LMD konf. darbai, 222–227
1997, Vilnius

Tam tikrų laiko eilučių modelių įvertinimas, naudojant netiesinio mažiausiųjų kvadratų uždavinio metodologiją

V. Šimonytė, V. Slivinskas (MII)

1. ĮVADAS

Vienas iš dažnai naudojamų modelių signalų apdorojimo srityje, sistemų identifikacijoje bei laiko eilučių analizėje yra kelių realiųjų sinusoidžių, stebimų triukšme, sumos modelis (žr., pvz., [3]). Jo pagrindiniai parametrai yra dažniai. Kitas populiarus modelis yra kelių silpstančių eksponenčių užtriukšminta suma (žr., pvz., [4]). Pagrindiniai tokio modelio parametrai yra gesimo koeficientai. Šių dviejų modelių kombinacija yra užtriukšminta silpstančių sinusoidžių suma (žr., pvz., [5], [6]). Jo pagrindiniai koeficientai yra gesimo koeficientai ir dažniai. Šiame darbe mes nagrinėjame modelių klasę, kuri apibendrina minėtųjų trijų modelių klases. Tai vadinamieji kartotiniai eksponentiniai-sinusiniai modeliai. Šios klasės modelis yra apibrėžiamas kaip kelių kartotinių silpstančių eksponenčių bei kelių kartotinių silpstančių sinusoidžių užtriukšminta suma. Šiame darbe nagrinėjamas kartotinio eksponentinio-sinusinio modelio parametrų įvertinimo uždavinys. Straipsnio struktūra yra tokia: nagrinėjamojo modelio aprašymas pateikiamas antrajame skyrelyje. Trečiajame skyrelyje nagrinėjamas netiesinis mažiausiųjų kvadratų uždavinys su atsiskiriančiais kintamaisiais. Ketvirtajame skyrelyje parodoma, kad kartotinio eksponentinio-sinusinio modelio parametrų įvertinimo uždavinys gali būti sprendžiamas kaip netiesinis mažiausiųjų kvadratų uždavinys su atsiskiriančiais kintamaisiais. Problemos, iškylančios taikant Levenbergo metodą nagrinėjamojo modelio „netiesinių" parametrų įvertinimui, yra nagrinėjamos penktajame skyrelyje.

2. SIGNALO MODELIS

Šiame skyrelyje mes pateikiame signalo modelį. Nagrinėjamos trys modelio formos: skaliarinė, vektorinė ir forma, kurioje naudojama vadinamoji bazinių signalų matrica.

2.1 *Skaliarinė forma.* Nagrinėkime modelį

$$x(t) = q(t) + e(t), \qquad t = 0, 1, \ldots, N-1,$$

čia $\{e(t)\}$ yra balto Gauso triukšmo seka, o $\{q(t)\}$ yra neužtriukšminto signalo reikšmių seka. Tarkime, kad neužtriukšmintas signalas q yra L signalų suma

$$q(t) = q_1(t) + \ldots + q_L(t), \tag{1}$$

iš kurių pirmieji L_1 yra kartotinės silpstančios eksponentės, t.y. tokio pavidalo signalai

$$q_1(t) = e^{\lambda_{1,l}t}(A_{l1} + A_{l2}(t) + \ldots + A_{lK_l}t^{K_l-1}), \tag{2}$$

čia $\lambda_{1,l} < 0$, $K_l \geqslant 1$, $A_{lk} \in \mathbf{R}$, $k = 1, \ldots, K_l$; $l = 1, \ldots, L_1$, o likusieji L_2 yra kartotinės silpstančios sinusoidės

$$q_{L_1+l}(t) = e^{\lambda_{2,l}t} \sum_{k=1}^{K_{L_1+l}} A_{L_1+l,k}t^{k-1} \sin(\omega_l t + \varphi_{L_1+l,k}), \tag{3}$$

čia $\lambda_{2,l} < 0$, $K_{L_1+l} \geqslant 1$, $A_{L_1+l,k} \in \mathbf{R}$, $\omega_l > 0$, $\varphi_{L_1+l,k} \in [-\pi, \pi[$, $k = 1, \ldots, K_{L_1+l}$; $l = 1, \ldots, L_2$ $(L = L_1 + L_2)$.

2.2 *Vektorinė forma.* Nagrinėjamąjį modelį galima užrašyti tokiu vektoriniu pavidalu:

$$\mathbf{x} = \mathbf{q} + \mathbf{e}$$

$$\mathbf{q} = \mathbf{q}_1 + \ldots + \mathbf{q}_L$$

čia

$$\mathbf{x} = (x(0), x(1), \ldots, x(N-1))^T,$$

$$\mathbf{q}_l = (q_l(0), q_l(1), \ldots, q_l(N-1))^T.$$

Vektoriai $\mathbf{e}$ ir $\mathbf{q}$ yra apibrėžiami analogiškai.

2.3 *Bazinių signalų matrica.* Nagrinėjamąjį modelį galima pateikti ir pavidalu, naudojančiu bazinių signalų matricą. Raide Φ pažymėkime $N \times (K_1 + \ldots + K_{L_1} + 2(K_{L_1+1} + \ldots + K_L))$ dydžio matricą, sudarytą iš dviejų matricų Φ_1 ir Φ_2: $\Phi = [\Phi_1\ \Phi_2]$. Φ_1 yra $N \times (K_1 + \ldots + K_{L_1})$ matrica su tokiais stulpeliais

$$\Phi_1 = [d^e_{11}, d^e_{12}, \ldots, d^e_{1K_1}, \ldots, d^e_{L_11}, d^e_{L_12}, \ldots, d^e_{L_1K_{L_1}}],$$

čia

$$d^e_{lk} = [\delta_{k1}, e^{\lambda_{1,l}}, 2^{k-1}e^{2\lambda_{1,l}}, \ldots, (N-1)^{k-1}e^{(N-1)\lambda_{1,l}}]^T,\ k = 1, \ldots, K_l;\ l = 1, \ldots, L_1.$$

Φ_1 atitinka kartotines silpstančias eksponentes. Φ_2 yra $N \times 2(K_{L_1+1} + \ldots + K_L)$ matrica

$$\Phi_2 = [d^c_{L_1+1,1}, d^s_{L_1+1,1}, \ldots, d^c_{L_1+1,K_{L_1+1}}, d^s_{L_1+1,K_{L_1+1}}, \ldots,$$
$$d^c_{L1}, d^s_{L1}, \ldots, d^c_{LK_L}, d^s_{LK_L}],$$

čia

$$d^c_{L_1+l,k} = [\delta_{k1}, e^{\lambda_{2,l}}\cos\omega_l, 2^{k-1}e^{2\lambda_{2,l}}\cos 2\omega_l, \ldots,$$
$$(N-1)^{k-1}e^{(N-1)\lambda_{2,l}}\cos(N-1)\omega_l]^T$$

$k = 1, \ldots, K_{L_1+l}$; $l = 1, \ldots, L_2$. Vektorius d^s_{lk} yra apibrėžiamas analogiškai (vienintelis skirtumas yra tas, kad „cos" reikia pakeisti į „sin", o δ_{k1} į 0). Matrica Φ_2 atitinka kartotines silpstančias sinusoides. Φ yra vadinama *bazinių signalų matrica.*

Naudojant šiuos pažymėjimus, neužtriukšmintos sumos vektorių $\mathbf{q}$ galima parašyti taip: $\mathbf{q} = \Phi\mathbf{a}$. Čia $\mathbf{a}$ yra $(K_1 + \ldots + K_{L_1} + 2(K_{L_1+1} + \ldots + K_L)) \times 1$ dydžio vektorius $\mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2]$, kur vektorius $\mathbf{a}_1$ yra tokio pavidalo:

$$\mathbf{a}_1 = [A_{11}, \ldots, A_{1K_1}, \ldots, A_{L_1 1}, \ldots, A_{L_1 K_1}]^T,$$

o vektorius $\mathbf{a}_2$ yra toks:

$$\mathbf{a}_2 = [A_{M1}\sin\varphi_{M1}, A_{M1}\cos\varphi_{M1}, \ldots, \quad A_{MK_M}\sin\varphi_{MK_M}, A_{MK_M}\cos\varphi_{MK_M}, \ldots, \\ A_{L1}\sin\varphi_{L1}, A_{L1}\cos\varphi_{L1}, \ldots, A_{LK_L}\sin\varphi_{LK_L}, A_{LK_L}\cos\varphi_{LK_L}]^T$$

čia $M = L_1 + 1$.

3. NETIESINIAI MAŽIAUSIŲJŲ KVADRATŲ UŽDAVINIAI SU ATSISKIRIANČIAIS KINTAMAISIAIS

Nagrinėkime netiesinį modelį

$$f(\mathbf{a}, \theta; n) = \sum_{m=1}^{M} a_m g_m(\theta; n),$$

čia $\{g_m(\cdot,\cdot\}$ yra tolydžios funkcijos, $\mathbf{a} \in \mathbf{R}^M, \theta \in \mathbf{R}^k$ (M ir k yra teigiami sveiki skaičiai), $n = 1, \ldots, N$. Šį modelį galima parašyti vektoriniu pavidalu $f(\mathbf{a}, \theta) = G(\theta)\mathbf{a}$. $G(\theta) = (G_{nm}(\theta))$ yra $N \times M$ matrica su $G_{nm}(\theta) = g_m(\theta, n)$, $m = 1, \ldots, M$, $n = 1, \ldots, N$. Reikia rasti parametrų $\mathbf{a}, \theta$ reikšmes, kurios minimizuoja funkcionalą

$$r(\mathbf{a}, \theta) = \sum_{n=1}^{N}(y_n - f(\mathbf{a}, \theta; n))^2 = \|\mathbf{y} - G(\theta)\mathbf{a}\|^2,$$

čia $\mathbf{y} = (y_1, \ldots, y_N)^T$ yra duomenų vektorius, $\|\cdot\|$ – Euklido norma: $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$. Esant matricos $G(\theta)$ lokaliai pastovaus rango prielaidai, šis uždavinys yra ekvivalentus funkcionalo

$$r_2(\theta) = \|P_{G(\theta)}^{\perp}\mathbf{y}\|^2,$$

(čia $P_{G(\theta)}^{\perp} = I_N - G(\theta)G^{+}(\theta)$ yra ortogonalusis projektorius į matricos $G(\theta)$ stulpelių erdvės ortogonalųjį papildinį, I_N – vienetinė $N \times N$ matrica) minimizavimo uždaviniui [1]. Suradus θ įvertį, kuris minimizuoja $r_2(\theta)$ vektoriaus $\mathbf{a}$ įvertis apskaičiuojamas taip: $\hat{\mathbf{a}} = G^{+}(\hat{\theta})\mathbf{y}$. Taigi kintamieji (parametrai) yra atskiriami ta prasme, kad pirmiausia įvertinamas „netiesinis" parametras θ , o po to „tiesinis" parametras $\mathbf{a}$. Tokie įvertinimo uždaviniai vadinami netiesinių mažiausiųjų kvadratų uždaviniais su atsiskiriančiais parametrais (žr., [1, 7]).

4. ĮVERTINIMO UŽDAVINIO FORMULAVIMAS

Grįžkime prie eksponentinio-sinusinio modelio (1)–(3). Antrame skyrelyje buvo parodyta, kad šio modelio vektorius $\mathbf{q}$ yra lygus bazinių signalų matricos Φ ir vektoriaus $\mathbf{a}$ sandaugai. Todėl kartotinio eksponentinio-sinusinio modelio parametrų įvertinimo uždavinys gali būti formuluojamas kaip optimalių parametrų $\mathbf{a}$ ir $\theta = [\lambda_{1,1}, \dots, \lambda_{1,L_1}, \lambda_{2,1}, \omega_1, \dots, \lambda_{2,L_2}, \omega_{L_2}]$, minimizuojančių netiesinį funkcionalą

$$r(\mathbf{a}, \theta) = \|x - \Phi(\theta)\mathbf{a}\|^2,$$

radimo uždavinys. Kaip tvirtinama 3-ame skyrelyje, esant matricos $\Phi(\theta)$ lokaliai pastovaus rango prielaidai, šis uždavinys yra ekvivalentus parametro $\hat{\theta}$, minimizuojančio funkcionalą

$$r_2(\theta) = \|P^{\perp}_{\Phi(\theta)}\mathbf{x}\|^2,$$

radimo uždaviniui (čia $P^{\perp}_{\Phi} = I_N - \Phi(\Phi^T\Phi)^{-1}\Phi^T = I_N - \Phi\Phi^+$ yra ortogonalusis projektorius į matricos Φ stulpelių erdvės ortogonalųjį papildinį). Tuomet $\hat{\mathbf{a}} = \Phi^+(\hat{\theta})x$.

5. LEVENBERGO METODO REALIZACIJA EKSPONENTINIO–SINUSINIO MODELIO ATVEJU

Levenbergo metodas, taikant jį funkcionalo $r_2(\cdot)$ minimizavimui, yra parametro θ pradinio įverčio tikslinimo iteratyvi procedūra:

$$\theta^{l+1} = \theta^l - (V^T(\theta^l)V(\theta^l) + c_l I_k)^{-1} V^T(\theta^l)\mathbf{b}(\theta^l), \quad l = 0, 1, \dots$$

kur $V(\theta) = DP^{\perp}_{\Phi(\theta)}\mathbf{x}$ yra $N \times k$ matrica (čia D žymi atvaizdžio Frešė išvestinę), $\mathbf{b}(\theta) = P^{\perp}_{\Phi(\theta)}\mathbf{x}$ – $N \times 1$ vektorius, I_k – $k \times k$ vienetinė matrica ($k = L_1 + 2L_2$), c_l- Levenbergo metodo konstanta l-oje iteracijoje (žr. [1]).

Iš pirmo žvilgsnio atrodo neaišku, kaip apskaičiuoti $V(\theta) = DP^{\perp}_{\Phi(\theta)}\mathbf{x}$. Pirmiausia reikia žinoti, kaip apskaičiuoti $DP^{\perp}_{\Phi(\theta)}$. Darbe [1] įrodyta, kad

$$DP^{\perp}_{\Phi} = -P^{\perp}_{\Phi}D\Phi B - (P^{\perp}_{\Phi}D\Phi B)^T,$$

čia $D\Phi$ yra trimatis tenzorius, sudarytas iš $k(N \times n)$ matricų, kiekvienoje iš kurių yra matricos Φ elementų dalinės išvestinės vienos iš k parametrų vektoriaus θ komponenčių atžvilgiu, o B yra tam tikra simetrinė apibendrinta matricos Φ atvirkštinė ($n = K_1 + \dots + K_{L_1} + 2(K_{L_1+1} + \dots + K_L)$). Tegu S_i yra $N \times n$ matrica, kuri lygi Φ išvestinei parametrų vektoriaus θ i-osios komponentės θ_i atžvilgiu. Tuomet $DP^{\perp}_{\Phi}$ yra sudaryta iš k $(N \times N)$ pavidalo

$$(DP^{\perp}_{\Phi})_i = -P^{\perp}_{\Phi}S_i B - (P^{\perp}_{\Phi}S_i B)^T$$

matricų. Taigi $DP^{\perp}_{\Phi}\mathbf{x}$ yra sudaryta iš k $(N \times 1)$ vektorių $(DP^{\perp}_{\Phi})_i\mathbf{x}$ $(i = 1, \dots, n)$. Vadinasi, $V(\theta)$ yra $N \times k$ matrica.

Kaip apskaičiuoti $P^{\perp}_{\Phi}S_i B$? Speciali matricos S_i struktūra įgalina žymiai supaprastinti tokios matricų sandaugos apskaičiavimą. Žemiau pateikiama teorema atskleidžia šį supaprastinimą.

TEOREMA. *Tegu* $1 \leqslant i \leqslant L_1$. *Tada* $N \times N$ *matrica* $P_\Phi^\perp S_i B$ *yra lygi* $N \times n$ *matricos* $P_\Phi^\perp S_i$ p_1*-ojo stulpelio ir* $n \times N$ *matricos* B p_1*-osios eilutės sandaugai* ($p_1 = K_1 + \ldots + K_i$). *Jei* $L_1 + 1 \leqslant i \leqslant L_1 + L_2$, *tai matrica* $P_\Phi^\perp S_i B$ *yra lygi matricos* $P_\Phi^\perp S_i$ *stulpelių su indeksais* $p_2 - 1, p_2$ *ir matricos* B *eilučių su tais pačiais indeksais sandaugai* ($p_2 = K_1 + \ldots + K_{L_1} + 2(K_{L_1+1} + \ldots + K_i)$).

Įrodymas. Nagrinėkime atvejį $1 \leqslant i \leqslant L_1$. Šiuo atveju matrica S_i yra lygi matricos Φ elementų dalinei išvestinei i-sios kartotinės silpstančios eksponentės gesimo koeficiento λ atžvilgiu. S_i matmenys yra tokie patys kaip matricos Φ, t.y. $N \times n$. Visi n matricos S_i stulpeliai yra lygūs nuliui, išskyrus stulpelius su indeksais $m_1 + 1, m_1 + 2, \ldots, m_1 + K_i$, čia $m_1 = K_1 + \ldots + K_{i-1}$. Matrica S_i turi tokią struktūrą: $S_i = [0_{N\times m_1}\; E\; 0_{N\times r_1}]$, čia $r_1 = K_{i+1} + \ldots + K_{L_1} + 2(K_{L_1+1} + \ldots + K_L)$, o E yra tokio pavidalo $N \times K_i$ matrica

$$E = \begin{pmatrix} 0 & \ldots & 0 & 0 \\ e^\lambda & \ldots & e^\lambda & e^\lambda \\ 2e^{2\lambda} & \ldots & 2^{K_i-1}e^{2\lambda} & 2^{K_i}e^{2\lambda} \\ \vdots & \ddots & \vdots & \vdots \\ (N-1)e^{(N-1)\lambda} & \ldots & (N-1)^{K_i-1}e^{(N-1)\lambda} & (N-1)^{K_i}e^{(N-1)\lambda} \end{pmatrix}$$

(kai $i = 1$, S_i neturi pirmojo nulinio bloko, o kai $i = L_1$, $r_1 = 2(K_{L_1+1} + \ldots + K_L)$). E atitinka tokį bazinių signalų matricos $N \times K_i$ dydžio bloką:

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ e^\lambda & e^\lambda & \ldots & e^\lambda \\ e^{2\lambda} & e^{2\lambda} & \ldots & 2^{K_i-1}e^{2\lambda} \\ \vdots & \vdots & \ddots & \vdots \\ e^{(N-1)\lambda} & (N-1)e^{(N-1)\lambda} & \ldots & (N-1)^{K_i-1}e^{(N-1)\lambda} \end{pmatrix}$$

Nesunku matyti, kad pirmieji K_i-1 matricos stulpeliai priklauso bazinių signalų matricos stulpelių erdvei. Todėl projektoriaus matricos $P_\Phi^\perp$ ir bet kurio iš šių stulpelių sandauga yra lygi nuliniam stulpeliui. Vadinasi, projektoriaus matricos $P_\Phi^\perp$ ir matricos S_i sandauga yra $N \times n$ matrica, kurios visi stulpeliai yra lygūs nuliniam stulpeliui, išskyrus stulpelį su indeksu $p_1 = K_1 + \ldots + K_i$. Taigi matricų sandauga $P_\Phi^\perp S_i \Phi^B$ yra lygi $N \times n$ matricos $P_\Phi^\perp S_i$ p_1-ojo stulpelio ir $n \times N$ matricos B p_1-osios eilutės sandaugai.

Atvejis $L_1 + 1 \leqslant i \leqslant L_1 + L_2$ yra tiriamas analogiškai.

IŠVADOS. Mes nagrinėjome vadinamąjį kartotinį eksponentinį-sinu-sinį modelį. Buvo parodyta, kad tokio modelio parametrų įvertinimo uždavinys yra ekvivalentus netiesiniam mažiausiųjų kvadratų uždaviniui su atsiskiriančiais kintamaisiais. Buvo aptartos problemos, iškylančios realizuojant Levenbergo metodą nagrinėjamojo modelio „netiesinių" parametrų įvertinimui, bei gauti kai kurie rezultatai, supaprastinantys metodo įgyvendinimą šiuo atveju.

LITERATŪRA

[1] G. Golub and V. Pereyra, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate, *SIAM Journal on Numerical Analysis*, **10** (1973), 413–432.

[2] A. Sen and M. Srivastava, *Regression Analysis: Theory, Methods and Applications*, Springer-Verlag, New York, 1990.

[3] J.-J. Fuchs, Multiscale identification of real sinusoids in noise, *Automatica*, **30** (1994), 147–155.

[4] M. R. Osborne and G. K. Smyth, A modified Prony algorithm for exponential function fitting, *SIAM J. Sci. Comput.*, **6** (1995), 119–138.

[5] T. Wigren and A. Nehorai, Asymptotic Cramer-Rao bounds of the parameters of damped sine waves in noise, *IEEE Trans. Signal Processing*, **39** (1991), 1017–1020.

[6] Y. -X. Yao and S. M. Pandit, Cramer-Rao lower bounds for a damped sinusoidal process, *IEEE Trans. Signal Processing*, **43** (1995), 878–885.

[7] G. J. S. Ross, *Nonlinear Estimation*, Springer-Verlag, New York, 1990.

Estimation of some time series models using methodology of a nonlinear least squares problem

V. Šimonytė, V. Slivinskas

We have considered the so-called multiple exponential-sinusoidal model. This model is defined as a sum of multiple damped exponentials and multiple damped sinusoids in noise. The problem of estimation of parameters of the model has been addressed. It has been shown that this problem can be solved as a nonlinear least squares problem whose variables separate. The problems arising in implementation of Levenberg method for estimation of „nonlinear" parameters of the model have been discussed, and some results simplifying implementation of the method in this case have been obtained.