

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Mada linearni vremenski invarijantni sistemi predstavljaju najčešći oblik opisa dinamičkog sistema, vrlo često je korisno ili nužno upotrebiti i druge opise.

Linearni vremenski varirajući modeli

Težinske funkcije

Ranije smo definirali linearni sistem kao onaj gdje linearna kombinacija ulaza vodi do iste linearne kombinacije odgovarajućih izlaza. Opšti linearni sistem se dakle može opisati sa:

$$y(t) = \sum_{k=1}^{\infty} g_t(k)u(t-k) + v(t) \quad (5.1)$$

Ako pišemo :

$$g_t(k) = \tilde{g}(t, t-k)$$

dobijemo:

$$y(t) = \sum_{s=-\infty}^{t-1} \tilde{g}(t, s)u(s) + v(t) \quad (5.2)$$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

gdje $\tilde{g}(t, s), t = s, s+1, \dots$ je odziv u trenutku t na jedinični impuls u trenutku s . Funkcija $\tilde{g}(t, s)$ je poznata kao težinska funkcija, pošto opisuje težinu koju ima ulaz u trenutku vremena s na izlaz u trenutku t . Opis (5.1) je potpuno analogan sa vremenski invarijantnim modelom (2.8) izuzev da sekvenca $g_t(k)$ nosi vremenski indeks t . Općenito, mogli bi uvesti vremenski varirajuću prenosnu funkciju sa:

$$G_t(q) = \sum_{k=1}^{\infty} g_t(k)q^{-k} \quad (5.3)$$

U praksi je lakše raditi sa vremenskom varijacijom u opisu modela u prostoru stanja.

Model u prostoru stanja sa vremenskim varijacijama

Vremenska varijacija u modelima prostora stanja (4.91) se jednostavno dobije na taj način da se dozvoli da matrice budu vremenski zavisne tj.

$$\begin{aligned} x(t+1, \theta) &= A_t(\theta)x(t, \theta) + B_t(\theta)u(t) + K_t(\theta)e(t) \\ y(t) &= C_t(\theta)x(t, \theta) + e(t) \end{aligned} \quad (5.4) \quad 2$$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Prediktor koji korespondira sa (4.86) se dobije kao:

$$\begin{aligned}\hat{x}(t+1, \theta) &= [A_t(\theta) - K_t(\theta)C_t(\theta)]\hat{x}(t, \theta) + B_t(\theta)u(t) + K_t(\theta)y(t) \quad (5.5) \\ \hat{y}(t|\theta) &= C_t(\theta)\hat{x}(t, \theta)\end{aligned}$$

Primjetimo da se ovo može pisati kao:

$$\hat{y}(t|\theta) = \sum_{k=1}^{\infty} w_t^u(k, \theta)u(t-k) + \sum_{k=1}^{\infty} w_t^y(k, \theta)y(t-k) \quad (5.6)$$

sa:

$$\begin{aligned}w_t^u(k, \theta) &= C_t(\theta) \prod_{j=t-k}^{t-1} [A_j(\theta) - K_j(\theta)C_j(\theta)] B_{t-k}(\theta) \\ w_t^y(k, \theta) &= C_t(\theta) \prod_{j=t-k}^{t-1} [A_j(\theta) - K_j(\theta)C_j(\theta)] K_{t-k}(\theta)\end{aligned} \quad (5.7)$$

Slično, mi bi mogli startati sa vremenski varirajućim modelom kao (4.84) i (4.85), gdje matrice A, B, C, R_1, R_{12} , i R_2 su funkcije od t . Odgovarajući prediktor će onda biti dat sa (4.94) i (4.95). Dva zajednička problema kod vremenski invarijantnih sistema vode ka vremenski zavisnim opisima: nejednako uzorkovanje i

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

i linearizacija. Ako su sistemi (4.62) i (4.63) uzorkovani u vremenskim trenutcima $t=t_k$, $k=1,2,\dots$, mi još uvijek možemo primjeniti formule uzorkovanja (4.66) do (4.68) da predjemo od t_k na t_{k+1} , koristeći $T_k = t_{k+1} - t_k$. Ako ovaj interval uzorkovanja nije konstantan, (4.67) će biti vremenski varirajući sistem. Sa ovim u vezi je i multirate sistem, tj sistem kod koga su različite varijable uzorkovane sa različitim brzinama. Tada $C_t(\mathbf{q})$ matrica u (5.4) će biti vremenski varirajuća da bi se dobila stanja koja su uzorkovana u trenutku t . Nedostajući izlazni podaci se takodjer mogu posmatrati kao neuniformno uzorkovanje.

Linearizacija nelinearnih sistema

Možda najčešće korištenje vremenski promjenljivih linearnih sistema je vezano za linearizaciju nelinearnog sistema oko neke trajektorije. Predpostavimo da je nelinearni sistem opisan sa:

$$\begin{aligned}x(t+1) &= f(x(t), u(t)) + r(x(t), u(t))w(t) \\ y(t) &= h(x(t)) + m(x(t), u(t))v(t)\end{aligned}\tag{5.8}$$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Prepostavimo također da članovi smetnji $\{w(t)\}$ i $\{v(t)\}$ su bijeli i mali, i da nominalno ponašanje sistema bez smetnji ($w(t)=0$ i $v(t)=0$), korespondira sa ulaznom sekvencom $u^*(t)$ i odgovarajućom trajektorijom $x^*(t)$. Zanemarujući nelinearne članove, diferencije :

$$\Delta x(t) = x(t) - x^*(t) \quad ;$$

$$\Delta y(t) = y(t) - h(x^*(t))$$

$$\Delta u(t) = u(t) - u^*(t)$$

su podvrgnute

$$\begin{aligned} \Delta x(t+1) &= F(t)\Delta x(t) + G(t)\Delta u(t) + \bar{w}(t) \\ \Delta y(t) &= H(t)\Delta x(t) + \bar{v}(t) \end{aligned} \quad (5.9)$$

gdje

$$F(t) = \left. \frac{\partial f(x, u)}{\partial x} \right|_{x^*(t), u^*(t)}, \quad G(t) = \left. \frac{\partial f(x, u)}{\partial u} \right|_{x^*(t), u^*(t)}$$

$$H(t) = \left. \frac{\partial h(x)}{\partial x} \right|_{x^*(t)}$$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Ovdje smo mi zanemarili kros članove sa smetnjom (kao $\Delta x \cdot v$), u svjetlu pretpostavke o malim poremećajima. U (5.9), $\bar{w}(t)$ i $\bar{v}(t)$ su bijele smetnje sa slijedećim kovarijantnim osobinama:

$$\begin{aligned} R_1(t) &= E\bar{w}(t)\bar{w}^T(t) = r(x^*(t), u^*(t)) Ew(t)w^T(t)r^T(x^*(t), u^*(t)) \\ R_2(t) &= E\bar{v}(t)\bar{v}^T(t) = m(x^*(t), u^*(t)) Ev(t)v^T(t)m^T(x^*(t), u^*(t)) \quad (5.10) \\ R_{12}(t) &= r(x^*(t), u^*(t)) Ew(t)v^T(t)m^T(x^*(t), u^*(t)) \end{aligned}$$

Ovaj model je sada linearni, vremenski varirajući aproksimativni opis od (5.8) u blizini nominalne trajektorije.

Modeli sa nelinearnostima

Nelinearna relacija izmedju ulazne sekvence i izlazne sekvence kao u (5.8) daje mnogo bogatije mogućnosti da opiše sistem.

Uvijek je mudro prvo pokušati da se iskoristi fizikalno sagledavanje karaktera mogućih nelinearnosti i konstruisati strukturu modela na taj način. Mi ćemo koristiti takozvano

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

semi fizikalno modeliranje, gdje podrazumjevamo korištenje jednostavnog fizikalnog razumjevanja da sagledamo suštinske nelinearnosti i da ih uključimo u modele.

Wienerovi i Hammersteinovi modeli

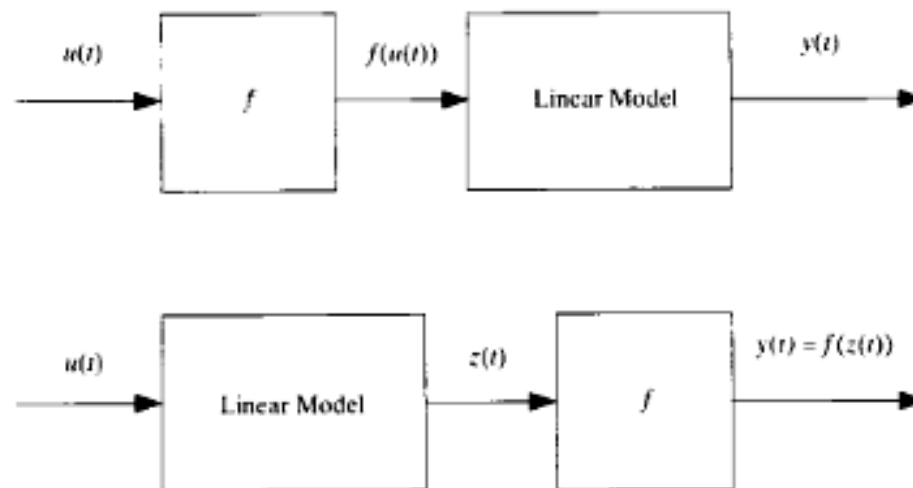
Često se susrećemo sa situacijom u kojoj je moguće opisati dinamiku sistema sa linearnim sistemom, ali postoje statičke nelinearnosti na ulazu i/ili izlazu sistema. Ovo je napr. slučaj kada su aktuatori i ventili nelinearni, napr. zbog zasićenja, ili ako senzori imaju nelinearne karakteristike. Model sa statičkom nelinearnosti na ulazu se naziva Hammerstein-ov model, dok govorimo o Wienerovom modelu ako je nelinearnost na izlazu, kao na slici br. 5.1. Kombinacija je onda poznata kao Wiener-Hammersteinov model. Parametrizacija takvih modela je prilično direktna. Posmatrajmo prvo Hammersteinov slučaj. Statička nelinarna funkcija $f(\cdot)$ se može parametrizirati bilo u terminima fizikalnih parametara, kao što je tačka zasićenja ili nivo zasićenja, ili u terminima crne kutije (black-box), kao koeficijenti spline funkcije.

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Ovo daje $f(\cdot, \eta)$. Ako linearni model je $G(q, \theta)$, procjenjeni izlazni model će biti:

$$\hat{y}(t|\theta, \eta) = G(q, \theta) f(u(t), \eta) \quad (5.11)$$

koji nije mnogo komplikovaniji od modela u prethodnom poglavlju.



Slika 5.1 Hammerstein model gore i Wienerov dole

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Linearni regresioni modeli

U (4.12) smo definirali strukturu linearnog regresionog modela, gdje je predikcija linearna po parametrima:

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \quad (5.12)$$

Da bi se opisala linearna diferentna jednačina, komponente vektora $\mathbf{f}(t)$ (tj. regresori) , su izabrani kao zakašnjele vrijednosti ulaza i izlaza. Kada koristimo (5.12), nije bitno kako je $\mathbf{f}(t)$ formirano, ono što je važno je da je to poznata vrijednost u trenutku t . Zbog toga joj možemo dozvoliti da sadrži arbitrarne transformacije mjerenih podataka. Neka y^t i u^t označavaju ulazne i izlazne sekvence do vremena t . Tada možemo pisati:

$$\hat{y}(t|\theta) = \theta_1\varphi_1(u^t, y^{t-1}) + \dots + \theta_d\varphi_d(u^t, y^{t-1}) = \varphi^T(t)\theta \quad (5.13)$$

sa arbitrarnim funkcijama \mathbf{f}_i prošlih podataka. Struktura (5.13) se može posmatrati kao konačno dimenzionalna paramaterizacija opšteg, nelinearnog prediktora.

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Ključno je kako izabrati funkcije $\varphi_i(u^t, y^{t-1})$. Postoji nekoliko mogućnosti.

Pokušaj sa black box ekspanzijom.

Mogli bi pokušati konstruisati regresore kao tipične (polinomijalne) kombinacije prošlih ulaza i izlaza i vidjeti da li je model u stanju da opiše podatke. Ovaj pristup normalno daje veliki broj mogućih regresora. Nekad je jednostavnije za Hammersteinov model, gdje mi možemo aproksimirati statičku nelinearnost sa polinomijalnom ekspanzijom:

$$f(u) = \alpha_1 u + \alpha_2 u^2 + \dots + \alpha_m u^m \quad (5.14)$$

Svaka potencija od u može zatim dati različitu dinamiku brojnika:

$$A(q)y(t) = B_1(q)u(t) + B_2(q)u^2(t) + \dots + B_m(q)u^m(t) \quad (5.15)$$

Ovo je jasno struktura linearnog regresionog modela.

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Korištenje fizikalnog sagledavanja.

Ponekad sagledavanje korištenjem zakona fizike, može otkriti koje su bitne nelinearnosti u sistemu. Ovo će onda sugerisati koje regresore treba pokušati u (5.13). Ovo smo već nazvali semi-fizikalno modeliranje.

Primjer: Kuća sa solarnim grijanjem

Posmatrajmo problem identifikacije dinamike kuće sa solarnim grijanjem. Nama je potreban model kako na temperaturu $y(t)$ utiče brzina pumpe i intenzitet sunčeve energije.

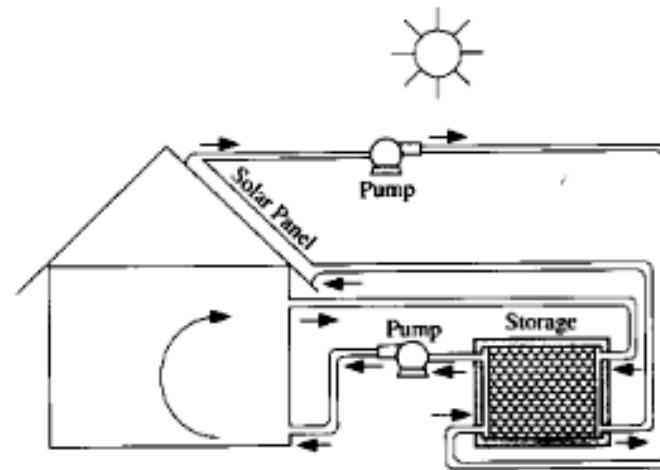


Figure 1.2 A solar-heated house.

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Direktan linearni model tipa (4.7) bi bio:

$$\begin{aligned} y(t) + a_1 y(t-1) + a_2 y(t-2) \\ = b_1 u(t-1) + b_2 u(t-2) + c_1 I(t-1) + c_2 I(t-2) \end{aligned} \quad (5.16)$$

Ovdje nismo koristili nikakvo sagledavanje fizikalnosti procesa grijanja, nego smo uveli black-box model (5.16) na ad hoc način. Mala analiza otkriva da linearni model nije realističan. Jasno, da efekti sunčevog intenziteta i brzine pumpe nisu aditivne osobine. Kada se pumpa isključi, sunce više neće uopšte uticati na temperaturu pohranjene tečnosti. Pogledajmo šta se dešava u sistemu grijanja. Uvedimo $x(t)$ kao temperaturu kolektora solarnog panela u trenutku t . Uz određena pojednostavljenja, fizikalnost se može opisati u diskretnom vremenu kako slijedi: Grijanje zraka u kolektoru $[= x(t+1) - x(t)]$ je jednako toploti koju daje sunce ($= d_2 I(t)$) umanjeno za gubitke toplote u okolinu ($= d_3 x(t)$), i umanjeno za dio toplote prenešen

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

u akumulator ($= d_0 x(t) u(t)$). , tj

$$x(t + 1) - x(t) = d_2 I(t) - d_3 x(t) - d_0 x(t) \cdot u(t) \quad (5.17)$$

Na isti način, porast temperature u akumulatoru toplote [$= y(t+1) - y(t)$] je jednak prenesenoj toploti [$= d_0 x(t) u(t)$] minus gubitci u okolinu [$= d_1 y(t)$] , tj.

$$y(t + 1) - y(t) = d_0 x(t) u(t) - d_1 y(t) \quad (5.18)$$

U jednačinama (5.17) i (5.18) koeficijenti d_k su nepoznate konstante, čije numeričke vrijednosti se trebaju odrediti. Temperatura $x(t)$ se ne mjeri, tako da mi prvo eliminišemo $x(t)$ iz (5.17) i (5.18). Ovo daje:

$$\begin{aligned} y(t) = & (1 - d_1)y(t - 1) + (1 - d_3) \frac{y(t - 1)u(t - 1)}{u(t - 2)} \\ & + (d_3 - 1)(1 - d_1) \frac{y(t - 2)u(t - 1)}{u(t - 2)} + d_0 d_2 u(t - 1) I(t - 2) \\ & - d_0 u(t - 1) y(t - 1) + d_0 (1 - d_1) u(t - 1) y(t - 2) \end{aligned} \quad (5.19)$$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Relacija izmedju mjenjenih vrijednosti y, u i I i parametara d_i je sada komplikovanija. Može se pojednostaviti putem reparametrizacije:

$$\begin{aligned}
 \theta_1 &= (1 - d_1) & \varphi_1(t) &= y(t - 1) \\
 \theta_2 &= (1 - d_3) & \varphi_2(t) &= \frac{y(t - 1)u(t - 1)}{u(t - 2)} \\
 \theta_3 &= (d_3 - 1)(1 - d_1) & \varphi_3(t) &= \frac{y(t - 2)u(t - 1)}{u(t - 2)} \\
 \theta_4 &= d_0 d_2 & \varphi_4(t) &= u(t - 1)I(t - 2) \\
 \theta_5 &= -d_0 & \varphi_5(t) &= u(t - 1)y(t - 1) \\
 \theta_6 &= d_0(1 - d_1) & \varphi_6(t) &= u(t - 1)y(t - 2) \\
 \theta^T &= [\theta_1 \ \theta_2 \ \dots \ \theta_6] & \varphi^T(t) &= [\varphi_1(t) \ \varphi_2(t) \ \dots \ \varphi_6(t)]
 \end{aligned} \tag{5.20}$$

Izraz (5.19) se može pisati kao prava linearna regresija

$$y(t) = \hat{y}(t|\theta) = \varphi^T(t)\theta \tag{5.21}$$

gdje imamo linearnu relaciju izmedju novih parametara \mathbf{q} i konstruisanih mjerenja $\mathbf{f}(t)$. Primjetimo da \mathbf{f} ne zavisi od \mathbf{q} ⁻¹⁴

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Cijena za ovo je da znanje o algebarskim relacijama između θ_i prema (5.20) je izgubljeno. Jednostavna struktura (5.21) daje razumno dobar model.

Nelinearni modeli u prostoru stanja

Opšti skup modela

Najopštiji opis konačno dimenzionalnog sistema je:

$$\begin{aligned}x(t+1) &= f(t, x(t), u(t), w(t); \theta) \\ y(t) &= h(t, x(t), u(t), v(t); \theta)\end{aligned}\tag{5.22}$$

Ovdje su $w(t)$ i $v(t)$ sekvence nezavisnih slučajnih varijabli i q označava vektor nepoznatih parametara. Problem određivanja prediktora na bazi (5.22) i na formalnim probabilističkim osnovama je značajan. U stvari, ovaj problem nelinearne predikcije je poznat da nema konačno dimenzionalno rješenje izuzev u izolovanim specijalnim slučajevima. Ipak, prediktori za

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

(5.22) se mogu konstruisati, ili sa ad hoc pristupom ili pomoću neke aproksimacije neostvarivog optimalnog rješenja. U tom slučaju, rezultirajući prediktor poprima oblik:

$$\hat{y}(t|\theta) = g(t, Z^{t-1}; \theta) \quad (5.23)$$

Sada, radi lakšeg označavanja uvodimo:

$$Z^t = (y^t, u^t) = (y(1), u(1) \dots y(t), u(t))$$

da označimo ulazno/izlazna mjerenja koja su raspoloživa u trenutku t . Ovo je forma u koju se stavlja model da bi se koristio za namjene identifikacije.

Model (5.23) se može komplementirati sa pretpostavkama o pridruženoj grešci predikcije:

$$\varepsilon(t, \theta) = y(t) - g(t, Z^{t-1}; \theta) \quad (5.24)$$

gdje je njegova kovarijantna matrica $\Lambda(t; \theta)$ a njegov PDF $f_\varepsilon(x, t; \theta)$.

Nelinearni simulacioni model

Naročito jednostavan način da se izvede prediktor iz (5.22) je da se zanemari procesni šum $w(t)$ i uzme:

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

$$\begin{aligned}x(t + 1, \theta) &= f(t, x(t, \theta), u(t), 0; \theta) \\ \hat{y}(t|\theta) &= h(t, x(t, \theta), u(t), 0; \theta)\end{aligned}\tag{5.25}$$

Ovakav prediktor se naziva simulacioni model, pošto je $\hat{y}(t|\theta)$ konstruisano simulirajući model bez šuma (5.22) koristeći stvarni ulaz. Simulacioni model je takodjer lako koristiti polazeći od predstave za kontinualno vrijeme:

$$\begin{aligned}\frac{d}{dt}x(t, \theta) &= f(t, x(t, \theta), u(t), 0; \theta) \\ \hat{y}(t|\theta) &= h(t, x(t, \theta), u(t), 0; \theta)\end{aligned}\tag{5.26}$$

Primjer: Delignifikacija drveta

Posmatrajmo problem smanjenja sadržaja lignina u otpacima drveta u kemijskoj mješavini. Ovo je proces kuhanja celuloze, da se dobije pulpa za izradu papira. Uvešćemo slijedeće oznake:

$x(t)$: koncentracija lignina u trenutku t

$u_1(t)$: absolutna temperatura u trenutku t

$u_2(t)$: koncentracija vodikovog sulfata $[\text{HSO}_3^-]$

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

$u_3(t)$: koncentracija vodika [H^+]

Osnovni zakoni kemije nam daju:

$$\frac{d}{dt}x(t) = -k_1 e^{-E_L/u_1(t)} [x(t)]^m \cdot [u_2(t)]^\alpha \cdot [u_3(t)]^\beta . \quad (5.27)$$

ovdje je E_L - Arrheniousova konstanta a k_1 , m , α i β su konstante pridružene sa reakcijom.

Simulacijom (5.27) sa mjerenim vrijednostima od $\{u_i(t), i = 1, 2, 3\}$ za date vrijednosti od $\theta^T = [E_L, k_1, m, \alpha, \beta]$ daje sekvencu odgovarajućih koncentracija lignina $\{x(t, \theta)\}$. U ovom slučaju je i izlaz sistema također koncentracija lignina, tako da je

$\hat{y}(t|\theta) = x(t, \theta)$. Ove prognozirane ili simulirane vrijednosti se mogu onda porediti sa stvarnim izmjerenim vrijednostima tako da greške udružene sa odredjenom specifičnom vrijednošću \mathbf{q} se mogu evaluirati.

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

Nelinearni black-box modeli: Osnovni principi

Kao što smo već vidjeli, model vrši mapiranje iz prošlih vrijednosti u prostor izlaza. U nelinearnom slučaju, ovo mapiranje ima opštu strukturu (5.23):

$$\hat{y}(t|\theta) = g(Z^{t-1}, \theta) \quad (5.28)$$

gdje smo izostavili eksplicitnu zavisnost od vremena. Ukoliko nemamo uvid u osobine sistema, mi trebamo potražiti parametrizacije od g koje su dovoljno fleksibilne da pokriju "sve vrste razumnog ponašanja sistema". Ovo će nam dati nelinearnu strukturu black-box modela što predstavlja nelinearni ekvivalent linearnih modela crne kutije.

Struktura za opšte mapiranje : Regresori

Sada je familija struktura modela opisana sa (5.28) suviše opšta, i korisno je pisati g kao sastavljanje dva mapiranja: jednog koje uzima povećani broj prošlih observacija Z^t i mapira ih u konačno dimenzionalni vektor $\mathbf{f}(t)$ fiksne dimenzije, i jedan

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

koji uzima ovaj vektor u prostor izlaza:

$$g(Z^{t-1}, \theta) = g(\varphi(t), \theta) ; \quad (5.29)$$

gdje

$$\varphi(t) = \varphi(Z^{t-1}) \quad (5.30)$$

Mi ćemo, kao i prije nazivati ovaj vektor \mathbf{f} regresionim vektorom, a njegove komponente se nazivaju regresorima. Možemo dozvoliti također opštiji slučaj da je formiranje regresora i samo po sebi parametrizirano:

$$\varphi(t) = \varphi(Z^{t-1}, \theta) \quad (5.31)$$

koje skraćeno pišemo $\mathbf{f}(t, \mathbf{q})$. Zbog jednostavnosti dodatni argument \mathbf{q} bit će korišten eksplicitno samo kada je to bitno za diskusiju.

Izbor nelinearnog mapiranja u (5.28) je time dekomponiran na dva parcijalna problema za dinamičke sisteme:

1. Kako izabrati regresioni vektor $\mathbf{f}(t)$ od prošlih vrijednosti ulaza

MODELI ZA VREMENSKI VARIRAJUĆE I NELINEARNE SISTEME

i izlaza

2. Kako izabrati nelinearno mapiranje $g(\mathbf{f}, \mathbf{q})$ iz regresorskog prostora u prostor izlaza.

Izbor regresora je naravno zavistan od aplikacije. Tipično, regresori se biraju na isti način kao i za linearne modele: prošla mjerenja i možda i prošli izlazi iz modela kao i prošle greške predikcije kao u (4.11), (4.20) i (4.40). Mi ćemo sada, analogno tabeli govoriti o NARX, NARMAX, NOE strukturama modela. Za nelinearni slučaj crne kutije, najčešće se koriste samo mjerene (ne i procjenjene) vrijednosti u regresorima, napr. NFIR I NARX.

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Vidjeli smo da vremenski invarijantni model može biti opisan ili prenosnim funkcijama ili preko odgovarajućih impulsnih odziva. Sada ćemo diskutirati metode koje imaju za cilj da odrede ove funkcije direktnim tehnikama bez da se prvo izabere neki skup mogućih modela. Ovakvi metodi se često zovu neprameterski pošto oni ne koriste eksplicitno konačno dimenzionalni parametarski vektor u potrazi za najboljim opisom. Diskutovaćemo određivanje prenosne funkcije $G(q)$ sa ulaza na izlaz. Pri tome ćemo pretpostavljati da sistem radi u otvorenoj konturi tj. $\{u(t)\}$ i $\{v(t)\}$ su nezavisne. Zatvorene konture obično dovode do problema kod primjene neprameterskih metoda.

Analiza tranzijentnih odziva i korelaciona analiza

Analiza impulsnog odziva

Ako je sistem koji je opisan sa (2.8):

$$y(t) = G_0(q)u(t) + v(t) \quad (6.1)$$

podvrgnut impulsnom ulazu:

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

$$u(t) = \begin{cases} \alpha, & t = 0 \\ 0, & t \neq 0 \end{cases} \quad (6.2)$$

tada će izlaz biti:

$$y(t) = \alpha g_0(t) + v(t) \quad (6.3)$$

po definiciji od G_0 i impulsnim odzivom $\{g_0(t)\}$. . Ako je nivo šuma nizak, moguće je odrediti koeficijente impulsnog odziva $\{g_0(t)\}$ iz eksperimenta sa impulsnim ulazom. Procjena će biti:

$$\hat{g}(t) = \frac{y(t)}{\alpha} \quad (6.4)$$

i sa greškama $v(t)/\alpha$. . Jednostavna ideja je analiza impulsnog odziva. Osnovna mana ove identifikacione metode je da mnogi procesi ne dozvoljavaju impulsne ulaze takve amplitude da je greška $v(t)/\alpha$ zanemarljiva u poredjenju sa koeficijentima impulsnog odziva. Nadalje ovakav oblik ulaza može dovesti do toga da sistem ispolji nelinearne efekte koji bi poremetili linearizirano ponašanje, koje smo prepostavili za model.

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Analiza odziva na step

Slično, odziv na step:

$$u(t) = \begin{cases} \alpha, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

primjenjen na (6.1) daje izlaz:

$$y(t) = \alpha \sum_{k=1}^t g_0(k) + v(t) \quad (6.5)$$

Oдавде, procjene za $g_0(k)$ se mogu dobiti kao:

$$\hat{g}(t) = \frac{y(t) - y(t-1)}{\alpha} \quad (6.6)$$

sa greškom procjene od $[v(t) - v(t-1)]/\alpha$. Ako mi zaista želimo da odredimo koeficijente impulsnog odziva koristeći (6.6), imaćemo prilično velike greške u većini praktičnih primjena. Medjutim, ukoliko je cilj da se odrede neke osnovne karakteristike kao što su vrijeme kašnjenja, statičko pojačanje, i dominantne vremenske konstante (tj. model u obliku (4.50), onda odzivi na

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

na step (6.5) mogu vrlo dobro obezbjediti te informacije sa dovoljnim stepenom tačnosti. U stvari, dobro poznate metode podešenja regulatora kao što je Ziegler-Nicholsovo pravilo su bazirane na informacijama o modelu koje se dobiju iz odziva na step.

Na bazi grafikona step odziva, neke karakteristične vrijednosti se mogu grafički konstruisati, pa se zatim mogu koristiti da se odrede parametri u modelu datog reda.

Korelaciona analiza

Posmatrajmo opis modela (6.1):

$$y(t) = \sum_{k=1}^{\infty} g_0(k)u(t - k) + v(t) \quad (6.7)$$

Ako je ulaz kvazi stacionarna sekvenca [vidjeti (2.59)] sa:

$$\overline{E}u(t)u(t - \tau) = R_u(\tau)$$

$$\overline{E}u(t)v(t - \tau) \equiv 0 \quad (\text{rad u otvorenoj konturi}) \quad 25$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

tada je:

$$\overline{E}_{y(t)u(t-\tau)} = R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k)R_u(k-\tau) \quad (6.8)$$

Ako je ulaz izabran kao bijeli šum, tako da:

$$R_u(\tau) = \alpha \delta_{\tau 0}$$

tada

$$g_0(\tau) = \frac{R_{yu}(\tau)}{\alpha}$$

Procjena impulsnog odziva se tada dobije iz procjene $R_{yu}(\tau)$ na primjer:

$$\hat{R}_{yu}^N(\tau) = \frac{1}{N} \sum_{t=\tau}^N y(t)u(t-\tau) \quad (6.9)$$

Ako ulaz nije bijeli šum, možemo procjeniti:

$$\hat{R}_u^N(\tau) = \frac{1}{N} \sum_{t=\tau}^N u(t)u(t-\tau) \quad (6.10)$$

i riješiti:

$$\hat{R}_{yu}^N(\tau) = \sum_{k=1}^M \hat{g}(k) \hat{R}_u^N(k-\tau) \quad (6.11)$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

za $\hat{g}(k)$. Ako je ulaz dostupan za manipulaciju, poželjno je izabrati ga tako da (6.10) i (6.11) postaju lako rješivi. Oprema za generisanje takvih signala i rješavanje $\hat{g}(k)$ je komercijalno raspoloživa.

Inače, najprirodniji način da se procjeni $g(k)$ kada ulaz nije "potpuno bijel" je da se skrati (6.7) na n , i tretira se kao n -ti red FIR modela (4.46) sa parametarskim metodama najmanje sume kvadrata. Drugi pristup je da se filtriraju i ulaz i izlaz sa predfilterom koji čini ulaz što je moguće više "bijelim" ("ulazno predbijeljenje") a zatim izračuna korelaciona funkcija (6.9) za ove filtrirane sekvence.

Analiza frekventnog odziva

Testiranje sa sinusnim signalom

Osnovna fizikalna interpretacija prenosne funkcije $G(z)$ je da kompleksni broj $\overline{G(e^{j\omega})}$ nosi informaciju o tome šta se dešava ulaznom sinusoidalnom signalu (vidjeti (2.32) do (2.34)). Imamo sada za (6.1) da sa:

$$u(t) = \alpha \cos \omega t, \quad t = 0, 1, 2, \dots \quad (6.12) \quad 27$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

dobijemo:

$$y(t) = \alpha |G_0(e^{i\omega})| \cos(\omega t + \varphi) + v(t) + \text{transient} \quad (6.13)$$

gdje:

$$\varphi = \arg G_0(e^{i\omega}) \quad (6.14)$$

Ova osobina također upućuje i na jednostavan način da se odredi $G_0(e^{i\omega})$:

Sa ulazom (6.12) odrediti amplitudu i fazni pomak rezultirajućeg izlaznog kosinusnog signala, i izračunati procjenu $\hat{G}_N(e^{i\omega})$ na bazi ove informacije. Ponoviti za niz vrijednosti frekvencija u interesantnom frekventnom području.

Ovaj metod je poznat kao frekventna analiza i predstavlja jednostavan metod za dobijanje detaljne informacije o linearnom sistemu.

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Frekventna analiza sa korelacionom metodom

Sa komponentom šuma $v(t)$ prisutnom u (6.13), može biti teško tačno odrediti $|G_0(e^{j\omega})|$ i \mathbf{f} pomoću grafičkih metoda. Pošto je interesantna komponenta od $y(t)$ kosinusna funkcija poznate frekvencije, moguće je nju korelirati iz šuma na slijedeći način. Formirajmo sume:

$$I_c(N) = \frac{1}{N} \sum_{t=1}^N y(t) \cos \omega t, \quad I_s(N) = \frac{1}{N} \sum_{t=1}^N y(t) \sin \omega t \quad (6.15)$$

Ubacujući (6.13) u (6.15), i ignorišući tranzijentni član, daje:

$$\begin{aligned} I_c(N) &= \frac{1}{N} \sum_{t=1}^N \alpha |G_0(e^{j\omega})| \cos(\omega t + \varphi) \cos \omega t + \frac{1}{N} \sum_{t=1}^N v(t) \cos \omega t \\ &= \alpha |G_0(e^{j\omega})| \frac{1}{2} \frac{1}{N} \sum_{t=1}^N [\cos \varphi + \cos(2\omega t + \varphi)] \\ &\quad + \frac{1}{N} \sum_{t=1}^N v(t) \cos \omega t \quad (6.16) \\ &= \frac{\alpha}{2} |G_0(e^{j\omega})| \cos \varphi + \alpha |G_0(e^{j\omega})| \frac{1}{2} \frac{1}{N} \sum_{t=1}^N \cos(2\omega t + \varphi) \\ &\quad + \frac{1}{N} \sum_{t=1}^N v(t) \cos \omega t \end{aligned}$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Drugi član teži ka nuli kako N teži ka beskonačnosti, a tako i treći član ako $v(t)$ ne sadrži čistu periodičnu komponentu frekvencije ω . Ako $\{v(t)\}$ je stacionarni stohastički proces takav da je:

$$\sum_0^{\infty} \tau |R_v(\tau)| < \infty$$

tada varijansa trećeg člana od (6.16) opada sa $1/N$. Slično:

$$I_s(N) = -\frac{\alpha}{2} |G_0(e^{i\omega})| \sin \varphi + \alpha |G_0(e^{i\omega})| \frac{1}{2} \frac{1}{N} \sum_{r=1}^N \sin(2\omega r + \varphi) + \frac{1}{N} \sum_{r=1}^N v(r) \sin \omega r \quad (6.17)$$

ova dva izraza sugerišu slijedeće procjene od $|G_0(e^{i\omega})|$ i od \mathbf{f} :

$$|\hat{G}_N(e^{i\omega})| = \frac{\sqrt{I_c^2(N) + I_s^2(N)}}{\alpha/2} \quad (6.18a)$$

$$\hat{\varphi}_N = \arg \hat{G}_N(e^{i\omega}) = -\arctan \frac{I_s(N)}{I_c(N)} \quad (6.18b)$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Ponavljajući proceduru za niz vrijednosti frekvencija, dobije se dobra slika za $G_0(e^{i\omega})$ za frekventni domen od interesa. Oprema koja vrši ovakvu frekventnu analizu pomoću korelacionog metoda je i komercijalno dostupna.

Prednost ovog metoda je da se lako može dobiti i Bodeov dijagram sistema. Glavni nedostatak je da mnogi industrijski procesi ne dozvoljavaju sinusoidalne ulaze u svom normalnom radu. Nadalje eksperiment se mora ponavljati za različite frekvencije što može rezultirati u dugim vremenima trajanja eksperimenta.

Relacija sa Fourier-ovom analizom

Poredjenje (6.15) sa definicijom (2.37):

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^N y(t)e^{-i\omega t} \quad (6.19)$$

pokazuje da:

$$I_c(N) - iI_s(N) = \frac{1}{\sqrt{N}} Y_N(\omega) \quad (6.20)$$

NEPARAMETARSKI VREMENSKI I FREKVENTNI METODI

Kao i u (2.46) nalazimo da , za (6.12):

$$U_N(\omega) = \frac{\sqrt{N\alpha}}{2}, \quad \text{if } \omega = \frac{2\pi r}{N} \text{ for some integer } r \quad (6.21)$$

Možemo preurediti (6.18) u obliku:

$$\hat{G}_N(e^{i\omega}) = \frac{\sqrt{N}Y_N(\omega)}{N\alpha/2} \quad (6.22)$$

koji koristeći (6.21), daje da je:

$$\hat{G}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)} \quad (6.23)$$

ovdje je ω tačno frekvencija ulaza. Poredjenjem sa (2.53) , mi takodjer nalazimo (6.23) najboljom procjenom.

7. PARAMETARSKI METODI PROCJENA

Predpostavimo da je izabran skup modela kandidata, i da je paramteriziran kao struktura modela koristeći vektor parametara \mathbf{q} . Traženje najboljeg modela unutar skupa sada postaje problem određivanja ili procjene \mathbf{q} . Postoje mnogi načini organizacije takvog traženja kao i različiti pogledi na to šta se treba tražiti.

Principi metoda procjene parametara.

Mi smo sada u situaciji da smo izabrali neku strukturu modela \mathcal{M} , sa specifičnim modellima $\mathcal{M}(\mathbf{q})$ parametriziranim koristeći vektor parametara $\theta \in D_{\mathcal{M}} \subset \mathbb{R}^d$. Skup modela definisan na ovaj način je:

$$\mathcal{M}^* = \{\mathcal{M}(\theta) | \theta \in D_{\mathcal{M}}\} \quad (7.1)$$

Prisjetimo se da svaki model predstavlja način procjenjivanja budućih izlaza. Prediktor može biti i linearni filter u obliku:

$$\mathcal{M}(\theta) : \hat{y}(t|\theta) = W_y(q, \theta)y(t) + W_u(q, \theta)u(t) \quad (7.2)$$

Ovo može korespondirati sa predikcijom jednog koraka unaprijed³³

PARAMETARSKI METODI PROCJENA

za opis sistema u obliku:

$$y(t) = G(q, \theta)u(t) + H(q, \theta)e(t) \quad (7.3)$$

u kojem slučaju:

$$W_y(q, \theta) = [1 - H^{-1}(q, \theta)], \quad W_u(q, \theta) = H^{-1}(q, \theta)G(q, \theta) \quad (7.4)$$

Prediktor može također biti nelinearni filter, u kojem slučaju mi ga pišemo kao opštu funkciju prošlih podataka Z^{t-1} .

$$\mathcal{M}(\theta) : \hat{y}(t|\theta) = g(t, Z^{t-1}; \theta) \quad (7.5)$$

Model $\mathbf{M}(\mathbf{q})$ može također sadržavati pretpostavke o karakteru pridruženih predikcionih grešaka, kao njihove varijanse ($\mathbf{I}(\mathbf{q})$) ili njihove raspodjele vjerovatnoća (PDF $f_e(x, \theta)$)

Mi smo također prikupili, ili ćemo prikupiti i blok podataka sa sistema:

$$Z^N = [y(1), u(1), y(2), u(2), \dots, y(N), u(N)] \quad (7.6)$$

Problem sa kojim se suočavamo je da se odlučimo kako ćemo koristiti podatke sadržane u Z^N da izaberemo pravu vrijednost³⁴

PARAMETARSKI METODI PROCJENA

vektora parametara $\hat{\theta}_N$, i time i pravi član $\mathcal{M}(\hat{\theta}_N)$ u skupu \mathcal{M}^* .

Formalno govoreći, mi treba da odredimo mapiranje iz podataka Z^N u skup $D_{\mathcal{M}}$:

$$Z^N \rightarrow \hat{\theta}_N \in D_{\mathcal{M}} \quad (7.7)$$

Takvo mapiranje je metod procjene parametara.

Evaluacija modela kandidata.

Mi tražimo test pomoću kojeg se mogu evaluirati sposobnosti različitih modela da opišu posmatrane podatke. Ranije smo naglasili da je osnovna funkcija modela upravo njegov aspekt predikcije, tako da ćemo ocjenjivati i njegov kvalitet sa ovog aspekta. Neka je predikciona greška data nekim modelom $\mathcal{M}(\theta_*)$

data sa :

$$e(t, \theta_*) = y(t) - \hat{y}(t|\theta_*) \quad (7.8)$$

Kada je skup podataka Z^N poznat, ove se greške mogu sračunati za $t=1,2,\dots,N$.

“Dobar” model je onaj koji je dobar u predikciji, tj. onaj koji proizvodi male greške predikcije kada se primjeni na opservirane podatke. Primjetimo da postoji značajna fleksibilnost u selekciji

PARAMETARSKI METODI PROCJENA

različitih prediktorskih funkcija, i ovo daje odgovarajuću slobodu u definisanju “dobrih” modela u terminima performanse predikcije.

Vodeći princip za procjenu parametara je:

Na bazi Z^t mi možemo izračunati predikcionu grešku $\varepsilon(t, q)$ koristeći (7.8). U trenutku $t=N$, izabрати $\hat{\theta}_N$ tako da predikcione greške $\varepsilon(t, \hat{\theta}_N)$, $t = 2, \dots, N$, postanu što je moguće manje (7.9).

Pitanje je kako definirati šta malo treba da znači. U nastavku ćemo opisati dva načina: jedan je da formiramo skalarnu normu ili funkciju kriterija koja mjeri veličinu ε . Drugi pristup je da tražimo da $\varepsilon(t, \hat{\theta}_N)$ bude nekolerirano sa datom sekvencom podataka. Ovo korespondira sa zahtjevom da izvjesne “projekcije” od $\varepsilon(t, \hat{\theta}_N)$ su nula.

Minimizacija predikcionih grešaka

Sekvenca predikcionih grešaka u (7.8) se može posmatrati kao vektor u \mathbf{R}^N . Veličina ovog vektora se može mjeriti koristeći bilo koju normu u \mathbf{R}^N kvadratičnu ili nekvadratičnu. Ovo ostavlja značajan broj izbora. Mi ćemo donekle ograničiti slobodu, posmatrajući samo slijedeći način evaluacije “kako velika” je

PARAMETARSKI METODI PROCJENA

sekvenca predikcionih grešaka: Neka je sekvenca predikcionih grešaka filtrirana kroz stabilni linearni filter $L(q)$:

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta), \quad 1 \leq t \leq N \quad (7.10)$$

Sada ćemo koristiti slijedeću normu:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \ell(\varepsilon_F(t, \theta)) \quad (7.11)$$

gdje je $\ell(\cdot)$ skalarna (tipično pozitivna) funkcija.

Funkcija $V_N(\theta, Z^N)$ je , za dato Z^N , je dobro definirana skalarna funkcija parametara modela q . Ona je prirodna mjera validnosti modela $\mathcal{M}(\theta)$. Procjena $\hat{\theta}_N$ je tada definirana mnimizacijom (7.11):

$$\hat{\theta}_N = \hat{\theta}_N(Z^N) = \arg \min_{\theta \in D_M} V_N(\theta, Z^N) \quad (7.12)$$

Ovdje $\arg \min$ znači : "minimiziranje argumenta funkcije". Ako minimum nije unikal, tada $\arg \min$ će označavati skup minimizirajućih argumenata. Mapiranje (7.7) je tada definirano implicitno sa (7.12). Ovaj način procjene q sadrži mnogo poznatih i korištenih procedura. Mi ćemo za njih koristiti opšti naziv: **identifikacione metode sa predikcionom greškom**(PEM).

PARAMETARSKI METODI PROCJENA

za familiju pristupa koji korespondiraju sa (7.12). Specifični metodi, sa specifičnim imenima koji im se dodjeljuju, se onda izvode kao specijalni slučajevi od (7.12), zavisno od izbora je $l(\cdot)$, izbora predfiltera $L(\cdot)$, izbora strukture modela, i u nekim slučajevima, od izbora metode za realizaciju minimizacije.

Izbor L

Efekat filtera L je da dozvoli dodatnu slobodu da se može reagovati na ne-momentalne osobine grešaka predikcije. Jasno je da, ukoliko je prediktor linearan i vremenski invarijantan, a y i u su skalari, tada je rezultat filtriranja ε , isti bilo da se prvo filtriraju ulazno/izlazni podaci a onda se primjene prediktori.

Jasno je medjutim da efekti visoko frekventnih smetnji, koji nisu bitni za problem modeliranja kao i sporo promjenljivi članovi se mogu otkloniti primjenom L. L dakle djeluje kao frekventni težinski faktor. Primjetimo još jedan specifičan aspekt filtriranja (7.10). Ako se koristi model (7.3), filtrirana greška $\varepsilon_F(t, \theta)$ je data sa:

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta) = [L^{-1}(q)H(q, \theta)]^{-1} [y(t) - G(q, \theta)u(t)] \quad (7.13)$$

PARAMETARSKI METODI PROCJENA

Efekat predfiltriranja je dakle identičan sa promjenom modela šuma sa $H(q, \mathbf{q})$ na:

$$\overline{H}_L(q, \theta) = L^{-1}(q)H(q, \theta) \quad (7.14)$$

Kada opisujemo i analiziramo metode koji koriste opšte modele šuma kod linearnih sistema, mi ćemo se obično ograničiti na $L(q) = 1$, pošto opcija predfiltriranja se ostvaruje slobodom u izboru $H(q, \mathbf{q})$.

Izbor I

Za izbor $I(\cdot)$, prvi kandidat bi bila kvadratična norma:

$$\ell(\varepsilon) = \frac{1}{2}\varepsilon^2 \quad (7.15)$$

i ovo je zaista standardni izbor. Pitanje robusnosti protiv loših podataka može pak zahtijevati i druge norme. Može se također pretpostaviti situacija da "najbolja" norma nije unaprijed poznata, tako da je razumno paramterizirati samu normu:

$$\ell(\varepsilon, \theta) \quad (7.16)$$

Često parametrizacija norme je nezavisna od modela parametrizacije

PARAMETARSKI METODI PROCJENA

$$\theta = \begin{bmatrix} \theta' \\ \alpha \end{bmatrix} : \ell(\varepsilon(t, \theta), \theta) = \ell(\varepsilon(t, \theta'), \alpha) \quad (7.17)$$

Vremenski varijabilne norme

Može se desiti da su mjerenja u različitim trenucima vremena sa varirajućom pouzdanošću. Razlog za ovo može biti u tome što se nivo šuma mijenja u vremenu ili što su neka mjerenja manje reprezentativna za osobine sistema. U takvim slučajevima mi smo motivirani da i norma \mathbf{I} bude vremenski varijabilna, tj.

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \ell(\varepsilon(t, \theta), \theta, t) \quad (7.18)$$

Na ovaj način, manje pouzdana mjerenja se mogu udružiti sa manjim težinama u kriteriju.

Mi ćemo često raditi sa kriterijem gdje je otežavanje ostvareno eksplicitno sa težinskom funkcijom $\beta(N, t)$:

$$V_N(\theta, Z^N) = \sum_{t=1}^N \beta(N, t) \ell(\varepsilon(t, \theta), \theta) \quad (7.19)$$

PARAMETARSKI METODI PROCJENA

Za fiksno N , N zavisnost od $\beta(N,t)$ nije bitna. Medjutim ako poredimo procjene $\hat{\theta}_N$ za različite N , kao naprimjer kod rekurzivne identifikacije, postaje interesantno diskutovati kako $\beta(N,t)$ varira sa N .

Multivarijabilni sistemi

Za multivarijabilne sisteme, ekvivalent kvadratičnom kriteriju je:

$$\ell(\varepsilon) = \frac{1}{2} \varepsilon^T \Lambda^{-1} \varepsilon \quad (7.27)$$

za neku simetričnu, pozitivno semidefinitnu matricu Λ dimenzije $p \times p$, koja otežava zajedno relativnu važnost komponenata od ε . Koji je najbolji izbor norme za Λ . Kao i u slučaju (7.16) vektor parametara \mathbf{q} se može proširiti da uključi komponente Λ , i funkcija \mathbf{l} , će biti odgovarajuća funkcija od \mathbf{q} .

Kao varijantu od kriterija (7.11), gdje je skalar $\mathbf{l}(\varepsilon)$ bio formiran za svako t , mi bi mogli prvo formirati $p \times p$ matricu:

$$Q_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \varepsilon(t, \theta) \varepsilon^T(t, \theta) \quad (7.28)$$

i uzeti da je kriterij skalarna funkcija ove matrice:

PARAMETARSKI METODI PROCJENA

$$V_N(\theta, Z^N) = h(Q_N(\theta, Z^N)) \quad (7.29)$$

Kriterij (7.27) je onda dobijen kao :

$$h(Q) = \frac{1}{2} \text{tr}(Q \Lambda^{-1}) \quad (7.30)$$

Linearna regresija i metod najmanjih kvadrata

Linearne regresije

Vidjeli smo iz ranijih izlaganja da su strukture modela linearne regresije vrlo korisne u opisu osnovnih linearnih i nelinearnih sistema. Linearna regresija koristi prediktor (5.67):

$$\hat{y}(t|\theta) = \varphi^T(t)\theta + \mu(t) \quad (7.31)$$

koji je linearan po \mathbf{q} . ovdje \mathbf{f} je vektor regresora, tj regresioni vektor. Podsjetimo se da za ARX strukturu (4.7) imamo:

$$\varphi(t) = [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n_a) \ u(t-1) \ \dots \ u(t-n_b)]^T \quad (7.32)$$

PARAMETARSKI METODI PROCJENA

U (7.31) $\mathbf{m}(t)$ je poznat kao vektor zavistan od podataka. Zbog jednostavnosti označavanja, uzećemo da je $\mathbf{m}(t) = 0$ u nastavku.

Kriterij najmanjih kvadrata

Sa (7.31) predikciona greška postaje:

$$\varepsilon(t, \theta) = y(t) - \varphi^T(t)\theta$$

a funkcija kriterija, koja rezultira iz (7.10) i (7.11), sa $L(q) = 1$, i $I(\varepsilon) = \frac{1}{2} \varepsilon^2$ je :

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} [y(t) - \varphi^T(t)\theta]^2 \quad (7.33)$$

Ovo je kriterij najmanjih kvadrata za linearnu regresiju (7.31). Osobina ovog kriterija je da je kvadratna funkcija po \mathbf{q} . Zbog toga, se može minimizirati analitički, što daje, pod uslovom da postoji inverzna vrijednost:

$$\hat{\theta}_N^{\text{LS}} = \arg \min V_N(\theta, Z^N) = \left[\frac{1}{N} \sum_{t=1}^N \varphi(t)\varphi^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \varphi(t)y(t) \quad (7.34)$$

je procjena najmanjih kvadrata (LSE).

PARAMETARSKI METODI PROCJENA

Uvedimo $d \times d$ matricu:

$$R(N) = \frac{1}{N} \sum_{t=1}^N \varphi(t) \varphi^T(t) \quad (7.35)$$

i d dimenzionalni vektor kolonu:

$$f(N) = \frac{1}{N} \sum_{t=1}^N \varphi(t) y(t) \quad (7.36)$$

U slučaju (7.32), $\mathbf{f}(t)$ sadrži zakašnjele ulazne i izlazne varijable, i ulazi veličina (7.35) i (7.36) će biti u obliku:

$$[R(N)]_{ij} = \frac{1}{N} \sum_{t=1}^N y(t-i) y(t-j), \quad 1 \leq i, j \leq n_a$$

i slično sume od $u(t-r) \cdot u(t-s)$ or $u(t-r) \cdot y(t-s)$ druge ulaze od $R(N)$. To jest, sastojće se od procjena kovarijansnih funkcija od $\{y(t)\}$ i $\{u(t)\}$. LSE se sada može izračunati koristeći samo procjene i zbog toga je u relaciji sa koralacionom analizom.

PARAMETARSKI METODI PROCJENA

Osobine LSE

Metod najmanjih kvadrata je specijalni slučaj identifikacionog metoda sa greškom predikcije (7.12). Pogledajmo neke heurističke karakteristike ovog metoda.

Predpostavimo da su observirani podaci ustvari bili generisani sa

$$y(t) = \varphi^T(t)\theta_0 + v_0(t) \quad (7.37)$$

za neku sekvencu $\{v_0(t)\}$. Možemo posmatrati \mathbf{q}_0 kao "pravu vrijednost" vektora parametara.

Kao i u izrazima (1.14) i (1.15) nalazimo da:

$$\lim_{N \rightarrow \infty} \hat{\theta}_N^{\text{LS}} - \theta_0 = \lim_{N \rightarrow \infty} R^{-1}(N) \frac{1}{N} \sum_{t=1}^N \varphi(t)v_0(t) = (R^*)^{-1}f^*,$$
$$R^* = \overline{E}\varphi(t)\varphi^T(t), \quad f^* = \overline{E}\varphi(t)v_0(t) \quad (7.38)$$

pod uslovom da su v_0 i \mathbf{f} kvazi stacionarni. Da bi LSE bio konzistentan tj. da $\hat{\theta}_N^{\text{LS}}$ konvergira ka \mathbf{q}_0 , potrebno je da:

PARAMETARSKI METODI PROCJENA

1. R^* je nesingularno. Ovo će biti obezbjedjeno sa ulaznim osobinama , kao u (1.17) i (1.18)
2. $f^* = 0$. Ovo će biti slučaj kada je , bilo:

a. $\{v_0(t)\}$ sekvenca nezavisnih slučajnih varijabli sa nultom srednjom vrijednošću (bijeli šum). Tada $v_0(t)$ neće zavisiti od toga šta se desilo do vremena $t-1$ i prema tome $E\varphi(t)v_0(t) = 0$.

b. ulazna sekvenca $\{u(t)\}$ je nezavisna od sekvence sa nultom srednjom vrijednošću $\{v_0(t)\}$ i $n_a = 0$ u (7.32). Tada $f(t)$ sadrži samo u-članove i time $E\varphi(t)v_0(t) = 0$.

Otežana suma najmanjih kvadrata

Kao i u (7.8) i (7.9) , različitim mjerenjima se mogu doznačiti različite težine u kriteriju najmanje sume kvadrata:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \alpha_t [y(t) - \varphi^T(t)\theta]^2 \quad (7.39)$$

ili:

$$V_N(\theta, Z^N) = \sum_{t=1}^N \beta(N, t) [y(t) - \varphi^T(t)\theta]^2 \quad (7.40)$$

PARAMETARSKI METODI PROCJENA

Izraz za rezultirajuću procjenu je analogan sa onim u (7.34):

$$\hat{\theta}_N^{LS} = \left[\sum_{t=1}^N \beta(N, t) \varphi(t) \varphi^T(t) \right]^{-1} \sum_{t=1}^N \beta(N, t) \varphi(t) y(t) \quad (7.41)$$

Multivarijabilni slučaj

Ako je izlaz $y(t)$ p -vektor i norma (7.27) se koristi, LS kriterij poprima oblik:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} [y(t) - \varphi^T(t)\theta]^T \Lambda^{-1} [y(t) - \varphi^T(t)\theta] \quad (7.42)$$

Ovo daje procjenu:

$$\hat{\theta}_N^{LS} = \left[\frac{1}{N} \sum_{t=1}^N \varphi(t) \Lambda^{-1} \varphi^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \varphi(t) \Lambda^{-1} y(t) \quad (7.43)$$

U slučaju da koristimo specifičnu parametrizaciju (4.56) sa \mathbf{q} kao $r \times p$ matricom;

$$\hat{y}(t|\theta) = \theta^T \varphi(t) \quad (7.44) \quad 47$$

PARAMETARSKI METODI PROCJENA

LS kriterij postaje:

$$V_N(\boldsymbol{\theta}, Z^N) = \frac{1}{N} \sum_{t=1}^N \|y(t) - \boldsymbol{\theta}^T \boldsymbol{\varphi}(t)\|^2 \quad (7.45)$$

sa procjenom:

$$\hat{\boldsymbol{\theta}}_N^{LS} = \left[\frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t) \boldsymbol{\varphi}^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \boldsymbol{\varphi}(t) y^T(t) \quad (7.46)$$

Da bi odredili $r \times p$ procjenjeno $\hat{\boldsymbol{\theta}}_N$, dovoljno je invertovati $r \times r$ matricu. U izrazu (7.43) \mathbf{q} je $r \times 1$ vektor i matrična inverzija se provodi na $r \times r$ matrici.

Procjena modela u prostoru stanja pomoću tehnike najmanjih kvadrata

Linearni sistem se može uvijek predstaviti u prostoru stanja u obliku kao u (4.84):

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + Du(t) + v(t) \end{aligned} \quad (7.54)$$

sa bijelim šumovima w i v . Alternativno, možemo predstaviti ulazno/izlaznu dinamiku kao u (4.80):

PARAMETARSKI METODI PROCJENA

$$\begin{aligned}x(t + 1) &= Ax(t) + Bu(t) \\y(t) &= Cx(t) + Du(t) + v(t)\end{aligned}\tag{7.55}$$

gdje šum na izlazu v , se može i obojiti. Da bi se procjenio takav model matrice se mogu parametrizirati ili na bazi fizikalnih razmatranja ili posmatrajući sistem kao crne kutije u kanoničkim opisima.

Predpostavićemo da nemamo uvida u neku specifičnu strukturu i da ćemo procjeniti bilo koje matrice A, B, C i D koje daju dobar opis ulazno/izlaznog ponašanja sistema. Pošto postoji beskonačan broj takvih matrica koje opisuju isti sistem, moraćemo da fiksiramo koordinatnu bazu realizacija u prostoru stanja. Predpostavimo za trenutak, da su u i y ne samo mjerljive, nego su takodjer i sekvence vektora stanja x . Ovo će fiksirati realizaciju koordinatne baze u prostoru stanja. Sada sa poznatim u , y i x , model (7.54) postaje linearna regresija: nepoznati parametri, svi ulazi u matrice u svim matricama, se mješaju sa mjerenim signalima u linearnim kombinacijama. Da bi ovo jasno vidjeli, neka su:

PARAMETARSKI METODI PROCJENA

$$Y(t) = \begin{bmatrix} x(t+1) \\ y(t) \end{bmatrix}, \quad \Theta = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$
$$\Phi(t) = \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}, \quad E(t) = \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}$$

Tada (7.54) se može pisati kao:

$$Y(t) = \Theta\Phi(t) + E(t) \quad (7.56)$$

Iz ovoga, svi elementi matrice u Θ se mogu procjeniti pomoću jednostavnog metoda najmanjih kvadrata (što u slučaju Gaussovskog šuma i poznate kovarijantne matrice koincidira sa metodom maksimalne sličnosti (maximum likelihood)). Kovarijantna matrica za $E(t)$ se također može procjeniti kao jednostavna suma kvadrata reziduala modela. Ovo će dati kovarijantne matrice kao i kros kovarijantnu matricu za w i v . Ove matrice će nam dozvoliti da izračunamo također i Kalmanov filter za (7.54). Primjetimo da ovo sve vrijedi i za multivarijabilne sisteme, tj. kada su ulazni i izlazni signali vektori. Problem je sada kako dobiti sekvencu vektora stanja x .

Neka je sistem dat sa predstavljanjem preko impulsnog odziva:

PARAMETARSKI METODI PROCJENA

$$y(t) = \sum_{j=0}^{\infty} [h_u(j)u(t-j) + h_e(j)e(t-j)] \quad (7.57)$$

gdje je u ulaz a e su inovacije. Neka je $\hat{y}(t|t-k)$ formalni k -koraka unaprijed prediktor definisan na taj način da je izbrisan doprinos u $y(t)$ od $e(j)$, $u(j)$, $j=t, \dots, t-k+1$:

$$\hat{y}(t|t-k) = \sum_{j=k}^{\infty} [h_u(j)u(t-j) + h_e(j)e(t-j)] \quad (7.58)$$

Nije napravljen pokušaj da se predvide ulazi $u(j)$: $j=t, \dots, t-k+1$, iz prošlih podataka. Definišimo:

$$\hat{Y}_r(t) = \begin{bmatrix} \hat{y}(t|t-1) \\ \vdots \\ \hat{y}(t+r-1|t-1) \end{bmatrix} \quad (7.59a)$$

$$\hat{Y} = \begin{bmatrix} \hat{Y}_r(1) & \dots & \hat{Y}_r(N) \end{bmatrix} \quad (7.59b)$$

Sada je slijedeće tačno kako $N \rightarrow \infty$

PARAMETARSKI METODI PROCJENA

1. Sistem (7.57) ima minimalni opis u prostoru stanja n -tog reda, ako i samo ako rang \hat{Y} je jednak n za sve $r \geq n$.

2. Vektor stanja bilo koje minimalne realizacije u inovacionim formama se može izahtrati kao linearna kombinacija od \hat{Y} koja formira osnovni red za \hat{Y} , tj.

$$x(t) = L\hat{Y}_r(t) \quad (7.60)$$

gdje $n \times p$ matrica L je takva da $L\hat{Y}$ spanira \hat{Y} , a p je dimenzija izlaznog vektora $y(t)$.

Jasno je da gornje činjenice nam dozvoljavaju da nadjemo pogodan vektor stanja iz podataka. Jedini preostali problem je da se procjeni k -koraka unaprijed prediktore. Korektan prediktor

$\hat{y}(t+k-1|t-1)$ je dat sa (7.58). Inovacija $e(j)$ se može pisati kao linearna kombinacija prošlih ulazno/izlaznih podataka. Prediktor prema tome može biti izražen kao linearna funkcija od $u(i)$, $y(i)$, $i \leq t-1$. Iz praktičnih razloga prediktor se aproksimira tako da samo zavisi od od fiksnog i konačnog broja prošlih podataka, kao s_1 prošlih izlaza i s_2 prošlih ulaza. To znači da će poprimiti oblik:

PARAMETARSKI METODI PROCJENA

$$\hat{y}(t+k-1|t-1) = \alpha_1 y(t-1) + \dots + \alpha_{s_1} y(t-s_1) \\ + \beta_1 u(t-1) + \dots + \beta_{s_2} u(t-s_2) \quad (7.61)$$

Ovaj prediktor može onda efikasno biti određen od strane druge projekcije najmanjih kvadrata, na ulazno/izlazne podatke.

Dakle, uspostavimo model:

$$y(t+k-1) = \theta_k^T \varphi_s(t) + \gamma_k^T U_\ell(t) + \varepsilon(t+k-1) \quad (7.62)$$

ili radeći sa svih r prediktora simultano:

$$Y_r(t) = \Theta \varphi_s(t) + \Gamma U_\ell(t) + E(t) \quad (7.63)$$

Ovdje:

$$\varphi_s(t) = \left[y^T(t-1) \dots y^T(t-s_1) \quad u^T(t-1) \dots u^T(t-s_2) \right]^T \quad (7.64a)$$

$$U_\ell(t) = \left[u^T(t) \dots u^T(t+\ell-1) \right]^T \quad (7.64b)$$

$$Y_r(t) = \left[y^T(t) \dots y^T(t+r-1) \right]^T \quad (7.64c)$$

$$\Theta = [\theta_1 \dots \theta_r]^T, \quad \Gamma = [\gamma_1 \dots \gamma_r]^T \quad (7.64d)$$

$$E(t) = \left[\varepsilon^T(t) \dots \varepsilon^T(t+r-1) \right]^T \quad (7.64e)$$

PARAMETARSKI METODI PROCJENA

Nadalje, l je broj, tipično jednak r , ulaznih vrijednosti čiji uticaj na $Y_r(t)$ se također mora uzeti u obzir. Sada Θ i Γ u (7.63) se mogu procjeniti koristeći najmanje kvadrate, za date $\hat{\Theta}_N$ i $\hat{\Gamma}_N$.
k-koračni prediktori unaprijed su dati sa:

$$\hat{Y}_r(t) = \hat{\Theta}_N \varphi_s(t) \quad (7.65)$$

Za dovoljno velike s , ovo će dati dobru aproksimaciju pravih prediktora.

Metod se sastoji od slijedećih koraka:

1. Izabrati s_1, r i l i formirati $\hat{Y}_r(t)$ u (7.65) i Y kao u (7.59).
2. Procjeniti rang n od Y i odrediti L u (7.60) tako da $x(t)$ korespondira sa dobro kondicioniranom bazom za to.
3. Procjeniti A, B, C, D i kovarijansne matrice šuma primjenjujući LS metod na linearnu regresiju (7.56).

Ovaj pristup daje vrlo korisne algoritme za procjenu modela, i pogodan je i za multivarijabilne sisteme. Algoritam također ima i numerički vrlo pouzdane implementacije, i tipično proizvodi procjenjene modele sa dobrim kvalitetom.

PARAMETARSKI METODI PROCJENA

Ako se želi, kvalitet modela se može poboljšati koristeći model kao početnu procjenu za metode predikcione greške (7.12). Tada model prvo treba transformisati u pogodnu paramterizaciju.

Metodi instrumentalne varijable

Instrumentalne varijable

Posmatrajmo ponovno linearni regresioni model (7.31)

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \quad (7.114)$$

Podsjetimo se da ovaj model sadrži nekoliko tipičnih modela linearnih i nelinearnih sistema. Procjena najmanjih kvadrata od \mathbf{q} je data sa (7.34) i može se izraziti kao:

$$\hat{\theta}_N^{\text{LS}} = \text{sol} \left\{ \frac{1}{N} \sum_{t=1}^N \varphi(t) [y(t) - \varphi^T(t)\theta] = 0 \right\} \quad (7.115)$$

Predpostavimo sada da se podaci mogu opisati kao u (7.37) sa:

$$y(t) = \varphi^T(t)\theta_0 + v_0(t) \quad (7.116)$$

LSE od $\hat{\theta}_N$ neće u tipičnim slučajevima težiti ka \mathbf{q}_0 pošto 55

PARAMETARSKI METODI PROCJENA

postoji korelacija između $v_0(t)$ i $\mathbf{f}(t)$. Pokušajmo sa opštim korelacionim vektorom $\mathbf{z}(t)$ u (7.115). U skladu sa opštom terminologijom kod identifikacije sistema, mi ćemo nazivati ovakvu aplikaciju (7.110) na linearnu regresiju kao metod instrumentalne varijable (IV). Elementi od \mathbf{z} se nazivaju instrumenti ili instrumentalne varijable. Ovo daje:

$$\hat{\theta}_N^{IV} = \text{sol} \left\{ \frac{1}{N} \sum_{t=1}^N \zeta(t) [y(t) - \varphi^T(t)\theta] = 0 \right\} \quad (7.117)$$

ili:

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N} \sum_{t=1}^N \zeta(t)\varphi^T(t) \right]^{-1} \frac{1}{N} \sum_{t=1}^N \zeta(t)y(t) \quad (7.118)$$

pod uslovom da postoji inverzna vrijednost. Da bi $\hat{\theta}_N$ težilo ka \mathbf{q}_0 za veliko N , vidimo iz (7.117) da tada $(1/N) \sum_{t=1}^N \zeta(t)v_0(t)$ treba da teži nuli. Da bi metod (7.117) bio uspješno primjenjiv na sistem (7.116), mi ćemo zahtjevati da instrumentalna varijabla $\mathbf{z}(t)$ ima slijedeće osobine (zamjenjujući srednje vrijednosti uzoraka sa očekivanjima):

$$\overline{E} \zeta(t)\varphi^T(t) \text{ be nonsingular} \quad (7.119)$$

$$\overline{E} \zeta(t)v_0(t) = 0 \quad (7.120)$$

PARAMETARSKI METODI PROCJENA

Ili rečeno riječima, instrumenti moraju biti korelirani sa regresionim varijablama ali nekorelirani sa šumom. Posmatrajmo sada moguće izbore instrumenata koji mogu biti podvrgnuti izrazima (7.119) i (7.120).

Izbor instrumenata

Predpostavimo da je (7.114) ARX model:

$$\begin{aligned} y(t) + a_1 y(t-1) + \dots + a_{n_a} y(t-n_a) \\ = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b) + v(t) \end{aligned} \quad (7.121)$$

Predpostavimo takodjer da pravi opis (7.116) korespondira sa (7.121) sa koeficijentima koji su indeksirani sa "nulom".

Prirodna ideja je da se generišu instrumenti slični sa (7.121) tako da se osigura (7.119), ali istovremeno ne dozvoliti im da na njih utiču $\{v_0(t)\}$. Ovo vodi ka:

$$\begin{aligned} \xi(t) = K(q) [-x(t-1) \quad -x(t-2) \dots \\ -x(t-n_a) \quad u(t-1) \dots u(t-n_b)]^T \end{aligned} \quad (7.122)$$

PARAMETARSKI METODI PROCJENA

gdje K je linearni filter i $x(t)$ se generiše od ulaza kroz linearni sistem:

$$N(q)x(t) = M(q)u(t) \quad (7.123)$$

Ovdje:

$$\begin{aligned} N(q) &= 1 + n_1q^{-1} + \dots + n_{n_n}q^{-n_n} \\ M(q) &= m_0 + m_1q^{-1} + \dots + m_{n_m}q^{-n_m} \end{aligned} \quad (7.124)$$

Većina instrumenata koji se koriste u praksi su generisana na ovaj način. Očito $\mathbf{z}(t)$ je dobijeno iz prošlih ulaza linearnim filtriranjem i može se konceptualno pisati kao:

$$\zeta(t) = \zeta(t, u^{t-1}) \quad (7.125)$$

Ako je ulaz generisan u otvorenoj konturi tako da ne zavisi od šuma $v_0(t)$ u sistemu, tada jasno je da (7.120) vrijedi. Pošto su i \mathbf{f} -vektor i \mathbf{z} -vektor generisani od iste ulazne sekvence (\mathbf{f} još dodatno sadrži efekte od v_0), može se očekivati da (7.119) će vrijediti i u opštem slučaju.

Jednostavan izbor instrumenata je da se prvo primjeni LS metod na (7.121) a zatim koristi LS model estimacije za N i M u (7.123)

Instrumenti se onda biraju kao i u (7.122) sa $K(q)=1$.

10. IZRAČUNAVANJE PROCJENA

U prethodnom poglavlju smo vidjeli tri osnovne procedure za procjenu parametara:

1. Pristup greške procjene kod kojeg neka funkcija $V_N(\theta, Z^N)$ se minimizira po \mathbf{q} .
2. Korelacioni pristup, kod kojeg sa neka jednačina $f_N(\theta, Z^N) = 0$ rješava po \mathbf{q} .
3. Pristup podprostora da se procjene modeli u prostoru stanja.

U nastavku ovog poglavlja ćemo vidjeti kako se ovi problemi najbolje rješavaju numeričkim metodama. U trenutku N , kada je skup podataka Z^N poznat, funkcije V_N i f_N su obične funkcije konačno dimenzionalnog realnog vektora parametara \mathbf{q} . Rješavanje problema se zbog toga svodi na standardna pitanja nelinearnog programiranja i numeričke analize. Ipak, vrijedi razmatrati probleme u okviru procjene vrijednosti parametara, jer ovo dodaje neku strukturu ovim funkcijama.

Metod podprostora (7.66) se može implementirati na nekoliko različitih načina, i ima mnogo opcija.

10. IZRAČUNAVANJE PROCJENA

Linearne regresije i najmanji kvadrati

Za linearne regresije, predikcija je data sa:

$$\hat{y}(t|\theta) = \varphi^T(t)\theta \quad (10.1)$$

Pristup predikcione greške sa kvadratnom normom primjenjen na (10.1) daje metod najmanjih kvadrata. Minimizacija elementa $\hat{\theta}_N^{\text{LS}}$ se onda može pisati kao u (7.34):

$$\hat{\theta}_N^{\text{LS}} = R^{-1}(N)f(N) \quad (10.2)$$

sa

$$R(N) = \frac{1}{N} \sum_{t=1}^N \varphi(t)\varphi^T(t) \quad (10.3)$$

$$f(N) = \frac{1}{N} \sum_{t=1}^N \varphi(t)y(t) \quad (10.4)$$

Alternativa je da se posmatra $\hat{\theta}_N^{\text{LS}}$ kao rješenje od:

$$R(N)\hat{\theta}_N^{\text{LS}} = f(N) \quad (10.5)$$

Ove jednačine su poznate kao normalne jednačine. Primjetimo da i osnovna jednačina (7.118) za IV metod je analogna (10.5)

10. IZRAČUNAVANJE PROCJENA

tako da većinu onoga što će biti izvedeno za LS metod se može takodjer primjeniti i na IV metod sa nekim adaptacijama. Matrica koeficijenata $R(N)$ u (10.5) može biti loše kondicionirana, naročito ako je njena dimenzija velika. Postoje metodi da se nadje $\hat{\theta}_N$ koji se mnogo bolje numerički ponašaju, koji nemaju normalne jednačine kao startnu tačku. Kod ovih metoda se ne formira matrica $R(N)$ pošto ona sadrži proizvode originalnih podataka. Umjesto toga, matrica R se formira sa osobinom:

$$RR^T = R(N)$$

Zbog toga se ova klasa metoda naziva “kvadratno-korjenskim algoritmima”, mada ovaj termin nije adekvatan jer se ne izračunavaju nikakvi kvadratni korjeni. Bilo bi ih korektnije nazvati “kvadratnim metodama” kod rješavanja (10.5).

Rješavanje LS procjena pomoću QR faktorizacija

Postoje različiti pristupi konstrukciji R , kao Hausholder transformacije, Gram-Schmidt procedure, i Cholesky dekompozicija. Mi ćemo koristiti metod QR faktorizacije od $n \times d$

10. IZRAČUNAVANJE PROCJENA

koja je deifinisana sa:

$$A = QR, \quad QQ^T = I, \quad R \text{ upper triangular} \quad (10.6)$$

Ovdje Q je n x n a R je n x d.

Da bi se primjenilo ovo na LS slučaj procjene parametara, napisaćemo opšti multivarijabilni slučaj (4.59) upomoću matrica, uvodeći:

$$\begin{aligned} \mathbf{Y}^T &= [y^T(1) \ \dots \ y^T(N)], \quad \mathbf{Y} \text{ is } Np \times 1 \\ \Phi^T &= [\varphi(1) \ \dots \ \varphi(N)], \quad \Phi \text{ is } Np \times d \end{aligned} \quad (10.7)$$

Ovdje je p=dim y. Tada LS kriterij se može pisati kao:

$$V_N(\theta, Z^N) = |\mathbf{Y} - \Phi\theta|^2 = \sum_{t=1}^N |y(t) - \varphi^T(t)\theta|^2 \quad (10.8)$$

Na normu očito ne utiče nikava ortonormalna transformacija koja se primjenjuje na vektor $\mathbf{Y} - \Phi\theta$. Zbog toga ako Q (pN x pN) je ortonormalno, tada je $QQ^T = I$.

$$V_N(\theta, Z^N) = |Q(\mathbf{Y} - \Phi\theta)|^2$$

10. IZRAČUNAVANJE PROCJENA

Uvedimo sada QR – faktorizaciju:

$$[\Phi \ Y] = QR, \quad R = \begin{bmatrix} R_0 \\ \dots \\ 0 \end{bmatrix} \quad (10.9)$$

Ovdje je R_0 je gornje trougaona matrica $(d+1) \times (d+1)$ matrica, koju smo dekomponovali kao:

$$R_0 = \begin{bmatrix} R_1 & R_2 \\ 0 & R_3 \end{bmatrix}, \quad R_1 \text{ is } d \times d, \quad R_2 \text{ is } d \times 1, \quad R_3 \text{ is scalar} \quad (10.10)$$

Ovo znači da :

$$V_N(\theta, Z^N) = |Q^T(Y - \Phi\theta)|^2 = \left| \begin{bmatrix} R_2 \\ R_3 \end{bmatrix} - \begin{bmatrix} R_1\theta \\ 0 \end{bmatrix} \right|^2 = |R_2 - R_1\theta|^2 + |R_3|^2$$

koje je očito minimizirano za:

$$R_1\hat{\theta}_N = R_2, \quad \text{giving} \quad V_N(\hat{\theta}_N, Z^N) = |R_3|^2 \quad (10.11)$$

Postoje tri važne prednosti ovog načina rješavanja LS procjena:

10. IZRAČUNAVANJE PROCJENA

1. Sa $R(N)$ kao u (10.3) , $R(N) = \Phi^T \Phi = R_1^T R_1$, tako da kondicioni broj (odnos izmedju najveće i najmanje singularne vrijednosti) od R_1 je kvadratni korjen od $R(N)$. Zbog toga (10.11) je mnogo bolje kondicionirano nego njegov ekvivalent (10.5).

2. R_1 je trougaona matrica tako da je lako riješiti jednačinu.

3. Ako je QR- faktorizacija izvršena za veličinu regresora d^* , tada rješenja i funkcije gubitaka za sve modele sa manje parametara – dobijene sa postavljanjem repnih parametara u q na nulu- se lako dobiju iz R_0 .

Primjetimo da se velika matrica Q nikada ne zahtjeva da se nadje $\hat{\theta}_N$ i funkcija gubitaka.

Sva informacija je sadržana u u “maloj” matrici R_0 . U Matlabu, značajna ušteda u računanju se dobije ako se uzme $R=\text{triu}(\text{qr}(A))$ da se izračuna (10.6) ako Q nam nije interesantno, a A ima mnogo više redova nego kolona.

10. IZRAČUNAVANJE PROCJENA

Početni uslovi : “Uprozoreni” podaci

Tipična struktura regresionog vektora $\mathbf{f}(t)$ je da se ona sastoji od pomjerenih podataka (uz eventualno rearanžiranje):

$$\varphi(t) = \begin{bmatrix} z(t-1) \\ \vdots \\ z(t-n) \end{bmatrix} \quad (10.12)$$

Ovdje je $z(t-1)$ r –dimenzionalni vektor. Na primjer, ARX model (4.11) sa $n_d = n_b = n$ daje (10.12) sa:

$$z(t) = \begin{bmatrix} -y(t) \\ u(t) \end{bmatrix}$$

dok AR- model za p -dimenzionalni proces $\{y(t)\}$ zadovoljava (10.12) sa $z(t)=-y(t)$.

Sa strukturom (10.12), matrica $R(N)$ u (10.3) će biti $n \times n$ blok matrica , čiji i -j blok je $r \times r$ matrica:

$$R_{ij}(N) = \frac{1}{N} \sum_{t=1}^N z(t-i)z^T(t-j) \quad (10.13)$$

10. IZRAČUNAVANJE PROCJENA

Ako imamo samo saznanje o $z(t)$ za $1 \leq t \leq N$, postavlja se pitanje šta da radimo sa nepoznatim početnim uslovima za $t \leq 0$ u (10.13). Mogu se uzeti dva pristupa:

1. Startati sumiranje u (10.13) i (10.14) u $t=n+1$ a ne u $t=1$. Tada sve sume (10.13) će uključiti samo poznate podatke.

2. Zamjeniti nepoznate početne vrijednosti sa nulama ("preduprozorenje" tj. "prewindowing"). Radi simetrije, repne (trailing) vrijednosti $z(t)$, $t= N+1, \dots, N+n$, se takodjer mogu zamjeniti sa nulama ("postwindowing") i sumiranje u (10.3) se proširuje do $N+n$. U ovom slučaju (10.5) su takodjer poznate kao Yule-Walker jednačine.

Često se dodaju i dodatni prozori podataka na obadva kraja zapisa da bi se "omekšali" efekti pripisanih nula.

U procesiranju govora ovakvi pristupi se nazivaju kovarijansnom metodom i autokorelacionim metodom.

Sa logičke tačke gledišta, pristup 1 je prirodniji. Pristup 2., pak daje specijalnu osobinu po kojoj blok (10.13) zavisi samo od razlike izmedju indeksa:

10. IZRAČUNAVANJE PROCJENA

$$R_{ij}(N) = R_{\tau}(N), \quad \tau = i - j$$

$$R_{\tau}(N) = \frac{1}{N} \sum_{t=\tau}^N z(t - \tau)z^T(t) \quad \tau \geq 0, \text{ analogously for } \tau < 0 \quad (10.14)$$

Ovo čini $R(N)$ blok Toeplitz matricom, koja daje lakši način rješavanja (10.5), kako ćemo vidjeti u nastavku.

Jasno, kada je $N \gg n$, razlika između dva pristupa postaje zanemarljiva.

Levinson-ov algoritam

Šift struktura (10.12) daje specifičnu strukturu za matricu $R(N)$. Postoji dosta literature koja koristi ovaj "brzi algoritam", a najpoznatiji je Levinsonov algoritam koji za AR model:

$$\hat{y}^n(t|\theta) = -a_1^n y(t-1) - \dots - a_n^n y(t-n) \quad (10.15)$$

se svodi na slijedeće relacije za izračunavanje procjenjenih koeficijenata modela:

10. IZRAČUNAVANJE PROCJENA

$$\begin{aligned}
 \hat{a}_k^{n+1} &= \hat{a}_k^n + \hat{\rho}_n \hat{a}_{n-k+1}^n, \quad k = 1, \dots, n \\
 \hat{a}_{n+1}^{n+1} &= \hat{\rho}_n \\
 V_{n+1} &= V_n + \hat{\rho}_n \alpha_n \\
 \hat{\rho}_n &= \frac{-\alpha_n}{V_n} \\
 \alpha_n &= R_{n+1} + \sum_{k=1}^n \hat{a}_k^n R_{n+1-k}
 \end{aligned} \tag{10.23}$$

Simbol “ ^ ” (hat) ovdje indicira stvarne procjene, bazirane na N podataka, nasuprot opštim parametrima modela α_k^n . Ovaj izraz nam dozvoljava da lako izračunamo \hat{a}_k^{n+1} iz \hat{a}_k^n . Sa početnim uslovima jednakim:

$$\begin{aligned}
 V_1 &= R_0 - \frac{R_1^2}{R_0} \\
 \hat{a}_1^1 &= \frac{-R_1}{R_0}
 \end{aligned} \tag{10.24}$$

imamo šemu za računanje procjena bilo kojeg reda. Primjetimo da prelazak sa \hat{a}_k^n na \hat{a}_k^{n+1} u (10.23) zahtjeva $4n+2$, sabiranja i množenja i jedno dieljenje. Zbog toga se algoritam naziva

10. IZRAČUNAVANJE PROCJENA

brzim algoritmom.

Numeričko rješenje pomoću iterativnih metoda traženja.

U opštem slučaju funkcija:

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \ell(\varepsilon(t, \theta), \theta) \quad (10.33)$$

ne može biti minimizirana analitičkim metodama. Također ne može ni jednačina:

$$0 = f_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \zeta(t, \theta) \alpha(\varepsilon(t, \theta)) \quad (10.34)$$

biti direktno riješena. U tom slučaju rješenje treba da se nadje iterativnim numeričkim tehnikama.

Numerička minimizacija

Metode numeričke minimizacije funkcije $V(\mathbf{q})$ iterativno ažuriraju procjenu tačke minimizacije. Ovo se obično radi na slijedeći način:

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} + \alpha f^{(i)} \quad (10.35)$$

10. IZRAČUNAVANJE PROCJENA

gdje je $\mathbf{f}^{(i)}$ pravac traženja baziran na informaciji o $V(\mathbf{q})$ prikupljenoj u prethodnim iteracijama, a α je pozitivna konstanta određena tako da se dobije odgovarajuće smanjenje u vrijednosti $V(\mathbf{q})$. Zavisno od informacije koju obezbjeđuje korisnik da se odredi $\mathbf{f}^{(i)}$, metodi numeričke minimizacije se mogu podijeliti u tri grupe:

1. Metode koji koriste samo vrijednost funkcije
2. Metode koje koriste i vrijednost funkcije V kao i njen gradijent.
3. Metode koje koriste vrijednost funkcije, njenog gradijenta i njenog Hessiana (matrica drugog izvoda).

Tipični član treće grupe odgovara Newtonovim algoritmima, gdje je korekcija u (10.35) izabrana u "Newton" ovom pravcu :

$$\mathbf{f}^{(i)} = - \left[\mathbf{V}''(\hat{\theta}^{(i)}) \right]^{-1} \mathbf{V}'(\hat{\theta}^{(i)}) \quad (10.36)$$

Najvažnija podklasa grupe 2 se sastoji od "quasi-Newton" ovih metoda, koje formiraju procjenu Hessiana i onda koriste (10.36).

Algoritmi grupe 1 ili formiraju gradijentne procjene putem diferentnih aproksimacija i nastavljaju kao quasi-Newton metodi

10. IZRAČUNAVANJE PROCJENA

ili imaju specifične paterne traženja.

Identifikacioni korisnik obično samo koristi program koji implementira neki od ovih algoritama. U svakom slučaju biće neophodno sračunati vrijednosti funkcija od (10.33) za svaku zahtjevanu vrijednost \mathbf{q} . Najveći problem leži u izračunavanju sekvence grešaka predikcije $\varepsilon(t, \mathbf{q}) = 1, \dots, N$. Ovo samo po sebi može biti i jednostavan ali i komplikovan zadatak.

Naprimjer, gradijent od (10.33) je :

$$V'_N(\theta, Z^N) = -\frac{1}{N} \sum_{t=1}^N \{ \psi(t, \theta) \ell'_p(\varepsilon(t, \theta), \theta) - \ell'_q(\varepsilon(t, \theta), \theta) \} \quad (10.37)$$

Ovdie, kao i obično, $\mathbf{y}(t, \mathbf{q})$ je $d \times p$ gradijentna matrica od $\hat{y}(t|\theta)$ ($p = \dim y$), po \mathbf{q} . Glavni računarski problem u (10.37) leži u izračunavanju sekvence od $\mathbf{y}(t, \mathbf{q})$, $t = 1, 2, \dots, N$. Za neke modele direktno računanje \mathbf{y} može biti vrlo kompleksno, i onda se treba poslužiti metodama minimizacije iz grupe 1.

10. IZRAČUNAVANJE PROCJENA

Računarski gradijenti

Da bi koristili formule iz prethodne sekcije, mi trebamo izraze za $\mathbf{y}(t, \mathbf{q})$ tj. gradijent predikcije. Obim posla potreban da se izračuna $\mathbf{y}(t, \mathbf{q})$ je vrlo zavistan od strukture modela, i ponekad se mora koristiti i numeričko diferenciranje.

Primjer 10.1 ARMAX struktura modela

Posmatrajmo ARMAX model (4.14). Prediktor je dat sa (4.18):

$$C(q)\hat{y}(t|\theta) = B(q)u(t) + [C(q) - A(q)]y(t) \quad (10.50)$$

Diferencirajući ovaj izraz po a_k daje:

$$C(q)\frac{\partial}{\partial a_k}\hat{y}(t|\theta) = -q^{-k}y(t) \quad (10.51)$$

Slično,

$$C(q)\frac{\partial}{\partial b_k}\hat{y}(t|\theta) = -q^{-k}u(t)$$

i

$$q^{-k}\hat{y}(t|\theta) + C(q)\frac{\partial}{\partial c_k}\hat{y}(t|\theta) = -q^{-k}y(t)$$

10. IZRAČUNAVANJE PROCJENA

Sa vektorom $\mathbf{f}(t, \mathbf{q})$ definisanim sa (4.2), ovi izrazi se mogu pogodnije pisati kao:

$$C(q)\psi(t, \theta) = \varphi(t, \theta) \quad (10.52)$$

Gradijent je tada dobijen filtriranjem "regresionog vektora" $\mathbf{f}(t, \mathbf{q})$ kroz filter $1/C(q)$. Ovaj filter je stabilan za sve \mathbf{q} za koje prediktor (10.50) je stabilan.

SISO model crne kutije

Većina formula za SISO modele crne kutije će biti uključena u tretman opšteg modela (4.33). Prediktor za ovaj model je dat sa (4.35):

$$\hat{y}(t|\theta) = \frac{D(q)B(q)}{C(q)F(q)}u(t) + \left[1 - \frac{D(q)A(q)}{C(q)} \right] y(t) \quad (10.53)$$

iz kojeg možemo naći kao i u prethodnom primjeru, da

$$\frac{\partial}{\partial a_k} \hat{y}(t|\theta) = -\frac{D(q)}{C(q)} y(t-k) \quad (10.54a)$$

$$\frac{\partial}{\partial b_k} \hat{y}(t|\theta) = -\frac{D(q)}{C(q)F(q)} u(t-k) \quad (10.54b)$$

10. IZRAČUNAVANJE PROCJENA

$$\begin{aligned}\frac{\partial}{\partial c_k} \hat{y}(t|\theta) &= -\frac{D(q)B(q)}{C(q)C(q)F(q)}u(t-k) + \frac{D(q)A(q)}{C(q)C(q)}y(t-k) \\ &= \frac{1}{C(q)}\varepsilon(t-k, \theta)\end{aligned}\quad (10.54c)$$

$$\begin{aligned}\frac{\partial}{\partial d_k} \hat{y}(t|\theta) &= \frac{B(q)}{C(q)F(q)}u(t-k) - \frac{A(q)}{C(q)}y(t-k) \\ &= -\frac{1}{C(q)}v(t-k, \theta)\end{aligned}\quad (10.54d)$$

$$\begin{aligned}\frac{\partial}{\partial f_k} \hat{y}(t|\theta) &= -\frac{D(q)B(q)}{C(q)F(q)F(q)}u(t-k) \\ &= \frac{D(q)}{C(q)F(q)}w(t-k, \theta)\end{aligned}\quad (10.54e)$$

gdje smo koristili ε , v i w , kako su definisani sa (4.37) do (4.39). Gradijent $\mathbf{y}(t, \mathbf{q})$ je sada dobijen filtriranjem regresionog vektora $\mathbf{f}(t, \mathbf{q})$ (definisanog u (4.40) kroz linearne filtere, mada različiti dijelovi od \mathbf{f} su udruženi sa različitim filterima u opštem slučaju. Jasno je da filteri uključeni u ovo su svi stabilni za $\theta \in D_M$.

U specijalnom slučaju modela sa izlaznom greškom, $A(q)=C(q)=D(q)=1$, dobićemo iz (10.54 b,e)

$$F(q)\psi(t, \theta) = \varphi(t, \theta) \quad (10.55)$$

10. IZRAČUNAVANJE PROCJENA

Lokalna rješenja i početne vrijednosti

Lokalni minimumi

Opšte numeričke sheme za minimizaciju i rješavanje jednačina, tipično imaju osobinu da, sa pogodno izbranom dužinom koraka integracije m , konvergiraju ka rješenju postavljenog problema. Ovo znači da će (10.48) i (10.49) konvergirati ka tački θ_N^* tako da:

$$f_N(\theta_N^*, Z^N) = 0 \quad (10.77)$$

dok (10.40) sa pozitivno definitnom R , konvergira ka lokalnom minimumu od $V_N(\theta, Z^N)$.

Za problem minimizacije, nas interesuje globalni minimum. Jednačina (10.77) može imati nekoliko rješenja. Jasno je da je inherentno svojstvo rutina traženja da se može garantirati samo konvergencija ka lokalnom rješenju problema. Da bi se našlo globalno rješenje, obično ne postoji drugi način nego startati iterativnu minimizacionu rutinu u različitim dopustivim početnim vrijednostima i porediti rezultate. Važna mogućnost je da se može koristiti neka preleminarna estimaciona procedura da se proizve-

10. IZRAČUNAVANJE PROCJENA

du dobre početne vrijednosti za minimizaciju. Kod validacije modela, model se ocjenjuje na osnovu njegovih performansi. Zbog toga, lokalni minimumi ne moraju uvijek kreirati probleme u praksi. Ako model prodje testove validacije, on će biti prihvatljiv model , čak i ako ne daje globalni minimum funkcije kriterija.

Problem “loših” lokalnih rješenja ima dva aspekta. Koncentrišimo se na problem lokalnih minimuma. Može se desiti da granična vrijednost funkcije kriterija , kako N teži ka beskonačnosti , $\bar{V}(\theta)$ ima takve lokalne minimume. Tada takodjer $V_N(\theta, Z^N)$ će imati takve minimume za veliko N . Egzistencija lokalnih minimuma od $\bar{V}(\theta)$ se može analizirati ali nema nekih značajnih rezultata u ovoj oblasti.

Drgi aspekt je , čak i ako $\bar{V}(\theta)$ ima samo jedan lokalni minimum (koji je jednak globalnom), funkcija $V_N(\theta, Z^N)$ može imati druge lokalne minimume zbog slučajne prirode podataka. Ovo je još teži problem za analitički tretman. Jedan izuzetak je metod linearne regresije najmanjih kvadrata, kod kojeg sa konstrukcijom funkcije kriterija ne možemo imati neglobalnih lokalnih minimuma bez obzira na osobine podataka.

10. IZRAČUNAVANJE PROCJENA

Rezultati za SISO modele crne kutije

Jedini analitički rezultati koji su raspoloživi o lokalnim rješenjima su za modele crne kutije pod pretpostavkom da sistem se može opisati unutar skupa modela : $S \in \mathcal{M}$. Izlistaćemo ove rezultate bez njihovih dokazivanja. Svi se svode na opšti SISO model (10.62) i odnose se na

$$\bar{V}(\theta) = \bar{E} \frac{1}{2} \varepsilon^2(t, \theta)$$

Zbog jednostavnosti, zvaćemo neglobalni , lokalni minimum kao “lažni minimum”:

- Za ARMA modele ($B \equiv 0, D \equiv F \equiv 1$) sve stacionarne tačke od $\bar{V}(\theta)$ su globalni minimumi
- Za ARAX modele ($C \equiv F \equiv 1$) nema lažnih lokalnih minimuma ako je odnos signala prema šumu dovoljno velik. Ako je vrlo mali, lažni lokalni minimumi postoje
- Ako je $A \equiv 1$, nema lažnih lokalnih minimuma ako je $n_f = 1$.
- Ako je $A \equiv C \equiv D \equiv 1$, nema lažnih lokalnih minimuma ako je ulaz bijeli šum. Za druge ulaze, pak, lažni minimumi mogu postojati.⁷⁷

10. IZRAČUNAVANJE PROCJENA

Za ARMAX model ($F \equiv D \equiv 1$) nije poznato da li postoje lažni lokalni minimumi. Praktično iskustvo sa različitim strukturama modela je da globalni minimum se obično nalazi bez mnogo problema za ARMAX modele. Za strukture sa izlaznom greškom, konvergencija ka lažnom minimumu se može često pojaviti.

Inicijalne vrijednosti parametara

Zbog mogućeg pojavljivanja neželjenih lokalnih minimuma u funkciji kriterija, vrijedi potrošiti malo vremena da se proizvedu dobre početne vrijednosti za iterativne procedure traženja. Takodjer, pošto Newton-ski tipovi metoda imaju dobre lokalne brzine konvergencije, ali ne i dovoljno brzine konvergencije kada su udaljenije od minimuma, ovi naponi se obično isplate time što će biti manje iteracija a time i kraće ukupno vrijeme računanja.

Za fizikalno parametrizirane strukture modela, najprirodnije je da se koristi fizikalni uvid u sistem da se obezbjede razumne početne vrijednosti. Takodjer, to nam omogućava da nadziremo i interagiramo sa iterativnim šemama traženja.

Za linearne strukture modela crne kutije, postoji nekoliko mogućnosti. Iz iskustva, slijedeća procedura je dobra početna⁷⁸

10. IZRAČUNAVANJE PROCJENA

procedura za opštu strukturu modela (10.62):

1. Primjeniti IV metod da se procjeni dinamička prenosna funkcija B/AF . Najčešće, jedno od A i F je jedinična matrica. Za sistem koji radi u otvorenoj konturi, prvo se može odrediti LS procjena jednog ARX modela, koja će se koristiti u generisanju instrumenata kao u (7.123). (10.79a)

2. Odrediti procjenu šuma u jednačini kao u (10.74). (10.79b)

3. Odrediti C i/ili D u (10.75) preko (10.76) nakon prvog AR koraka visokog reda, da se nadje \hat{e} , što je nepotrebno ako je $C=1$). Red AR modela se može izabrati kao suma svih redova modela u (10.62) tako da se balansira računarski napor. (10.79c)

U slučaju da $S \in \mathcal{M}$ ovo će dovesti inicijalne procjene parametara proizvoljno blisko pravim vrijednostima kako N raste. Odavde metodi koji su razmatrani će nas efikasno dovesti do globalnog minimuma kriterija. U ovom slučaju ćemo imati proceduru koja je

10. IZRAČUNAVANJE PROCJENA

globalno konvergentna ka globalnom minimumu za dovoljno veliko N .

Za nelinearnu strukturu crne kutije (10.61) postoji nekoliko tehnika. Jednostavna je ona da se posije ("seed") veliki broj fiksnih vrijednosti nelinearnih parametara \mathbf{b}_k i \mathbf{g}_k i da se procjene odgovarajući α pomoću linearnih najmanjih kvadrata. Procjene α_k koje su najznačajnije (relativno u odnosu na njihove procjenjene standardne devijacije) se onda biraju i odgovarajuće vrijednosti \mathbf{b}_k i \mathbf{g}_k se koriste kao početne vrijednosti za obezbjedjenje Gauss-Newton-ovog iterativnog traženja.

Inicijalni uslovi filtera

Filteri (10.53)-(10.54) kao i (10.56) zahtjevaju inicijalne vrijednosti $\mathbf{f}(0, \mathbf{q})$ da se inicijaliziraju. U slučaju da filteri imaju konačni impulsni odziv, što se dešava samo za ARX specijalan slučaj, možemo čekati da inicijaliziramo filtere sve dok nije poznato dovoljno prošlih vrijednosti. Ovo je ono što smo nazvali pristup 1 u izrazu (10.13). U opštem slučaju, mi trebamo strategiju da možemo da radimo sa nepoznatim početnim uslovima.

10. IZRAČUNAVANJE PROCJENA

Ovo nije bilo intenzivnije diskutovano u literaturi , ali se može ukazati na slijedeće pristupe:

1. Uzeti $\mathbf{f}(0, \mathbf{q})=0$
2. Izabrati $\mathbf{f}(0, \mathbf{q})$ tako da prvi $\hat{y}(t|\theta), t = 1, \dots, \dim \varphi$ tačno se podudara sa $y(t)$.
3. Uvesti $\mathbf{f}(0, \mathbf{q})=\eta$ kao parametar i procjeniti ga zajedno sa \mathbf{q} .
4. Procjeniti ili "procjeniti unatrag" (backforecast) $\mathbf{f}(0, \mathbf{q})$ iz podataka na taj način da se odgovarajući filteri izvršavaju inverzno (unatrag) u vremenu.

Za modele gdje je tranzijent prediktorskog filtera prekratak u poredjenju sa dužinom zapisa podataka, nije bitno koji od gornjih pristupa se izabere. Medjutim kod sporo opadajućih tranzijenata, prva dva metoda mogu imati vrlo negativan uticaj na kvalitet modela. Ovo je naročito uočljivo za OE modele, gdje model šuma će izabrati rezidualne iz lošeg tranzijentnog ponašanja.