Aalto-yliopisto
Perustieteiden korkeakoulu
Matematiikan ja systeemianalyysin laitos

MS-E2129 Systeemien identifiointi

9. harjoituksen ratkaisut

1. (i) Tarkasteltava malli voidaan siis kirjoittaa muodossa

$$y(t) = ay(t-1) + bu(t-1) + e(t),$$

eli

$$y(t) = \begin{bmatrix} y(t-1) & u(t-1) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + e(t) = \phi^T(t)\theta + e(t).$$

Nyt käytössämme on havainnot $y(t)$, $t = 1, \ldots, N$, jolloin kokoamalla $y(t)$:t, $\phi(t)$:t ja virhetermit $e(t)$ vektoreiksi ja matriiseiksi y, ϕ ja e saadaan, että

$$e = y - \phi\theta,$$

missä siis

$$\phi = \begin{bmatrix} y(1) & u(1) \\ y(2) & u(2) \\ \vdots & \vdots \\ y(N-1) & u(N-1) \end{bmatrix}$$

Nyt PNS-estimointi tuottaa parametriestimaatin

$$\hat{\theta} = (\phi^T\phi)^{-1}\phi^T y.$$

Tarkastellaan ensin harhaisuus, eli se, onko $E\hat{\theta} = \theta$. Nyt kannattaa kirjoittaa parametriestimaatti hieman eri tavalla:

$$\hat{\theta} = (\phi^T\phi)^{-1}\phi^T(\phi\theta + e) = \theta + (\phi^T\phi)^{-1}\phi^T e,$$

eli siis harha

$$\begin{aligned}
\theta - \hat{\theta} &= -(\phi^T\phi)^{-1}\phi^T e \\
&= -\begin{bmatrix} \sum y(i)^2 & \sum y(i)u(i) \\ \sum y(i)u(i) & \sum u(i)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y(i)e(i+1) \\ \sum u(i)e(i+1) \end{bmatrix} \\
&= \frac{1}{\det \phi^T\phi} \begin{bmatrix} -\sum u(i)^2 & \sum y(i)u(i) \\ \sum y(i)u(i) & -\sum y(i)^2 \end{bmatrix} \begin{bmatrix} \sum y(i)e(i+1) \\ \sum u(i)e(i+1) \end{bmatrix} \\
&= \frac{1}{\det \phi^T\phi} \begin{bmatrix} -\sum u(i)^2 \sum(y(i)e(i+1)) + \sum(y(i)u(i)) \sum(u(i)e(i+1)) \\ \sum(y(i)u(i)) \sum(y(i)e(i+1)) - \sum y(i)^2 \sum(u(i)e(i+1)) \end{bmatrix}.
\end{aligned}$$

Nyt esimerkiksi

$$\sum_{i=1}^{N-1} y(i)u(i) \sum_{i=1}^{N-1} u(i)e(i+1) = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} y(i)u(i)u(j)e(j+1),$$

joten edellisissä summissa esiintyy ristitermejä $y(i)e(k)$, missä $i > k$. Näiden odotusarvo ei ole nolla, joten estimaatti on harhainen.

Kuitenkin, jos $N \to \infty$ niin

$$\theta - \hat{\theta} \to \frac{1}{\mathrm{Var}[y]\mathrm{Var}[u] - \mathrm{Cov}[y,u]^2} \begin{bmatrix} -\mathrm{Var}[u]\mathrm{R}_{ey}(1) + \mathrm{Cov}[y,u]\mathrm{R}_{eu}(1) \\ \mathrm{Cov}[y,u]\mathrm{R}_{ey}(1) - \mathrm{Var}[y]\mathrm{R}_{eu}(1) \end{bmatrix} = 0,$$

koska $\mathrm{R}_{eu}(1) = 0$ ja $\mathrm{R}_{ey}(1) = 0$ ($e(t)$ on valkoista kohinaa ja $y(t)$ sekä $u(t)$ eivät voi riippua tulevasta $e(t+1)$). Siten estimaatti on tarkentuva.

(ii) Nyt mallina on

$$y(t) = ay(t-1) + bu(t-1) + e(t) - ae(t-1).$$

Koska suuretta $e(t)$ ei voida mitata, ei parametria a voida estimoida suoraan termistä $ae(t-1)$, vaan tämä sisällytetään virhetermiin. Merkitään $v(t) = e(t) - ae(t-1)$, jolloin

$$y(t) = ay(t-1) + bu(t-1) + v(t),$$

missä $v(t)$ on autokorreloitunut virhetermi.

Kuten edellä saadaan PNS-estimaatiksi

$$\hat{\theta} = (\phi^T\phi)^{-1}\phi^T v,$$

eikä siten tässäkään tapauksessa estimaatti ole harhaton. Lisäksi virhetermi v ei nyt ole valkoista kohinaa, jolloin $\mathrm{R}_{ey}(1) \neq 0$ (sekä $y(t)$ että $e(t+1)$ riippuvat $e(t)$:stä) ja siten estimaatti ei ole tarkentuva.

Perusongelma on se, että virheen mallia ei huomioida asianmukaisesti.

2. Mallin ennustevirhe voidaan kirjoittaa muodossa

$$CF\varepsilon(t) = FDy(t) - BDu(t),$$

jota derivoimalla saadaan (huomaa, että ε on funktio parametreista):

$$\begin{aligned}
\frac{\partial}{\partial b_i} &: CF\frac{\partial\varepsilon(t)}{\partial b_i} = -Du(t-i) \\
\frac{\partial}{\partial c_i} &: F\varepsilon(t-i) + CF\frac{\partial\varepsilon(t)}{\partial c_i} = 0 \\
\frac{\partial}{\partial d_i} &: CF\frac{\partial\varepsilon(t)}{\partial d_i} = Fy(t-i) - Bu(t-i) \\
\frac{\partial}{\partial f_i} &: C\varepsilon(t-i) + CF\frac{\partial\varepsilon(t)}{\partial f_i} = Dy(t-i).
\end{aligned}$$

Tästä saadaan ratkaistua kysytyt derivaatat:

$$
\begin{aligned}
\frac{\partial \varepsilon(t)}{\partial b_i} &= \frac{1}{CF}(-Du(t-i)) = -\frac{D}{CF}u(t-i)\\
\frac{\partial \varepsilon(t)}{\partial c_i} &= \frac{1}{CF}(-F\varepsilon(t-i)) = -\frac{1}{C}\varepsilon(t-i)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \varepsilon(t)}{\partial d_i} &= \frac{1}{CF}(Fy(t-i) - Bu(t-i))\\
&= \frac{1}{CFD}(FDy(t-i) - BDu(t-i))\\
&= \frac{1}{CFD}(CF\varepsilon(t-i)) = \frac{1}{D}\varepsilon(t-i)\\
\frac{\partial \varepsilon(t)}{\partial f_i} &= -\frac{1}{CF}(-C\varepsilon(t-i) + Dy(t-i))\\
&= \frac{1}{CF^2}(-CF\varepsilon(t-i) + FDy(t-i))\\
&= \frac{1}{CF^2}(-BDu(t-i)) = \frac{BD}{CF^2}u(t-i).
\end{aligned}
$$

3. $V(\theta)$:n neliöllinen approksimaatio pisteen θ^k ympäristössä:

$$V(\theta) \approx Q_V(\theta, \theta^k) = V(\theta^k) + V'(\theta^k)(\theta - \theta^k) + \frac{1}{2}V''(\theta^k)(\theta - \theta^k)^2.$$

Tämän minimi saadaan gradientin nollakohdasta:

$$
\begin{aligned}
\frac{\partial Q_V(\theta, \theta^k)}{\partial \theta} &= V'(\theta^k) + V''(\theta^k)(\theta - \theta^k) = 0\\
\Rightarrow \theta - \theta^k &= -V''(\theta^k)^{-1}V'(\theta^k)\\
\Rightarrow \theta &= \theta^k - V''(\theta^k)^{-1}V'(\theta^k)
\end{aligned}
$$

Tämä on sama kuin Newton-Raphson algoritmi vakioaskeleella $\alpha_k = 1$.

4. a) Mittausjärjestely on esitetty kuvassa 1. Nyt τ saadaan ratkaistua helposti yhtäaikaa (x, y, z):n kanssa yhtälöryhmästä:

$$
\begin{aligned}
(s_{11} - x)^2 + (s_{12} - y)^2 + (s_{13} - z)^2 &= (r_1 - \tau)^2\\
(s_{21} - x)^2 + (s_{22} - y)^2 + (s_{23} - z)^2 &= (r_2 - \tau)^2\\
(s_{31} - x)^2 + (s_{32} - y)^2 + (s_{33} - z)^2 &= (r_3 - \tau)^2\\
(s_{41} - x)^2 + (s_{42} - y)^2 + (s_{43} - z)^2 &= (r_4 - \tau)^2,
\end{aligned}
$$

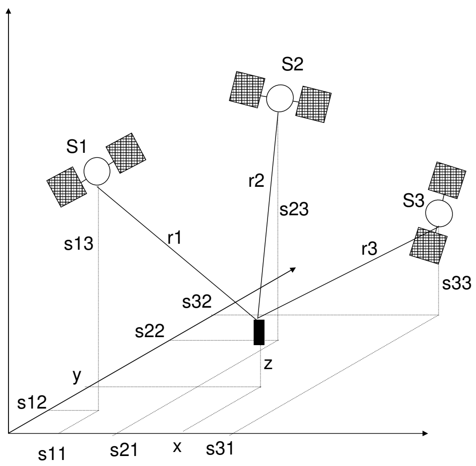
eli 3d-paikannukseen tarvitaan neljän satelliitin mittaustulos.

b) Oletetaan nyt, että satelliitteja on $N > 4$ kappaletta, ja mittauksissa kohinaa. Satelliiteilta saadaan etäisyysmittaukset

$$R_i = \sqrt{(s_{i1} - x)^2 + (s_{i2} - y)^2 + (s_{i3} - z)^2} + \tau + \varepsilon_i.$$

Siten meillä on malli

$$R = f(S, X) + \varepsilon,$$



Kuva 1: Tehtävän 4 GPS mittausjärjestely

missä $R = \begin{bmatrix} R_1 & \dots & R_N \end{bmatrix}^T$ ja paikkamatriisi

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ \vdots & \vdots & \vdots \\ s_{N1} & s_{N2} & s_{N3} \end{bmatrix}$$

ja $X = \begin{bmatrix} x & y & z & \tau \end{bmatrix}^T$. Estimoidaan X iteratiivisella linearisoinnilla. Linearisoidaan malli:

$$\Delta R = \frac{\partial f}{\partial X}(S, X^*)\Delta X + \varepsilon \text{ (tämä on muotoa } y = \phi\theta + e),$$

missä $\Delta R = R - R^*$, $\Delta X = X - X^*$ ja tähdellä indeksoinnit viittaavat edellisen iteraatiokierroksen ratkaisuihin.

1. Laske $\frac{\partial f}{\partial X}(S, X^*)$
2. Laske ΔR mittausten avulla.
3. Oletetaan $\varepsilon_i \sim N(0, \sigma_i^2)$ sekä ε_i ja ε_j riippumattomiksi kun $i \neq j$. Laske painomatriisi $W = \mathrm{diag}\left(\frac{1}{\sigma_i^2}\right)$.
4. Laske $\Delta X = \left(\frac{\partial f}{\partial X}^T W \frac{\partial f}{\partial X}\right)^{-1} \frac{\partial f}{\partial X} W \Delta R$, eli yleistetty PNS (GLS) parametriestimaatti lineaarisoidulle systeemille.
5. Päivitä $X^* = X^* + \Delta X$
6. Jos $\|\Delta X\|$ "pieni", lopeta; Muuten mene kohtaan 1.

Huomaa, että menetelmässä tarvitaan alkuarvaus X:lle. Jos paikannin liikkuu hitaasti voidaan alkuarvauksena käyttää edellisen hetken paikkatietoa. Muuten joudutaan approksimoimaan mittaushistoriasta paikantimen liiketilaa.

5. Lasketaan ensin $\hat{b}$ derivoimalla $V = \sum_{t=1}^{N} \lambda^{N-t}(y(t)-bu(t))^2$ ja etsimällä nollakohta:

$$\begin{aligned}
\frac{\partial V}{\partial b} = \sum 2\lambda^{N-t}(y(t) - bu(t))(-u(t)) = 0 \\
\sum \lambda^{N-t} y(t)u(t) - b \sum \lambda^{N-t} u(t)^2 = 0 \\
\hat{b} = \frac{\sum \lambda^{N-t} y(t)u(t)}{\sum \lambda^{N-t} u(t)^2}.
\end{aligned}$$

Tästä voidaankin sitten laskea varianssi

$$\begin{aligned}
\mathrm{Var}[\hat{b}] &= \mathrm{E}[(b - \hat{b})^2] \\
&= \mathrm{E}\left[\left(b - \frac{\sum \lambda^{N-t}(bu(t) + e(t))u(t)}{\sum \lambda^{N-t} u(t)^2}\right)^2\right] \\
&= \mathrm{E}\left[\left(b - b\frac{\sum \lambda^{N-t} u(t)^2}{\sum \lambda^{N-t} u(t)^2} - \frac{\sum \lambda^{N-t} e(t)u(t)}{\sum \lambda^{N-t} u(t)^2}\right)^2\right] \\
&= \mathrm{E}\left[\frac{(\sum \lambda^{N-t} e(t)u(t))^2}{(\sum \lambda^{N-t} u(t)^2)^2}\right] \\
&= \frac{\mathrm{E}\left[(\sum \lambda^{N-t} e(t)u(t))(\sum \lambda^{N-s} e(s)u(s))\right]}{(\sum \lambda^{N-t} u(t)^2)^2} \\
&= \frac{\mathrm{E}\left[(\sum \sum \lambda^{N-t}\lambda^{N-s} e(t)e(s)u(t)u(s))\right]}{(\sum \lambda^{N-t} u(t)^2)^2} \\
&= \frac{\sum \lambda^{2(N-t)} u(t)^2}{(\sum \lambda^{N-t} u(t)^2)^2}, \text{ koska } \mathrm{E}[e(t)e(s)] = \delta_{t,s}.
\end{aligned}$$

Kun $\lambda = 1$, niin $\mathrm{Var}[\hat{b}] = \frac{1}{\sum u(t)^2}$ ja $\lim_{N\to\infty} \mathrm{Var}[\hat{b}] = 0$.

Näytetään sitten, että jos $\lambda < 1$, niin on olemassa u, s.e. varianssi ei mene nollaan. Valitaan $u(t) = a$. Tällöin

$$\mathrm{Var}[\hat{b}] = \frac{\sum \lambda^{2(N-t)} a^2}{(\sum \lambda^{N-t} a^2)^2} = \frac{1}{a^2}\frac{\sum \lambda^{2(N-t)}}{(\sum \lambda^{N-t})^2} = \frac{1}{a^2}\frac{\sum (\lambda^2)^{t-1}}{(\sum \lambda^{t-1})^2}.$$

Käyttäen geometrisen sarjan kaavaa, saadaan

$$\lim_{N\to\infty} \mathrm{Var}[\hat{b}] = \frac{1}{a^2}\frac{1/(1-\lambda^2)}{(1/(1-\lambda))^2} = \frac{1}{a^2}\frac{(1-\lambda)^2}{1-\lambda^2} = \frac{1-\lambda}{a^2(1+\lambda)}.$$

6. Tehtävän mittausdata (sisäänmeno ja ulostulo) ovat kurssin kotisivulla tallennettuina tiedostossa las09t7.mat.

   Olkoot A, B, C, D, E ja F taaksepäinsiirto-operaattorin polynomeja, ja näitä vastaavat asteluvut n_A, n_B, n_C, n_D, n_E ja n_F. Tällöin:

   a) ARX-malli:
   $$A(q)y(t) = B(q)u(t) + e(t).$$
   Mallin voit estimoida Matlabissa esimerkiksi funktiolla arx. Mallin valinnassa voit käyttää apuna Matlab-funktiota selstruc.

b) OE-malli:

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t).$$

Matlabissa käytä funktiota oe.

c) ARMAX-malli:

$$A(q)y(t) = B(q)u(t) + C(q)e(t).$$

Käytä Matlabin funktiota armax.

d) BJ-malli:

$$y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t).$$

Matlab-funktio bj.

Lopuksi voit tarkastella funktiota las09t7gen, jossa näytetään miten ulostulo on generoitu. Havaitset, että todellinen järjestelmä on epälineaarinen ja siten mikään malli ei oikein istu dataan.