efektywny numerycznie algorytm obliczania funkcji wiarygodności dla modelu VARMA istnieje (Shea, 1989). Brak jednak dotychczas algorytmu wyznaczania gradientu.

7. Podsumowanie

W artykule przedstawiono metody estymacji parametrów modelu ciągłego procesu stochastycznego na podstawie pomiarów dyskretnych w czasie. Przedstawiono algorytm obliczania funkcji wiarygodności w oparciu o bazowy model dyskretny. Następnie przedstawiono kowariancyjnie równoważną realizację modelu dyskretnego, prowadzącą do modelu ARMA. Pokazano, w jaki sposób model ARMA można wykorzystać do konstrukcji efektywnego numerycznie algorytmu estymacji parametrów modelu ciągłego. W końcu przedstawiono problemy estymacji parametrów ciągłych modeli wektorowych oraz podano metody ich rozwiązania.

8. Podziękowanie

Autor wyraża wdzięczność p. Profesorowi Mieczysławowi Brdysiowi z The University of Birmingham za stworzenie znakomitych warunków do pracy w School of Electronic & Electrical Engineering.

LITERATURA

1. Anderson B. D. O, and J.B Moore (1979), Optimal Filtering, Prentice-Hall

2. Åström K. J. (1970), Introduction to Stochastic Control Theory, Academic

3. Bergstrom A. R. ed. (1976), Statistical Inference in Continuous Time Economic Models, North-Holland

4. Błachuta M. (1994), Metody identyfikacji stacjonarnych i niestacjonarnych modeli ARMA, Zeszyty Naukowe Politechniki Śląskiej, s. Automatyka, przyjęte do druku

5. Błachuta M. and A. Polański (1987), On time invariant representations of discrete random processes, IEEE Trans. on Auto. Control, vol. AC-32, pp. 1125-1127

6. Błachuta M., Polański A. (1990), Reprezentacje dyskretnych procesów Gaussa- Markowa a symetryczne rozwiązania równania Riccatiego, Archiwum Automatyki i Telemechaniki, vol. 35, pp. 177-189

7. Burshtein D. (1993), An efficient algorithm for calculating the likelihood and likelihood gradient of ARMA models, IEEE Trans. on Auto. Control, vol. AC-38, pp. 336-340

8. Chen H. F. and L. Guo (1991), Identification and Stochastic Adaptive Control, Birkhauser

9. Gardner G., A. C. Harvey and G. D. A. Phillips (1980), Algorithm AS 154. An algorithm for exact maximum likelihood estimation of autoregressive-moving average models by means of Kalman filtering, Appl. Statist., vol. 29, pp. 311-322

10. Gikhman I. I. and A. V. Skorokhod (1969), Introduction to the theory of random processes, Saunders

11. Gikhman I.I and A. V. Skorokhod (1972), Stochastic differential equations, Springer

12. Hannan E. J. (1988), The estimation of the order of an ARMA process, Ann. Statist, vol. 8, pp. 1071-1081

13. Hannan E. J. and L. Kavalieris (1983), Linear estimation of ARMA processes, Automatica, vol. 19, pp. 447-448

14. Hannan E. J. and J. Rissanen (1982), Recursive estimation of mixed autoregressive-moving average order, Biometrika, vol. 69, pp. 81-94

15. Kohn R. and C. F. Ansley (1982), Computing the likelihood and their derivatives for a Gaussian ARMA model, J. Statist. Computn Simuln, vol. 15, pp. 229-263

16. Kushner H. (1971), Introduction to Stochastic Control, Holt, Rinehart and Winston

17. Kwakernaak H. and R. Sivan (1972), Linear Optimal Control Systems, Wiley

18. Maine R. E. and K. Iliff (1981), Formulation and implementation of a practical algorithm for parameter estimation with process and measurement noise, SIAM J. Appl. Math., vol. 41, pp. 558-579

19. Mélard G. (1984), A fast algorithm for the exact likelihood of autoregresive-moving average models, Algorithm AS 197, Applied Statistics, vol. 33, pp. 104-113

20. Middleton R. H. and G. C. Goodwin (1990), Digital Control and Estimation - A Unified Approach, Prentice Hall

21. Moore J. B. (1988), Convergence of continuous time stochastic ELS parameter estimation, Stochastic processes and their applications, vol. 27, pp. 195-215

22. Van Schuppen J. H. (1983), Convergence results for continuous time adaptive stochastic filtering algorithms, J. of Math. Analysis and Applic., vol. 96, pp. 209-225

23. Saha D. C and G. P. Rao (1983), Identification of Continuous Dynamical Systems - the Poisson Moment Functional Approach, Springer

24. Sargan J. D. (1974), Some discrete approximations to continuous time stochastic models, J. Roy. Statist. Soc., ser. B, vol. 36, pp. 74-90

25. Shea B. L. (1989), Algorithm AS 242, The exact likelihood of a vector autoregresive moving average model, Appl. Statist., vol. 38, pp. 161-204

26. Sinha N. K. and G. P. Rao, ed. (1991), Identification of Continuous-Time Systems, Methodology and Computer Implementation, Kluwer

27. Söderström T., (1984): On computing continuous time counterparts to ARMA models, Raport UPTEC 84 100 R, Uppsala University

28. Söderström T., (1991): On computing stochastic continuous-time models from ARMA models, Int. J. Control, vol. 53, pp. 1311-1326

29. Unbehauen H. and G. P. Rao (1987), Identification of Continuous Systems, North-Holland

30. Wymer C. R. (1972), Econometric estimation of stochastic differential equation systems, Econometrica, vol. 40, pp. 565-577

Recenzent: Doc. dr hab. inż. Zbigniew Nahorski
IBS PAN Warszawa

Wpłynęło do Redakcji dnia 15.05.1994

Abstract

In the paper Maximum Likelihood parameter estimation algorithms based on discrete-time samples of continuous-time stochastic processes with a rational spectral density are addressed. For the considered class of processes stochastic state-space equations are formulated, their solutions are defined and discrete-time models are evaluated. Formulation of a continuous-time parameter estimation problem based on these discrete-time models is referred to as a natural formulation. An algorithm for calculation of the likelihood function which bases on the basic discrete-time model and Kalman filtering is presented. Then a covariance equivalent realization which leads to the ARMA model is introduced. Difficulties connected with using an estimated ARMA model for determination of continuous-time parameters are highlighted. The ARMA model is then indirectly used in a computationally efficient method of continuous-time parameters estimation. Finally, estimation problems in multivariate continuous-time models are stated and solved.