15. IZBOR IDENTIFIKACIONIH KRITERIJA

Opšti aspekti

Izbor identifikacionog metoda je jedna od važnih odluka u okviru identifikacije sistema.

U dosadašnjim razmatranjima su razvijena i analizirana tri osnovna pristupa identifikaciji, od kojih je svaki bio udružen sa nekim dizajn varijablama:

1. *Pristup sa greškom predikcije (7.12)*

- $l(\cdot)$: norma
- H : set modela šuma, uključujući predfilter L(q)

2. *Korelacioni pristup (7.110)*

- $\alpha(\cdot)$: funkcija oblika (shaping function)
- L(q) : predfilter
- $\zeta(t, \vartheta)$: korelacioni vektor

15. IZBOR IDENTIFIKACIONIH KRITERIJA

3. *Pristup podprostora za procjenu modela u prostoru stanja* (7.66):

- φ_s (t) : korelacioni vektor, koji korespondira sa regresorima za koje su k- koraka unaprijed prediktori odredjeni
- r : horizont maksimalne predikcije
- W_1 i W_2 : težinske matrice u (10.127)
- R : postmultiplikaciona matrica u izrazu (10.128)

Izbor pristupa kao i izbor dizajn varijabli unutar tog pristupa su odredjeni sa nizom razmatranja kao:

Aplikabilnost

Pristup preko greške predikcije ima tu prednost da se može primjeniti na sve strukture modela, linearne i nelinearne, skrojene (tailor made prema fizikalnim razmatranjima) ili black-box parametrizirane. Metoda je jednako vrijedna i za sisteme koji rade u otvorenoj kao i u zatvorenoj konturi. Kôd minimizacije je u suštini isti, samo je izračunavanje prediktora i njegovog gradijenta specifično za svaku strukturu modela.

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Korelacioni pristup je takodjer, u principu , generalno primjenjljiv, ali se najčešće koristi za linearnu black-box familiju modela. Većinom se koristi u okviru IV metode za ARX modele. Primjena u zatvorenoj konturi zahtjeva specijalnu pažnju u izboru instrumenata.

Metod podprostora je specifično dizajniran za black-box linearne sisteme u formi modela u prostoru stanja. Za ovaj metod je takodjer potrebno koristiti specijalna rješenja za rad u zatvorenoj konturi.

Razmatranja bajesa

Da li će primjenjeni metod dati nebajesovanu procjenu u slućaju da $\mathcal{S} \in \mathcal{M}?$. Svi navedeni metodi za sve generičke izbore dizajn varijabli će dati konzistentnost u slućaju rada sistema u otvorenoj konturi. Pristup preko predikcione greške ima dodatnu prednost da garantira konzistentnost i za sisteme koji rade u zatvorenoj konturi, u slućajevima kada struktura modela (uključujući i model šuma) sadrži i sam sistem odnosno njegov istinski model.

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Ako $\mathcal{S} \notin \mathcal{M}$, da li raspodjela bajesa može biti jasno objašnjena i na nju se može uticati?

I ovdje metodi predikcije greške imaju jasnu prednost. Izraz (8.71) i ostali jasno pokazuju u kojem smislu model aproksimira istinski sistem u slućaju linearnog sistema. Za modele sa fiksnim šumom i otvorenom spregom, lako je kontrolisati frekventno naglašavanje putem predfiltriranja. Aproksimacioni aspekti korelacionih metoda se mogu napisati, ali su oni manje transparentni. Tačna priroda aproksimacionih osobina metoda podprostora i kako na njih utiču dizajn varijable još uvijek nije jasno istraženo.

Razmatranja varijanse i robusnosti

Optimizacija (tj. minimizacija) varijanse je jednostavniji problem nego optimizacija bajesa. Razlog leži u tome da mi imamo eksplicitne relacije, kako na varijansu utiču dizajn varijable u slućajevima predikcione greške i korelacionih metoda. Za metode podprostora, nije još uvjek jasno kako

15. IZBOR IDENTIFIKACIONIH KRITERIJA

dizajn varijable utiču na varijansu.

Znamo iz ranijih razmatranja da je teoretska Cramer-Rao donja granica asimptotski dostižljiva pomoću metoda maksimalne vjerovatnosti (maximuim likelihood), tako da znamo unaprijed odgovore na slijedeća pitanja vezana za optimizaciju varijanse:

- najbolji izbor $\ell(x) = -\log f_e(x)\,(f_e(\cdot)$ koji je PDF od istinskih inovacija
- najbolji izbor od modela šuma/predfiltera koji je jednak istinskom opisu šuma (vjerovatno procjenjenog)
- najboljeg izbora IV metode koju treba uzetu jednaku sa metodom greške predikcije koji joj prethodi.

Medjutim, MLE (maximum likelihood) još uvjek ne mora biti najbolji pristup u svim slučajevima. Razlozi mogu biti u efektima bajesa koji su ranije diskutovani ili su procjene osjetljive na prethodno znanje koje može biti neprecizno.

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Lakoća računanja

Metod podprostora ima važnu prednost da njegovi algoritmi ne sadrže iterativno traženje. Oni se mogu implementirati koristeći numerički robustne algoritme. IV metod ima slićnu prednost da on može procjeniti dinamiku linearnog sistema (ali ne i osobine šuma) bez iterativnog traženja. Metodi greške predikcije, izuzev u slućaju linearne regresije, se moraju osloniti na iterativne metode traženja, i mogu biti zarobljeni u pogrešne rezultate koji korespondiraju sa lokalnim minimumima.

Izbor norme: Optimalna norma

Ako se fokusiramo na aspekte varijanse, cilj je da izaberemo l da se $\kappa(\ell, f_e)$ minimizira u normi :

$$\kappa(\ell) = \kappa(\ell, f_e) = \frac{\int (\ell'(x))^2 f_e(x) dx}{(\int \ell''(x) f_e(x) dx)^2} \qquad (15.2)$$

gdje skalar κ zavisi samo od l(x) i od raspodjele istinskih inovacija e_0 (t). Neka je njihova PDF funkcija f_e (x) .

15. IZBOR IDENTIFIKACIONIH KRITERIJA

U izrazu (15.2) prim (‘) i dvostruki prim (“) označavaju diferenciranje po argumentu x.

Primjetimo da je ovaj problem, nezavistan od samog problema identifikacije, specifičnog modela koji se koristi, itd.

Lemma

$$\kappa(\ell, f_e) \geq \kappa(-\log f_e, f_e), \qquad \forall \ell \qquad (15.3)$$

Dokaz lemme

Imamo parcijalnu integraciju:

$$\int \ell''(x) f_e(x) dx = -\int \ell'(x) f_e'(x) dx$$

Cauchy-jeva nejednakost daje:

$$\left[\int \ell''(x) f_e(x) dx\right]^2 = \left[\int \ell'(x) \frac{f_e'(x)}{f_e(x)} \cdot f_e(x) dx\right]^2$$

$$\leq \int [\ell'(x)]^2 f_e(x) dx \cdot \int \left[\frac{f_e'(x)}{f_e(x)}\right]^2 f_e(x) dx$$

15. IZBOR IDENTIFIKACIONIH KRITERIJA

sa jednakošću kada je:

$$\ell'(x) = C_1 \frac{f_e'(x)}{f_e(x)}$$

što dokazuje da:

$$\ell(x) = C_1 \log f_e(x) + C_2$$

daje minimum od (15.3)

Lemma nam kaže da je najbolji izbor:

$$\ell_{\mathrm{opt}}(\varepsilon) = -\log f_e(\varepsilon) \qquad (15.4a)$$

što se može posmatrati kao potvrda činjenice da metod maksimalne vjerovatnosti je asimptotski efikasan. Lemma se odnosi na sluaj sekvence stacionarnih inovacija $\{e_0\ (t)\}$. Ako raspodjela e_0 (t) zavisi od t , f_e (x, t), onda optimalna norma takodjer varira u vremenu:

15. IZBOR IDENTIFIKACIONIH KRITERIJA

$$\ell_{\mathrm{opt}}(\varepsilon, t) = -\log f_e(\varepsilon, t) \qquad (15.4b)$$

Ovo slijedi iz činjenice da (15.45b) daje procjenu maksimalne vjerovatnosti (MLE). Ako su inovacije Gaussovske, sa poznatom varijansama, tada (15.45b) nam kaže da koristimo kvadratičnu normu, skaliranu sa inverznom vrijednošću varijansi inovacija.

Ono što smeta kod ovih rezultata je da PDF f_e može biti nepoznata. Postoje dva načina rješavanja ovog problema: da se simultano procjeni f_e ili da se izabere "ℓ" koje je neosjetljivo na različite vrijednosti f_e koje mogu biti raspoložive.

Adaptacija norme

U prvom slučaju, mi ćemo uključiti dodatne parametre $\alpha : \ell(\varepsilon, \alpha)$ tako da dozvole prilagodjenje norme. Ako imamo dobru procjenu za f_e u (15.4) , tada pristup preko adaptivne norme će biti dobro riješenje.

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Osjetljivost optimalne norme

Optimalno skaliranje varijanse $\kappa(\ell, f)$ može biti prilično osjetljivo u odnosu na PDF f.To znači da skalar $\kappa(-\log f_e, f)$ kao funkcija od f, može imati vrlo oštar minimum kod $f=f_e$. Ovo je prikazano u slijedećem primjeru:

Primjer 15.1 Osjetljivost optimalne norme.

Neka nominalna PDF f_e je normalna sa varijansom 1.

$$f_e(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \varphi(x)$$

Tada $-\log f_e(x) - \frac{1}{2}x^2$ (zanemarujući konstantni član) je:

$$\kappa(-\log f_e, f_e) = \frac{\int x^2 \varphi(x) dx}{[\int \varphi(x) dx]^2} = 1 \qquad (15.5)$$

Predpostavimo sada da greške predikcije sa vrlo malom vjerovatnoćom mogu poprimiti neku veliku vrijednost. Ovo može, napr. odgovarati situaciji sa greškama u opremi za mjerenje

15. IZBOR IDENTIFIKACIONIH KRITERIJA

ili u komunikacionoj opremi. Takve podatke smo nazvali iskačućim (outliers). Mi takodjer predpostavljamo da je ε skoro normalno, ali sa vjerovatnočom $\frac{1}{2}10^{-3}$ može poprimiti vrijednost 100, i sa tom istom vjerovatnočom i vrijednost -100. Stvarno f je tada:

$$f(x) = (1 - 10^{-3})\varphi(x) + 10^{-3}\left[\tfrac{1}{2}\delta(x - 100) + \tfrac{1}{2}\delta(x + 100)\right] \qquad (15.6)$$

Ovo daje:

$$\kappa(-\log f_e, f) = (1 - 10^{-3}) + 10^{4} \cdot 10^{-3} = 10.999 \qquad (15.7)$$

Varijansa time postaje 11 puta veća, mada promjena vjerovatnoće u absolutnom iznosu je vrlo mala.

Robusnost norme

Jasno je da takva osjetljivost na stvarnu vrijednost PDF f_e nije prihvatljiva u praktičnim slučajevima. Adaptacija norme nije riješenje u većini slučajeva, pošto konačan broj podataka ne mora dati dovoljno tačnu procjenu najbolje norme.

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Umjesto toga, mi moramo tražiti norme koje su robusne u odnosu na nepoznate varijacije u PDF. Ovo je dobro razvijena oblast u statistici.

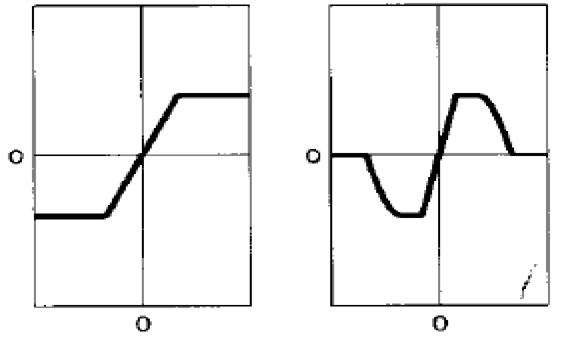
Korisna formalizacija je da se traži norma l koja minimizira najveću varijansu skaliranja koja se može pojaviti kod neke klase PDF-ova:

$$\ell_{\text{opt}} = \arg \min_{\ell} \max_{f \in \mathcal{F}} \kappa(\ell, f) \qquad (15.8)$$

Ova norma daje tkz. minimax M- procjenu. Problem (15.8) sa $\kappa(\ell, f)$ datim sa (15.2) je variacioni problem čije riješenje zavisi samo od familije f funkcija.

Tipične familije ovih funkcija su u okviru normalnih raspodjela. Riješenja ovakvih problema imaju karakterističnu osobinu da $\ell'(x)$ se ponaša kao x za malo x a onda se zasičuje, i može čak dalje da teži ka nuli kako x raste. Neki tipični oblici krivih ovih funkcija su prikazani na slijedećoj slici 15.1.

15. IZBOR IDENTIFIKACIONIH KRITERIJA



Slika br. 15.1

Nastavak primjera 15.1

Neka je $\ell_*(x)$ takva da:

$$\ell_*'(x) = \begin{cases} x, & |x| < 4 \\ 4, & x \geq 4 \\ -4, & x \leq -4 \end{cases}$$

Tada sa f kao u (15.6) :

$$\kappa(\ell_*, f) = \frac{0.999 \int_{|x| \leq 4} x^2 \varphi(x) dx + 0.999 \int_{|x| > 4} 16 \cdot \varphi(x) dx + 0.001 \cdot 16}{\left[\int_{|x| \leq 4} \varphi(x) dx\right]^2}$$

$$\approx 1.015$$

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Smanjenje u varijansi u poredjenju sa (15.7) je drastično. Provjerimo takodjer šta gubimo u optimalnosti u odnosu na to kada bi istinska PDF zaista bila normalna:

$$\kappa(\ell_*, f_e) = \frac{\int_{|x|\leq 4} x^2\varphi(x)dx + \int_{|x|>4} 16\cdot\varphi(x)dx}{\left[\int_{|x|\leq 4}\varphi(x)dx\right]^2} \approx 1.0001$$

Cijena povećane varijanse za normalni slućaj je vrijedna da se plati da se dobije otpornost na male varijacije u vrijednosti PDF.

Preporučena vrijednost robustne norme bi bila:

$$\ell'_R(x) = \begin{cases} x & |x| < \rho\cdot\hat{\sigma} \\ \rho\cdot\hat{\sigma} & x > \rho\cdot\hat{\sigma} \\ -\rho\cdot\hat{\sigma} & x < -\rho\cdot\hat{\sigma} \end{cases} \qquad (15.9)$$

Ovdje je $\hat{\sigma}$ procjenjena standardna deviiacija predikcionih grešaka, dok je ρ skalar u opsegu $1 \leq \rho \leq 1.8$. . Procjena $\hat{\sigma}$. sa svoje strane treba biti robustna tako da nije poremećena sa iskačućim (outliers) vrijednostima. Preporućena procjena je:

15. IZBOR IDENTIFIKACIONIH KRITERIJA

$$\hat{\sigma} = \frac{\text{MAD}}{0.7} \qquad (15.10)$$

Ovdje je MAD = medijan od $\{|\varepsilon(t) - \tilde{\varepsilon}|\}$ sa $\tilde{\varepsilon}$ kao medijan od $\{\varepsilon(t)\}$

Otkrivanje iskačućih vrijednosti (outliers)

Iskakači kao oni u promjeru 15.1 su drastični , ali se mogu često otkriti sa vizuelnim pregledom zapisa podataka. Dobra je praksa, čak kada se koristi i robusna norma, da se iscrtaju podaci prije nego što se koriste za identifikaciju. Iskakači (outliers) se najlakše mogu otkriti u iscrtavanjima (plotovima) reziduala $\varepsilon(t, \hat{\theta}_N)$

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Rezime

Metod greške predikcije (PEM) treba da bude osnovni pristup kod identifikacije sistema.

Ovi metodi imaju tri osnovne prednosti:

1. Primjenjljivost na modele sa opštom strukturom
2. Optimalna asimptotska tačnost kada se istinski sistem može opisati unutar strukture modela.
3. Razumne aproksimacione osobine kada se istinski sistem ne može predstaviti unutar strukture modela.

4. Za datu strukturu modela $\hat{y}(t|\theta)$, PEM se može sumirati kako slijedi:

Izabrati predfilter L(q). Nakon toga formirati kriterij:

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

$$\varepsilon_F(t, \theta) = L(q)[y(t) - \hat{y}(t|\theta)]$$

$\ell(\cdot)$ given by (15.9)

15. IZBOR IDENTIFIKACIONIH KRITERIJA

Za linearni model crne kutije, formirati početnu procjenu $\hat{\theta}_N^{(1)}$ sa procedurom (10.79) koja je ranije opisana. Nakon toga minimizirati V_N iterativno koristeći prigušeni Gauss-Newtonov metod.

Ipak, još uvjek je tačno da drugi metodi mogu biti preferirajuči u izvjesnim slučajevima, naročito za linearne sisteme sa nekoliko izlaza. To zahtjeva strukturu modela sa mnogo parametara, i onda su metodi podprostora dobra alternativa.

Ovi metodi imaju prednost u tome što dozvoljavaju procjenu koristeći numerički robusna izračunavanja bez iterativnog traženja.

Glavna prednost IV metoda je njihova jednostavnost. Često je vrlo poželjno da se koristi IV metoda, kao i metod podprostora , za prvo brzo pricjenjivanje prenosne funkcije sistema. Ona se kasnije može poboljšati sa PEM metodom ako je potrebno.

Teško je reći koji pristup je najbolji. Najbolje je imati sve ove metode u jednom toolboksu kao što je SIT, i onda koristiti različite metode i kroz validaciju ih porediti i vidjeti koji daje najbolje rezultate.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Izbor strukture modela i validacija modela

Izbor odgovarajuće strukture modela $\mathcal{M}$ i je najkritičniji dio za uspješnu identifikaciju. Ovaj izbor mora biti baziran i na razumjevanju identifikacione procedure kao i na uvidu i znanju o sistemu kojeg treba identificirati.

Nakon što je izabrana struktura modela, identifikaciona procedura obezbjedjuje specifičan model u ovoj izabranoj strukturi. Ovaj model može biti najbolji respoloživi, ali je krucijalno pitanje je da li je taj model dovoljno dobar za namjenu za koju ćemo ga koristiti. Testiranje da li je dati model odgovarajući, je poznato kao *validacija modela.*

Opšti aspekti izbora strukture modela

Put ka izboru strukture modela uključuje najmanje tri koraka:

1. Izabrati tip skupa modela.

Ovaj izbor uključuje naprimjer selekciju izmedju nelinearnih i linearnih modela, izmedju ulazno-izlaznih, modela crne kutije i fizikalno parametriziranih modela u prostoru stanja, itd.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

2. Izabrati veličinu skupa modela

Ovo uključuje pitanja kao što je selekcija reda modela u prostoru stanja, stepena polinoma u modelu

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

Ovo takodjer uključuje i problem koje varijable uključiti u opis modela. Mi treba da izaberemo $\mathcal{M}$ i iz datog rastučeg lanca struktura :

$$\mathcal{M}_1 \subset \mathcal{M}_2 \subset \mathcal{M}_3 \ldots \qquad (16.2b)$$

Ovaj problem se naziva problemom selekcije redoslijeda

3. Izabrati način parametrizacije modela .

Kada smo se odlučili za skup modela $\mathcal{M}^*$ (napr. za model u prostoru stanja sa izabranim stepenom), preostaje da ga parametriziramo , tj. da nadjemo pogodnu strukturu modela $\mathcal{M}$ čiji je rang jednak $\mathcal{M}^*$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Kvalitet modela

Kvalitet rezultirajućeg modela može, napr. biti mjeren pomoću kriterija srednjih kvadrata $J(\mathcal{D})$, gdje dizajn varijable D uključuju strukturu modela $\bar{M}$.

Iz ranijih izlaganja smo vidjeli da je pogodno da razdvojimo srednje kvadratnu grešku u doprinos bajesa i varijanse tj.

$$J(\mathcal{D}) = J_B(\mathcal{D}) + J_P(\mathcal{D}) \qquad (16.4)$$

Možemo dakle izabrati M tako da se i bajes i varijansa drže malim. Medjutim ovo su često konfliktni zahtjevi. Da bi se smanjio bajes, najčešće se moraju koristiti veće i fleksibilnije strukture modela, koje zahtjevaju više parametara.

Pošto varijansa tipično raste sa brojem procjenjenih parametara, najbolja struktura modela je kompromis izmedju:

• *Fleksibilnosti* : Korištenje struktura modela koje nude dobre mogućnosti opisivanja mogućih različitih sistema. Fleksibilnost se može dobiti bilo korištenjem mnogo parametara ili postavljajući ih u ‘’strateške pozicije’’ (16.5)

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Štedljivost (Parsimonija): To znači da ne treba da koristimo nepotrebno mnogo parametara , tj. treba da budemo štedljivi (‘’parsimonični’’)sa parametrizacijom modela. (16.6)

Kompromis se može objektivno formalizirati kao minimizacija od (16.4) u odnosu na strukture modela.

Cijena modela

Cijena modela je udružena sa naporom da se on sračuna, tj. da se izvrši minimizacija u :

$$\hat{\theta}_N = \arg\min_{\theta \in D_{\mathcal{M}}} V_N(\theta, Z^N)$$

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \ell\,(\varepsilon(t, \theta), \theta, t) \qquad (7.155)$$

odnosno riješi jednačina:

$$\varepsilon_F(t, \theta) = L(q)\varepsilon(t, \theta)$$

$$\hat{\theta}_N = \operatorname*{sol}_{\theta \in D_{\mathcal{M}}} \left[f_N(\theta, Z^N) = 0 \right]$$

$$f_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \zeta(t, \theta)\alpha(\varepsilon_F(t, \theta)) \qquad (7.156)$$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Ove aktivnosti su umnogome zavisne od strukture modela koja utiče na:

- *Kompleksnost algoritma :* Riješavanje da se dobije $\hat{\theta}_N$ uključuje evaluaciju grešaka predikcije $\varepsilon(t,\theta)$ i njihovih gradijenata $\psi(t,\theta)$ za niz vrijednosti ϑ. Obim posla koji je udružen sa ovim evaluacijama zavisi kritično od izbora ᴍ.
- *Osobine funkcije kriterija:* Obim rada da se riješi za $\hat{\theta}_N$ takodjer zavisi od toga koliko evaluacija funkcije kriterija i njenog gradijenta je potrebno. Ovo je odredjeno ‘’oblikom’’ funkcije kriterija da li (7.155) ili (7.156), nejedinstvenošću minimuma, mogućim nepoželjnim lokalnim rješenjima, itd.

‘’Oblik’’ funkcije kriterija, sa svoje strane je rezultat izbora norme l(.) , i kako $\varepsilon(t,\theta)$ zavisi od ϑ . tj. od strukture modela.

Postoji takodjer i cijena pridružena sa korištenjem modela. Kompleksni model visokog reda je mnogo teži da se koristi u simulacijama i dizajnu sistema upravljanja.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Često ovo povećanje reda modela daje samo marginalno poboljšanje u kvalitetu modela u odnosu na jednostavniji model nižeg reda, i možda ne vrijedi više cijene koju treba platiti za model većeg reda.

Kao posljedica toga:

- Namjeravano korištenje modela koji će se dobiti identifikacijom (16.9) će takodjer uticati na strukturu modela.

Opšta razmatranja

Finalni izbor strukture modela će biti kompromis izmedju navedenih aspekata (16.5) do (16.9). Tehnike i razmatranja koja se koriste kada se evaluiraju ovi aspekti se mogu podjeliti u različite kategorije:

- *A priori razmatranja*: Izvjesni aspekti su nezavisni od skupa podataka Z^N i mogu se a priori evaluirati, prije nego što se podaci izmjere.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

- *Tehnike bazirane na preleminarnoj analizi podataka.*

Sa podacima koji su raspoloživi, neka testiranja i evaluacije Z^N se mogu provesti, koja daju uvid u moguće i pogodne strukture podataka. Ove tehnike ne zahtjevaju nužno računanja kompletnog modela.

- Poredjenje različitih struktura modela : Prije nego što je izabrana konačna struktura modela, preporučuje se da se proba sa različitim strukturama modela i porede kvalitet i cijena koju ti modeli nude. Ovo će zahtjevati izračunavanje i poredjenje nekoliko modela.

- *Validacija datog modela* : Bez obzira na to kako je dati model dobijen, možemo uvjek koristiti Z^N da evaluiramo da li je vjerovatno da će model služiti svojoj namjeni. Ako je neki odredjeni model prihvaćen, mi smo takodjer implicitno i potvrdili i izbor njegove strukture modela kojoj pripada.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

A PRIORI ramatranja

Tip modela

Izbor koji tip modela će se koristiti je prilićno subjektivan i uključuje nekoliko pitanja koja su nezavisna od skupa podataka Z^N . To je najčešće rezulatat kompromisa izmedju aspekata koji su prije navedeni, kombinirani sa nekim iracionalnim faktorima kao raspoloživost kompjuterskih programa i familijarnost sa nekim tipovima modela.

- Kompromis izmedju štedljivosti (parsimonije) i fleksibilnosti je u srcu svakog problema identifikacije. Kako ćemo dobiti dobro fitovanje sa podacima sa što manje parametara? Odgovor obićno će biti da koristimo a priori znanje o sistemu, intuiciju, i ingenioznost. Ove činjenice podcrtavaju to da se identifikacija ne može realizovati kao potpuno automatizovana procedura.

Problem minimizacije izraza (16.4) favorizuje fizički parametrizirane modele. Zavisiće od našeg uvida i razumjevanja procesa da li je razumno graditi dobro osnovane i detaljne fizički parametrzirane strukture modela.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Ovo je dakle problem koji zavisi od aplikacije.

Za fizikalni sistem, a priori informacija se tipično može inkorporirati u model sa kontinualnim vremenom. Ovo znači da izračunavanje $\varepsilon(t,\theta)$ i minimizacija (7.155) postaje mukotrpan posao i sa aspekta programskog napora kao i vremena izračunavanja. Aspekti algoritamske kompleksnosti kao i oblik funkcije kriterija favorizuju stoga model crne kutije (black box). Pod ovim mislimo prije svega na model kao:

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

koji adaptira svoje parametre prema podacima, bez da nameće bilo kakve fizikalne interpretacije vrijednosti ovih parametara.

Opšti savjet je da se pokušaju prvo jednostavne stvari. U sofisticirane strukture modela treba ići samo onda ako jednostavne strukture ne zadovolje testove validnosti. Naročito linearni regresioni modeli kao :

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

$$\hat{y}(t|\theta) = \theta^T \varphi(t) = \varphi^T(t)\theta \qquad (4.12)$$

vode ka jednostavnim i robusnim šemama minimizacije (tj. metod najmanjih kvadrata). Ovi su modeli zbog toga često dobar prvi izbor za problem identifikacije.

Takodjer moramo primjetiti da korištenje fizikalnog modela i apriori znanja ne znači nužno da treba uvjek da konstruišemo neke dobro izgledajuće i dopadljive kontinualne modelne strukture. Analiza prirode relacija izmedju signala mjerenja može nam takodjer dati neke ideje za izbor strukture modela.

Potrebno je provjeriti da li neka nelinearna transformacija podataka (kao u primjeru solarno grijane kuće) ili neka logaritamska transformacija će omogućiti da lakše fitujemo podatke u linearni model. Takodjer i nelinerani efekti aktuatora i senzora mogu biti poznati i mogu se koristiti, kao i informacije o nelinearnosti da se redefiniraju ulazni i izlazni podaci.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Red modela

Rješavanje problema (16.2) obićno zahtjeva pomoć iz podataka. Medjutim , fizikalni uvid kao i namjena identificiranog modela će nam često reći koji je opseg reda modela koji se treba razmatrati. Takodjer, čak kada podaci nisu ni evaluirani, poznavajući N i kvalitet podataka će indicirati koliko parametara je razumno da se procjeni. Sa malo vrijednosti podataka, nije razumno pokušavati odrediti model u kompleksnoj strukturi modela.

U relaciji sa ovim je problem koliko različitih vremenskih skala dozvoliti da jedan te isti model koristi. Iz iskustva je poznato da iz niza razloga, može biti teško jednim modelom opisati više od tri dekade unutar frekventnog područja. Razmatranja o brzini sampliranja, odgovarajučoj pobudi i dužini zapisa podataka, sugerišu da ne treba pokrivati više od tri dekade vremenskih konstanti unutar jednog eksperimenta. Ako je sistem krut, tako da sadrži vrlo udaljene jedne od drugih vremenske konstante od interesa, zaključak je da treba realizovati dva ili više modela, od kojih svaki pokriva odgovarajući opseg frekvencija, i svaki sampliran sa odgovarajućim najpogodnijim intervalom sampliranja. Za visoko frekventni model, nisko frekventna dinamika za sve praktične namjene izgleda kao da su to integratori (čiji broj je jednak broju polova iznad broja nula kod niskih frekvencija).

Sa druge strane, visoko frekventna dinamika izgleda kao statička (trenutačna) relacija sa aspekta nisko-frekventnog modela. U tom slučaju treba uvesti član bez kašnjenja b_0 u(t) u ovaj model.

Parametrizacija modela

Pitanje parametrizacije modela je u suštini numeričko pitanje. Mi tražimo parametrizaciju modela koja je dobro kondicionirana tako da zaokruživanja i ostale numeričke greške u jednom parametru imaju mali uticaj na ulazno-izlazno ponašanje modela. Ovo je problem koji je davno uočen i u oblasti digitalnog filtriranja , ali ne toliko i u literaturi na temu identifikacije. U stvari, standardne ulazno-izlazne strukture modela u obliku prenosnih funkcija , mogu biti vrlo osjetljive na numeričke greške. Izbor parametrizacije linearnog modela se u suštini svodi na izbor nekoliko predstava u prostoru stanja. Model u obiku diferentne jednačine korespondira sa osmotrivošću kanonske forme. Drugi izbori varijabli stanja, kao što su digitalni filteri ili ladder/lattice filteri, daju bolje kondicionirane parametrizacije. Neki autori kao Middleton i Goodwin su se zalagali da se parametrizacija realizuje po

$$\delta = 1 - q^{-1}$$

a ne po q^{-1} da bi se prevazišao ovaj problem.

Selekcija strukture modela na bazi preleminarne analize podataka

Pod preleminarnom analizom podataka, ćemo podrazumjevati izračunavanje koje ne uključuje odredjivanje kompletnog modela sistema**.** Takva se analiza može pokazati korisnom za nalaženje pogodnih struktura modela.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Procjenjivanje tipa modela

Općenito, podacima podržana selekcija strukture modela, se pojavljuje kao nedovoljno istražena oblast. Izuzetak je odredjivanje reda kod linearnih struktura. Razumljivo je da razne neparametarske tehnike mogu biti korisne da se nadju pogodne nelinearne transformacije podataka, kao i da ukažu na tip zavisnosti izmedju mjerenih varijabli koji može postojati i treba ga razmotriti.

Specifičan problem čini izuzetak od ovog pravila, a to je testiranje za efekte nelinearnosti. To je pitanje: da li je vjerovatno da se podaci mogu objasniti sa linearnim relacijama ili se zahtjeva nelinearna struktura modela? Takvi testovi se baziraju na relacijama izmedju viših (više od drugog reda) korelacija i spektara, koji slijede iz linearnih opisa.

Procjenjivanje reda

Red linearnog sistema se može procjeniti na mnoge različite načine. Metodi koji su bazirani na preleminarnoj analizi podataka se svrstavaju u slijedeće kategorije

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

1. Ispitivanje procjene spektralne analize od prenosne funkcije

2. Testiranje ranga u sampliranim kovarijansnim matricama

3. Korelacione varijable

4. Ispitivanje informacione matrice

Pogledajmo ukratko svaki od ovih pristupa

1. **Procjena spektralne analize** :

Neparametarska procjena prenosne funkcije $\hat{G}_N(e^{i\omega})$ će dati vrijednu informaciju o rezonantnim pikovima i savijanju na visokim frekvencijama (high frequency roll-off) kao i o faznim pomacima.

Sve ovo daje sugestiju koji red modela sistema će biti potreban da dâ adekvatan opis interesantnog dijela dinamike sistema čiji se model identificira.

Primjetimo u kontekstu ovoga da Bodeovi plotovi u diskretnom vremenu daju neke artefaktove (greške) kod njihove interpretacije u terminima polova i nula, u poredjenju sa Bode plotovima za kontinualne sisteme. Zato treba biti pažljiv sa ovim posmatranjima .

2. Testiranje rangova u kovarijantnim matricama:

predpostavimo da je istinski sistem opisan sa :

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n)$$
$$= b_1 u(t-1) + \cdots + b_n u(t-n) + v_0(t) \qquad (16.10)$$

za neku sekvencu šuma $\{ v_0 (t)\}$. Predpostavimo takodjer da je n najmanji broj za koji ovo vrijedi (tj. ‘’n je istinski red sistema’’). Neka je :

$$\varphi_s(t) = [-y(t-1) \ldots -y(t-s) \quad u(t-1) \ldots u(t-s)]^T \qquad (16.11)$$

Predpostavimo prvo da $v_0(t) \equiv 0$. Tada (16.10) implicira da matrica:

$$R^s(N) = \frac{1}{N} \sum_{t=1}^{N} \varphi_s(t) \varphi_s^T(t) \qquad (16.12)$$

će biti nesingularna za s ≤ n (pod uslovom da se {u (t)} perzistentno pobudjuje) , i singularna za $s \geq n+1$. $\kappa(s) = \det R^s(N)$

se dakle može koristiti kao test veličina za red modela.

U slučaju da je šum $\{v_0(t)\}$ prisutan u (16.10), (16.12) se može koristiti , sa pogodnim pragom, uz uslov da je odnos signal-šum dovoljno visok. Ako ovo nije slučaj, Woodside sugeriše da se koristi ‘’poboljšana’’ matrica:

$$\hat{R}^s(N) = R^s(N) \quad \hat{\sigma}^2 R_v \tag{16.13}$$

gdje $\hat{\sigma}^2 R_v$ je procjenjeni uticaj $v_0(t)$ na R^s (N).

Bolja alternativa, kada uticaj $v_0(t)$ nije zanemarljiv , je da se koriste drugi korelacioni vektori. Ako su $\{v_0(t)\}$ i $\{u(t)\}$ nekorelirani, mogli bi koristiti

$$\zeta_s(t) = [\,u(t-1) \quad u(t-2) \ldots u(t-2s)\,]^T \tag{16.14}$$

i naći da:

$$\overline{R}^s_\zeta(N) = \overline{E}\varphi_s(t)\zeta_s^T(t) \tag{16.15}$$

je nesingularno za s ≤ n i singularno za $s \geq n+1$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Zamjenjujući $\overline{\overline{E}}$ sa srednjom vrijednošću sampla, daje koristan iznos za testiranje.

Ako je poznato da je $\{v_0(t)\}$ sa pokretnom srednjom vrijednošću reda r, tako da y(t-r-1) i $v_0(t)$ su nekorelirani, možemo koristiti takojder :

$$\zeta_s(t) = \varphi_s(t - r) \tag{16.16}$$

ili bilo koju kombinaciju takvih korelatora.

3. Korelacione varijable

Problem odredjivanja reda (16.2b) je u tome da li da se uključi još jedna varijabla u strukturu modela ili ne. Ova varijabla bi mogla biti y(t-n-1) u (16.10) (problem odredjivanja istinskog reda) ili varijabla mjerene smetnje w(t). U svakom slućaju, pitanje je da li ova nova varijabla može doprinjeti objašnjavanju izlazne varijable y(t). Ovo se mjeri korelacijom izmedju y(t) i w(t). Ipak, da bi odbili mogući odnos izmedju y(t) i w(t) koji je već uzet u obzir sa strukturom modela nižeg reda, korelacija treba biti mjerena izmedju w(t) i onog šta još treba biti objašnjeno. (napr. reziduali $\varepsilon(t,\hat{\theta}_N) = y(t) - \hat{y}(t|\hat{\theta}_N)]$). Ovo je poznato kao kanonska korelacija ili paralelna korelacija u regresionoj analizi.

Možemo takodjer primjetiti da odredjivanje reda modela u prostoru stanja, tj. koliko singularnih vrijednosti su značajne, je takodjer poseban test.

4. Informaciona matrica:

Ukoliko je red modela precjenjen u nekim strukturama modela, globalna i lokalna identifikabilnost će biti izgubljena. Ovo znači da $\psi(t,\theta)$ neće imati puni rang kod $\theta = \theta^*$ (tj. granične vrijednosti). i time informaciona matrica

$$M_N = \frac{1}{\kappa_0} \cdot \sum_{t=1}^{N} E\psi(t,\theta_0)\psi^T(t,\theta_0) \qquad (7.89)$$

će biti singularna. Pošto Gauss-Newtonov algoritam traženja koristi inverznu vrijednost informacione matrice, prirodna testna vrijednost, da se ustanovi da li je red modela suviše visok, će biti kondicioni broj ove matrice.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Analogna situacija se javlja kada se koristi IV metod. Tada matrica

$$R_{\zeta}(N) = \frac{1}{N}\sum_{t=1}^{N}\zeta(t)\varphi^T(t)$$

u

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

će biti singularna kada je red precijenjen. Testiranje kondicioniranja ove matrice je prirodno inkorporirano u IV pristup.

Poredjenje struktura modela

Najprirodniji pristup traženju pogodne strukture modela je jednostavno da se testira više različitih i da se porede rezultirajući modeli. Model koji će se evalurati će biti generički označen sa $m = \mathcal{M}(\hat{\theta}_N)$

On je procjenjen unutar strukture modela M , za koju se predpostavlja da ima $d_{\mathcal{M}} = \dim \theta$ slobodnih parametara. Pod estimacionim podacıma mı podrazumjevamo podatke koji su bili korišteni da se procjeni m, dok pod validacionim podacima mi podrazumjevamo svaki skup podataka koji je raspoloživ a nije bio korišten da se izgradi niti jedan od modela koji se evaluira.

Šta porediti?

Postoji naravno više načina da se evaluira model. Mi ćemo opisati evaluacije i poredjenja koja se baziraju na skupovima podataka sa sistema. Općenito govoreći, ovi testovi trebaju da pokažu relevantne karakeristike za model koji se izvodi, tako da je poželjno da su ovi skupovi podataka prikupljeni pod uslovima koji su bliski onima koji se namjeravaju koristiti za rad sistema koji se identificira. Testovi modela su ustvari testovi o tome kako dobro model reprodukuje ove podatke iz procesa.

Mi ćemo općenito raditi sa modelnim predikcijama sa k-koraka unaprijed $\hat{y}_k(t|m)$ kao osnovom za poredjenja. Pod ovim mi podrazumjevamo da je $\hat{y}_k(t|m)$ izračunato iz prošlih podataka:

$$u(t-1), \ldots, u(1), \; y(t-k), \ldots, y(1) \qquad (16.17)$$

koristeći model m . Slučaj kada je k jednako ∞ odgovara korištenju samo prošlih ulaza ti čistoj simulaciji. Mi koristimo za ovaj slučaj označavanje: $\hat{y}_\infty(t|m) = \hat{y}_s(t|m)$. Slično, uvodimo oznaku $\hat{y}_1(t|m) = \hat{y}_p(t|m)$ za standardni jedan korak unaprijed prediktor. Za linearni model : $y = \hat{G}u + \hat{H}e$ mi ćemo imati:

$$\hat{y}_s(t|m) = \hat{G}(q)u(t) \qquad (16.18a)$$

$$\hat{y}_p(t|m) = \hat{H}^{-1}(q)\hat{G}(q)u(t) + \left(1 - \hat{H}^{-1}(q)\right) y(t) \qquad (16.18b)$$

$$\hat{y}_k(t|m) = \hat{W}_k(q)\hat{G}(q)u(t) + \left(1 - \hat{W}_k(q)\right) y(t) \qquad (16.18c)$$

gdje je $\hat{W}_k$ odredjeno iz :

$$\begin{aligned} W_k(q) &\triangleq 1 - q^{-k}\tilde{H}_k(q)H^{-1}(q) = \left[H(q) - q^{-k}\tilde{H}_k(q)\right]H^{-1}(q) \\ &= \overline{H}_k(q)H^{-1}(q) \end{aligned} \qquad (3.29)$$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Za model izlazne greške H(q)=1 , jasno nema razlike izmedju izraza u (16.18). Inače, primjetimo da postoji značajna konceptualna razlika izmedju y_s i y_p . Ovaj drugi ima y(t-1) a prvi y vrijednosti raspoloživih i može dati uklapanje (fit) koji izgleda dobar, mada njegov model može biti loš.

Primjer 16.1 Trivijalni model

Posmatrajmo model :

$$m: \quad \hat{y}(t|\theta) = y(t-1)$$

On će predvidjeti slijedeći izlaz da je jednak prethodnom. Za zapis podataka koji se brzo samplira $\hat{y}_p(t|m)$ se praktično neće razlikovati od y(t). Sa druge strane $\hat{y}_s(t|m) \equiv 0$, tako da je model neupotrebljiv za simulaciju.

Za opšti model

$$\mathcal{M}: \theta \rightarrow g(t, Z^{t-1}; \theta) \in \mathcal{M}^*; \theta \in D_{\mathcal{M}} \subset \mathbf{R}^d \qquad (5.66)$$

simulirani izlaz je definisan rekurzivno kao:

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

$$\hat{y}_s(t|m) = g(t, Z_s^{t-1}, \hat{\theta}_N) \tag{16.19}$$

$$Z_s^{t-1} = \left\{\hat{y}_s(t-1|m), u(t-1), \hat{y}_s(t-2|m), u(t-2), \ldots, \hat{y}_s(1|m), u(1)\right\}$$

Za aplikacije u sistemima upravljanja, procjenjeni izlaz nad intervalom vremena koji korespondira sa dominatnom vremenskom konstantom, će biti adekvatna varijabla koju treba posmatrati. Simulirani izlaz može biti instruktivan, pošto je to zahtjevniji zadatak da se reprodukuje izlaz samo iz ulaza.

Za nestabilni model, očito je da moramo da budemo oprezni.

Sada se modeli mogu evaluirati ili putem vizuelne inspekcije iscrtavanja (plota) y(t) i $\hat{y}_k(t|m)$ ili pomoću numeričke vrijednosti:

$$J_k(m) = \frac{1}{N}\sum_{t=1}^{N}\left|y(t) - \hat{y}_k(t|m)\right|^2 \tag{16.20}$$

Mi ćemo takodjer koristiti označavanje $J_p = J_1$ i $J_s = J_\infty$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Korisno je dati neku normaliziranu mjeru za ovaj fit. Predpostavimo da y je bio detrendiran na nultu srednju vrijednost i definirajmo:

$$R^2 = 1 - \frac{J_k(m)}{\frac{1}{N}\sum_{t=1}^{N} |y(t)|^2} \tag{16.21}$$

Tada R je onaj dio varijacije izlaza koji je objašnjen modelom, i često je izražen u %.

Mjera kvaliteta J_k (m) će zavisiti od stvarnih podataka iz zapisa za koje se pravi poredjenje. Zbog toga je prirodno da se razmatra oćekivana vrijednost ove mjere, gdje je oćekivanje uzeto u odnosu na podatke, posmatrajući model kao fiksnu , determinističku vrijednost:

$$\overline{J}_k(m) = E J_k(m) \tag{16.22}$$

Ovo daje mjeru kvaliteta za dati model. Sada, $m = \mathcal{M}(\hat{\theta}_N)$ je samo po sebi slučajna varijabla, koja je procjenjena iz podataka sa šumom.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Oćekivanje od modela fitovanja u odnosu na $\hat{\theta}_N$ daje mjeru kvaliteta za strukturu modela M :

$$\overline{J}_k(\mathcal{M}) = E\overline{J}_k(\mathcal{M}(\hat{\theta}_N)) \qquad (16.23)$$

Primjetimo da kod modela sa linearnom regresijom, mjera J_p (m) se može izračunati simultano za mnogo modela. Jedini zahtjev je da su modeli dobijeni brisanjem repnih (trailing) regresora. Ovo slijedi iz izraza :

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

koji pokazuje da je norma k-tog reda od R_2 daje porast $J_p(m_1) - J_p(m_2)$ kada k-ti parametar je otklonjen iz strukture modela.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Poredjenje modela i svježih setova podataka: kros validacija

Nije iznenadjujuće da će model biti u stanju da reprodukuje procjenjene podatke. Pravi test je da li će on biti u stanju da takodjer opiše svjež set podataka sa procesa. Sugestivan i atraktivan način poredjenja dva različita modela m_1 i m_2 je da se evaluira njihova performansa na validacionim podacima, tj. izračunavanjem $J_k (m_i)$ u (16.20). Mi ćemo nakon toga dati prednost onom modelu koji pokaže bolju performansu. Takve procedure su poznate kao kros-validacije i razvijeno je nekoliko varijanata.

Jedna atraktivna osobina kros validacionih procedura je njihov pragmatičan karakter: poredjenje ima smisla i bez bilo kakvih probabilističkih argumenata i bez bilo kakvih predpostavki o istinskom sistemu. Njihov jedini nedostatak je da moramo da pohranimo svjež set podataka za validaciju, i zbog toga ne možemo koristiti sve informacije koje imamo u zapisima da izgradimo modele tj. za estimaciju.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Za linearne regresije možemo koristiti J_p (m) na slijedeći način: Neka je $\hat{y}_p(t|m_t)$ izračunato za model m_t koji je procjenjen iz svih podataka izuzev za observacije (y(t),φ(t)) unutar strukture modela M. Formirajmo J sumiranjem nad svim odgovarajućim kvadratima grešaka. Tada J je mjera snage predikcije ove strukture modela u kros-validacionom smislu, a da nisu nikakvi podaci izgubljeni u fazi estimacije. Procedura se naziva PRESS (prediction sum of squares).

Poredjenje modela na setovima podataka iz druge ruke (second-hand) : evaluacija oćekivanog uklapanja (fita)

Prava mjera kvaliteta za model m je oćekivani kriterij $\bar{J}_k$ u (16.22). Ako model je evaluiran nad validacionim podacima, observacija J_k je rezonska i nebajesovana procjena od $\bar{J}_k$.

Ovo je razlog zašto se preferira evaluacija modela na validacionim podacima**.**

Ako koristimo estimacione podatke za poredjenje, tada J_k nije više nebajesovana procjena $\bar{J}_k$. Ovo znači da vrijednost J_p je jednaka vrijednosti identifikacionog kriterija:

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

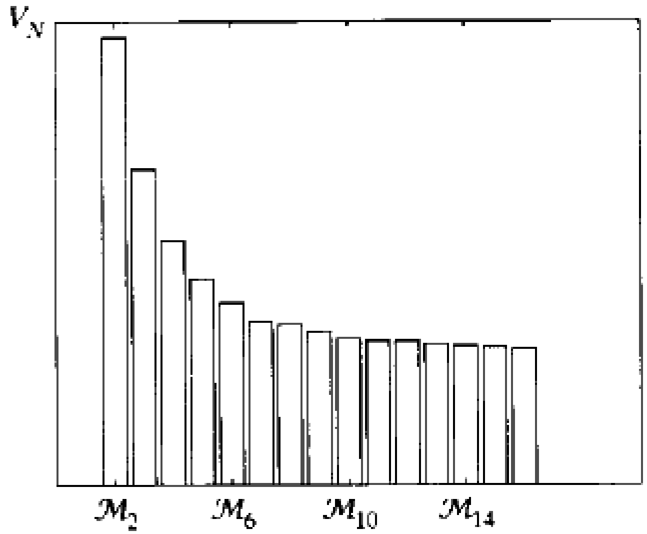
$$J_p(m) = \frac{1}{N}\sum_{t=1}^{N}\left|y(t) - \hat{y}(t|\hat{\theta}_N)\right|^2 = \min_{\theta}\frac{1}{N}\sum_{t=1}^{N}\left|y(t) - \hat{y}(t|\theta)\right|^2 \quad (16.24)$$

Pragmatičan pregled

Model dobijen u većoj strukturi modela će automatski dati manju vrijednost kriterija fitovanja, pošto je minimizaciona vrijednost dobijena minimizacijom nad većim setom podataka. Kako se struktura modela povećava, kao i u (16.2b), minimalna vrijednost kriterija će se ponašati kako je prikazano na slici 16.1: ona je monotono opadajuća funkcija fleksibilnosti strukture modela. U početku, vrijednost V_N opada pošto model preuzima sve više relevantnih osobina od podataka. Ali čak i nakon što je struktura modela dostignuta koja dozvoljava korektan opis sistema, vrijednost V nastavlja da opada, sada zato što dodatni (nepotrebni) parametri podešavaju sebe prema karakteristikama specifične realizacije šuma. Ovo je poznato kao prefitovanje (overfit) i ovaj dodatno poboljšani fit nema neke vrijednosti za nas, pošto ćemo mi primjeniti model na podatke sa različitim realizacijama šuma.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Jasno je da je smanjenje od ‘’overfita’’ manje značajno od smanjenja koje je izazvano kada su u model uključene neke značajnije karakteristike. Mi ćemo stoga gledati na nadjemo ‘’koljeno’’ na krivoj prikazanoj na slici br. 16.1



Slika br. 16.1 Minimalna vrijednost funkcije gubitka kao funkcije od veličine strukture modela (16.2b) V_N =min V_N (ϑ)

Dobra je praksa da se nacrta ova kriva da se dobije subjektivi osjećaj da li je poboljšani fit značajan i vrijedan truda.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Validacija modela

Procedura procjene parametara izabire ‘’najbolji’’ model unutar izabrane strukture modela. Suštinsko pitanje je sada da li ovaj ‘’najbolji model’’ je ‘’dovoljno dobar’’. Ovo je problem validacije modela. Ovo pitanje ima nekoliko aspekata:

1. Da li se model slaže dovoljno dobro sa observiranim podacima sa procesa?
2. Da li je model dovoljno dobar za našu namjenu?
3. Da li model opisuje ‘’istinski sistem’’

Općenito metod da se odgovori na ova pitanja je da se suprostavi model $\mathcal{M}(\hat{\theta}_N)$ sa što je moguće više informacija sa pravog sistema, koliko je to praktično. Ovo uključuje a priori znanje, eksperimentalne podatke i iskustvo u korištenju modela. Tehnike validacije modela se uglavnom fokusiraju na prvo pitanje.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Validacija sa aspekta namjene modela

Ono što je bitno u inženjerskoj praksi je odgovor na pitanje 2. Postoji uvjek neka namjena sa kojom se provodi modeliranje. Razlog može biti da se zahtjeva model radi dizajna i sinteze regulatora, predikcije, ili simulacije. Krajnja validacija je tada da se testira da li problem koji je motivirao razvoj i dobijanje modela je zadovoljavajuće riješen sa dobijenim modelom. Ako regulator baziran na modelu daje zadovoljavajuće upravljanje, tada je model ‘’validan’’ , bez obzira na formalne aspekte koji se mogu postaviti. Često nije moguće, ili je isuviše skupo i opasno testirati sve moguće modele u odnosu na namjeravanu namjenu modela. Umjesto toga treba razviti načine izgradnje povjerenja u razvijeni model.

Raspoloživost fizikalnih parametara

Za strukturu modela koji je parametriziran po fizikalnim parametrima**,** prirodna i važna validacija je da se suprostave procjenjene vrijednosti iz modela i njihove varijanse sa onim što je razumno na osnovu apriori znanja. Takodjer je dobra praksa da se evaluira osjetljivost ulazno-izlaznog ponašanja u odnosu na ove parametre da se provjeri njihova praktična identifikabilnost.

Konzistentnost ulazno-izlaznog ponašanja modela

Za modele crne kutije, mi fokusiramo naš interes na njihove ulazno-izlazne osobine. Za linearne sisteme mi ih obićno prikazujemo kroz Bode-ove dijagrame. Za nelinearne modele, oni bi se obićno anlizirali putem simulacije. Uvjek je dobra praksa da se evaluiraju i porede različiti linearni modeli u Bode-ovim plotovima, po mogućnosti sa procjenjnim varijansama prevedenim u intervale povjerenja od $\hat{G}$ i $\hat{H}$.

Poredjenja izmedju procjena spektralne analize ı Bode-ovih plotova izvedenih iz parametarskih modela su vrlo korisna , jer su ovi formirani i dobijeni iz različitih polaznih predpostavki.

Općenito ako stvarni sistem ne pripada skupu modela, mi dobijamo aproksimaciju čiji će karakter zavisiti od eksperimentalnih uslova, korištenih predfiltera i strukture modela. Time, poredjenjem Bode-vih plotova dobijenih metodama greške predikcije u različitim strukturama, kao i sa različitom predfilterima, sa metodama podprostora i spektralne analize, će dati dobar osjećaj da li su u njima ‘’uhvaćene’’ bitne osobine dinamike sistema koji se identificira.

Redukcija modela

Jedna procedura koja testira da li je model jednostavan i odgovarajući opis sistema je da se primjeni neka od tehnika redukcije modela. Ako se red modela može reducirati bez da značajnije utiče na ulazno-izlazne karakteristike, tada polazni model je nepotrebno kompleksan.

Intervali povjerenja parametara

Drugа procedura za provjeru da li tekući model sadrži isuviše mnogo parametara je da se poredi procjena sa odgovarajućom procjenjenom standardnom devijacijom. Ako interval povjerenja sadrži nulu, mi bi mogli razmatrati da li bi ovaj parametar trebao biti isključen. Ovo je obićno relevantno ako odgovarajući parametar odražava fizikalnu strukturu, kao napr. red modela ili vremensko kašnjenje. Ako su sve procjenjene standardne devijacije veliki iznosi, informaciona matrica je bliska singularnoj. Ovo je takodjer indikacija da je red modela suviše velik.

Simulacija i predikcija

Mi smo u ranijim izvodjenjima koristili sposobnosti modela da reprodukuje ulazno-izlazne podatke u smislu simulacija ali i predikcija kao glavne alate za poredjenja. Takvi plotovi, kao i numerički iznosi uklapanja (fitovanja) koji su im pridruženi, su takodjer vrlo privlaćni za evaluaciju modela. Iz ovoga mi tačno vidimo koje je karakteristike u stanju da model reprodukuje, a koje nije uspjeo da ‘’ulovi’’. Razlike mogu biti uzrokovane šumom ili greškama modela, i mi ćemo samo vidjeti njihov kombinovani uticaj na kvalitet reprodukcije ponašanja modela prema realnim procesima.

Ako bi imali nezavisnu procjenu nivoa šuma, onda bi bili u stanju reći iz $J_k(m)$ iz izraza (16.20) koji je iznos greške modela.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Analiza rezuduala

Izostavljeni dijelovi podataka iz procesa modeliranja (leftovers), tj. onaj dio podataka koje model nije mogao da reprodukuje - nazivaju se rezidualima:

$$\varepsilon(t) = \varepsilon(t, \hat{\theta}_N) = y(t) - \hat{y}(t|\hat{\theta}_N) \tag{16.52}$$

Jasno je da ovi reziduali nose informaciju o kvalitetu modela.

U nastavku ćemo diskutovati formalne metode da se izvedu zaključci o validnosti modela iz analize reziduala.

Pragmatične tačke posmatranja

Mi u suštini imamo skup podataka Z^N bilo da su to estimacioni ili validacioni podaci, i nominalni model m. Mi želimo da znamo kvalitet modela, koji u stvari znači kako će on biti u stanju da repredukuje novi set podataka. Jednostavna i pragmatična polazna tačka je da izračunamo osnovnu statistiku za reziduale iz modela:

$$S_1 = \max_t |\varepsilon(t)|, \quad S_2^2 = \frac{1}{N}\sum_{t=1}^{N} \varepsilon^2(t) \tag{16.53}$$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Intuitivno korištenje ove statistike bi onda bilo slijedeće: ‘’Ovaj model nije nikada proizveo veći rezidual od S_1 (ili jednu prosječnu grešku od S_2) za sve podatke koje smo vidjeli. Vjerovatno je da će te granice zadržati i za sve buduče podatke’’.

Sada, korištenje statistike kao što je (16.53) ima jednu implicitnu invarijantnu predpostavku: Reziduali ne zavise od onoga što se može promjeniti. Od specijalnog interesa je , naravno, da oni ne zavise od specifičnog ulaza koji je korišten u Z^N. Ako bi zavisili, vrijednost (16.53) bi bila ograničena , pošto model treba raditi za čitav opseg mogućih ulaza. Da bi ovo provjerili, rezonski je da analiziramo kovarijansu izmedju reziduala i prošlih ulaza:

$$\hat{R}_{\varepsilon u}^{N}(\tau) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t) u(t - \tau) \qquad (16.54)$$

Ako su ovi iznosi mali, imamo razloga da vjerujemo da mjere (16.53) mogu biti relevantne i kada se model primjeni i na druge ulaze.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Drugi način izražavanja važnosti da je $\hat{R}_{\varepsilon u}^{N}$ malo je slijedeći: Ako postoje tragovi prošlih ulaza u rezidualima, onda postoji dio u y(t) koji potiče od prošlih ulaza i koji nije bio korektno prepoznat i uzet u obzir u modelu m. Prema tome model je mogao biti poboljšan.

Slično, ako mi nadjemo korelaciju izmedju samih reziduala, tj. ako iznosi :

$$\hat{R}_{\varepsilon}^{N}(\tau) = \frac{1}{N}\sum_{t=1}^{N}\varepsilon(t)\varepsilon(t-\tau) \qquad (16.55)$$

nisu mali za $\tau \neq 0,$, tada dio od ε(t) je mogao biti procjenjen iz prošlih podataka. Ovo znači da je y(t) mogao biti bolje procjenjen, što je ponovno znak nedostatka modela.

Za formalniji pristup, mi ćemo motivirati kriterij estimacije kao metod maksimalne vjerovatnosti (ML), predpostavljajući da su izlazni podaci generisani prema izrazu :

$$y(t) = g(t, Z^{N}; \hat{\theta}_{N}) + \varepsilon(t) \qquad (16.56)$$

gdje $\varepsilon(t)$ ima osobine koje su nezavisne jedne od drugih i od prošlih podataka. Pitanje validacije modela u odnosu na podatke je sada: ‘’Da li je vjerovatno da je zapis podataka Z^N bio generisan od strane modela datog sa (16.56)?’’. Ovo pitanje je ekvivalentno sa pitanjem:

‘’Da li je vjerovatno da

$$\varepsilon(t) = y(t) - g(t, Z^N; \hat{\theta}_N) \qquad (16.57)$$

je sekvenca nezavisnih slučajnih varijabli sa PDF $f_e(x, t; \hat{\theta}_N)$?’’.

Jasno je da (16.55) i (16.54) su osnova za dio odgovora.

Test bjeline (whiteness test)

Brojevi $\hat{R}_\varepsilon^N(\tau)$ nose informaciju o tome da li reziduali se mogu posmatrati kao bijeli. Da se dobije predstava o tome kako veliki mogu biti ovi brojevi kada bi $\varepsilon(t)$ i istinski bio bijeli šum, mi ćemo rezonovati na slijedeći način:

Predpostavimo da je $\{\varepsilon(t)\}$ sekvenca bijelog šuma, sa nultom srednjom vrijedošću i varijansom λ . Tada slijedi da je:

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\begin{bmatrix}\varepsilon(t-1)\\ \vdots \\ \varepsilon(t-M)\end{bmatrix}\varepsilon(t) \in AsN(0, \lambda^2 \cdot I)$$

k-ti red ovog vektora je $\sqrt{N}\hat{R}_{\varepsilon}^{N}(k)$. Pod predpostavkom da je ε bijeli šum, ovo znači da:

$$\frac{N}{\lambda^2}\sum_{\tau=1}^{M}\left(\hat{R}_{\varepsilon}^{N}(\tau)\right)^2$$

treba biti asimptotski $\chi^2(M)$ distriburano. Zamjenjujući nepoznato λ sa jednom očiglednom procjenom te vrijednosti neće ovo promjeniti asimptotski (ustvari respodjela postaje F-raspodjela). Test na bjelinu će sada biti, da li :

$$\zeta_{N,M} = \frac{N}{\left(\hat{R}_{\varepsilon}^{N}(0)\right)^2}\sum_{\tau=1}^{M}\left(\hat{R}_{\varepsilon}^{N}(\tau)\right)^2 \qquad (16.58)$$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

će zadovoljiti test da je $\chi^2(M)$ distribuirano, tj. provjeravajući da li $\zeta_{N,M} < \chi^2_\alpha(M)$, je α nivo od $\chi^2(M)$ raspodjele.

Pored ovoga testa bjeline, mogu se provesti i dodatni testovi, kao naprimjer broj promjena predznaka, od ε(t) kao i histogram test za raspodjelu od ε.

Nezavisnost izmedju reziduala i prošlih ulaza

Da bi istražili koji zahtjevi trebaju biti udruženi sa (16.54), definisaćemo:

$$r_M^N = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \varepsilon(t)\varphi(t), \quad \varphi(t) = \begin{bmatrix} u(t - M_1) \\ \vdots \\ u(t - M_2) \end{bmatrix}. \qquad (16.59)$$

$$M = M_2 - M_1 + 1 = \dim\varphi$$

Primjetimo da k-ta komponenta od r je jednaka $\sqrt{N}\,\hat{R}_{\varepsilon u}^N(k + M_1 - 1)$.

data sa (16.54). Ako su ε nezavisni od φ i mogu biti napisani kao:

$$\varepsilon(t) = \sum_{k=0}^{\infty} f_k e(t-k), \quad f_0 = 1, \quad e(t) \text{ white noise with } Ee^2(t) = \lambda \quad (16.60)$$

tada slijedi da:

$$r_M^N \in AsN(0, \lambda P), \quad P = \overline{E}\tilde{\varphi}(t)\tilde{\varphi}^T(t), \quad \tilde{\varphi}(t) = \sum_{k=1}^{\infty} f_k \varphi(t+k) \qquad (16.61)$$

Može se pokazati da (k, ℓ) element od P se mogu takodjer izraziti kao:

$$\sum_{\tau=-\infty}^{\infty} R_\varepsilon(\tau) R_u(\tau - (k-\ell)), \; R_\varepsilon(\tau) = E\varepsilon(t)\varepsilon(t-\tau), \; R_u(\tau) = \overline{E}u(t)u(t-\tau) \qquad (16.62)$$

Sada (16.61) implicira da:

$$\zeta_{N,M}^u = \frac{1}{\lambda} r_M^N P^{-1} [r_M^N]^T \in As\chi^2(M) \qquad (16.63)$$

ako je ε nezavisno od ulaza. Dakle $\zeta_{N,M}^u$ je korektan iznos da se podvrgne $\chi^2(M)$ testu. Primjetimo da mi treba da procjenimo model (16.60) da bi mogli da formiramo ovaj iznos. Ako se predpostavi za ε da je bijeli šum, ili je zadovoljio test (16.58), izračunavanje $\zeta_{N,M}^u$ je pojednostavljeno.

Jednostavno korištenje (16.54) je da se smatra samo jedno dato τ , kako slijedi:

Iz gornjih relaciia sliiedi da:

$$\sqrt{N}\hat{R}^{N}_{\varepsilon u}(\tau) \in AsN(0, P_1). \quad P_1 = \sum_{k=-\infty}^{\infty} R_{\varepsilon}(k)R_{u}(k) \qquad (16.64)$$

Ako N_{α} označava α nivo od N(0,1) raspodjele, mogli bi provjeriti da li je zadovoljeno:

$$\left|\hat{R}^{N}_{\varepsilon u}(\tau)\right| \leq \sqrt{\frac{P_1}{N}}N_{\alpha} \qquad (16.65)$$

Ako ne, hipoteza da ε(t) i u(t-τ) su nezavisni treba biti odbačena.

Jedan privlačan način da se provede test je da se nacrta $\hat{R}^{N}_{\varepsilon u}(\tau)$ kao funkcija od τ. Pošto P_1 u (16.55) ne zavisi od τ, granice povjerenja će biti horizontalne linije. Takav plot otkriva vrijedan uvid u korektnost strukture modela. Ako napr. vremensko kašnjenje u iznosu dva sampla je predpostavljeno u modelu, a stvarno kašnjenje je jedan sampl, tada jasna

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

korelacija izmedju u(t-1) i ε(t) će se pokazati. Kada ispitujemo plot od $\hat{R}^N_{\varepsilon u}(\tau)$ treba uočiti slijedeće činjenice:

1. Korelacija izmedju u(t-τ) i ε(t) za negativne τ je indikacija izlaznog feedbacka u ulazu , a ne da struktura modela nije dobra.

2. Metod najmanjih kvadrata konstruiše $\hat{\theta}_N$ takvo da je $\varepsilon(t, \hat{\theta}_N)$ nekorelirano sa regresorima. Mi sada imamo $\hat{R}^N_{\varepsilon u}(\tau) = 0$ for $\tau = 1, \ldots, n_b$ automatski, za strukturu modela, kada se analiza provodi na estimacionim podacima.

Ovo znači da moramo biti pažljivi kada se biraju brojevi M_1 i M_2 u φ(t). Ako se koriste estimacioni podaci, zajedno sa ARX modelom sa redom n_a i n_b , mi treba da imamo da je $M_1 > n_b$

Takodjer, prirodno je da se uzme da je $M_1 > 0$, ako se samo slućajna zavisnost od prošlih ulaza testira.

Nezavisnost izmedju u i ε se može mjeriti i u drugim terminima. Nemodelirani nelinearni efekti mogu, naprimjer, biti uoćeni u razbacanim plotovima parova $(\varepsilon(t), u(t-\tau))$, ili kroz korelaciju izmedju nelinearnih transformacija u i ε.

Test za dinamičke sisteme

Testiranje korelacije izmedju prošlih ulaza i reziduala je prirodno provesti da bi se evaluiralo da li model ‘’pokupio’’ bitan dio (linearne) dinamike iz u u y. Za dinamički model, rezultati testova se mogu efikasnije i efektivnije vizuelizirati ako ih gledamo kao procjene rezidualne dinamike ili kao model modela greške (*model error model*):

$$\varepsilon(t) = G_\varepsilon(q)u(t) \tag{16.66}$$

Ustvari, ako je ulaz bijel, tada $\hat{R}^N_{\varepsilon u}(\tau)$ su aproksimativno komponente procjene $\hat{\theta}_N$ dobijene iz FIR modela:

$$\varepsilon(t) = \theta^T \varphi(t)$$

sa φ dato sa (16.59), tj. impulsnim odzivom od (16.66). Za slučaj ne-bijelog ulaza, bit će lakše evaluirati plot procjene impulsnog odziva od $\hat{G}_\varepsilon$ nego od procjene korelacije $\hat{R}^N_{\varepsilon u}$

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

pošto je ova druga procjena koralacije afektirana sa internom korelacijom od u.

Čak i efikasnije, za potrebe sistema upravljanja, bi bilo prikazivanje frekventne funkcije procjene $\hat{G}_{\varepsilon}(e^{i\omega})$, zajedno sa procjenjenim regionima povjerenja. Ovo daje sliku koje frekventne opsege model nije ‘’ulovio’’ , u ulazno-izlaznom ponašanju. Zavisno od namjene modela, model bi se mogao prihvatiti kao validan, čak ako (16.65) nije ispunjeno, ukoliko se greške pojavljuju u frekventnim opsezima koja su od manjeg ili nikakvog interesa.

Primjer 16.3 Analiza reziduala za dinamičke sisteme

Sistem:

$$\begin{aligned} y(t) - 1.2y(t-1) - 0.15y(t-1) + 0.35y(t-3) \\ = u(t-1) + 0.5u(t-2) + e(t) - e(t-1) + 0.4e(t-2) \end{aligned}$$

je simuliran sa 500 samplova sa ulazom koji se sastoji od sinusoida izmedju 0.3 i 0.6 rad/sec i sa Gaussovskim šumom sa varijansom 1.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

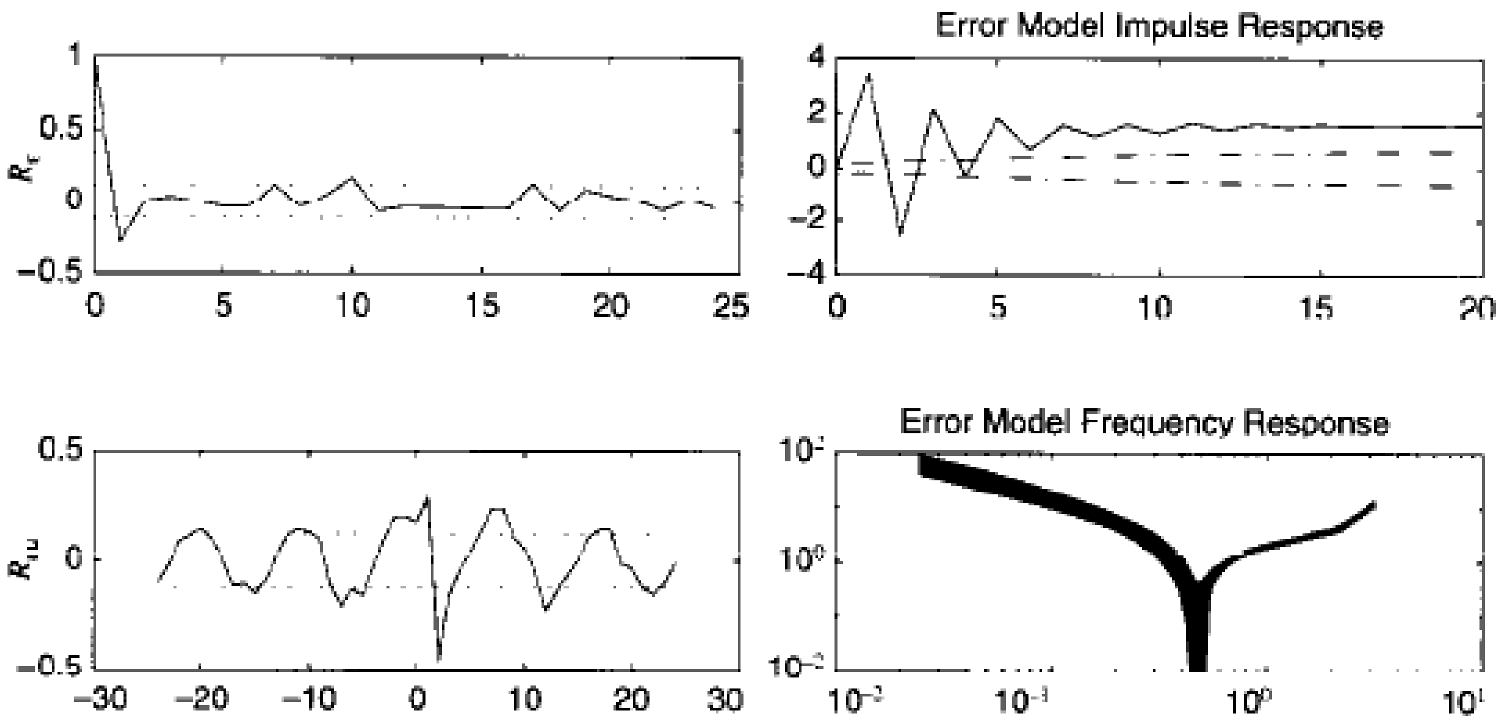
Iz ovih podataka je identificiran ARX model drugog reda m. Validacioni set podataka je generisan koristeći slučajni binarni ulaz sa rezonantnim vrhom od oko 0.3 rad/sec.

Naredna slika br. 16.2 pokazuje rezultat konvencionalne analize reziduala, kada je m bio podvrgnut ovim podacima.

Slika 16.2b pokazuje impulsne i frekventne odzive modela modela greške (16.66) , procjenjenog kao 10-ti red ARX modela. Jasno je da frekventni plot modela greške daje mnogo precizniju informaciju o kvalitetu modela sa aspekta korištenja modela za potrebe upravljanja.

Na slijedećoj slici 16.3 amplitude Bode plotova modela i istinskog sistema su poredjene. Iz ovih poredjenja vidimo da je informacija iz modela greške iz validacije vrlo pouzdana.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

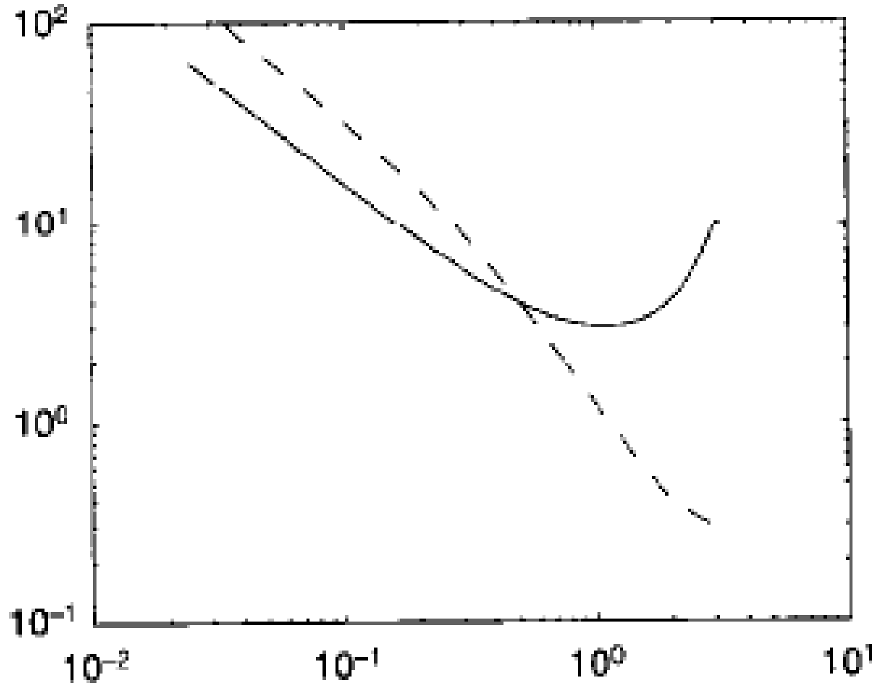


(a) konvencionalna rezidualna analiza korelacionih funkcija

(b) Model modela greške procjenjen iz validacionih podataka

Slika br. 16.2 Validacija ARX modela drugog reda koristeći validacione podatke. Crta-tačka linije označavaju intervale povjerenja. Za plot u frekventnom domenu interval povjerenja je označen sjenčenjem regiona

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA



Slika br. 16.3 Amplitudni Bode-ov plot modela m (nacrtan punom linijom) i plot istinskog (tačnog) sistema (ctkana linija)

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Kritična evalacija podataka

Reziduali $\varepsilon(t, \hat{\theta}_N)$ će nam takodjer reči , kada se ubace u funkciju uticaja :

$$S(t) = \overline{R}_t^{-1}(N)\psi(t, \hat{\theta}_N)\ell'(\varepsilon(t, \hat{\theta}_N))$$

$$\overline{R}_t(N) = \sum_{\substack{k=1 \\ k \neq t}}^{N} \psi(k, \hat{\theta}_N)\ell''(\varepsilon(k, \hat{\theta}_N))\psi^T(k, \hat{\theta}_N) \qquad (15.12)$$

koje tačke podataka su imale veliki uticaj na procjene. Pouzdanost ovih tačaka treba biti kritički evaluirana kao dio procedure validacije modela. Uvjek je dobra praksa da se iscrta $\varepsilon(t, \hat{\theta}_N)$ i da se ispitaju podaci za iskačuće vrijednosti (outliers) kao i na ‘’loše podatke’’.

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

Rezime

Model ‘’Istinskog’’ (tačnog) sistema je jedan apstraktni entitet koji ne može biti dostignut u praktičnim modeliranjima.

Mi se moramo zadovoljiti sa djelomićnim opisima koji su zadovoljavajući sa aspekta namjene modela. Ponekad to znači da treba da radimo sa nekoliko modela istog sistema koji trebaju biti korišteni za različite radne tačke sistema, ili za različite vremenske skale u radu modela, itd.

U ovom poglavlju su opisani razni metodi pomoću kojih se mogu naći pogodne strukture modela i pomoću kojih mi možemo odbaciti ili razviti povjerenje u specifični model.

Medju a priori razmatranjima, navedimo princip : ‘’pokušati najprije sa jednostavnim stvarima’’. Ovo obično znači da treba početi sa testiranjem jednostavnih linearnih regresija, kao što su ARX modeli u linearnim strukturama, kao i varijante sa nelinearnim transformacijama podataka na bazi fizikalnog uvida, gdje je god to moguće i adekvatno.

Za validaciju modela, mi smo opisali niz metoda različite prirode. To su:

16. IZBOR STRUKTURE MODELA I VALIDACIJA MODELA

- Poredjenje linearnih modela dobijenih pod različitim uslovima u različitim strukturama modela (uključujući procjene spektralne analize) u Bode plotovima.
- Poredjenje mjerenih i simuliranih izlaza iz modela za modele dobijene iz različitih struktura.
- Testiranje reziduala na nezavisnost od prošlih ulaza i moguće na bijelinu šuma
- Nadziranje intervala povjerenja za procjene parametara za repne (trailing) i vodeće (leading) nulte vrijednosti u polinomima prenosnih funkcija, kao i za mogući gubitak lokalne identifikabilnosti.

Konačno, treba biti naglašena i subjektivna komponenta u validaciji modela. Tehnike koje su predstavljene trebaju biti posmatrane kao savjetodavne za korisnika. Konačna odluka je ipak na korisniku. Ili po riječima autora Drapera i Smitha ‘’Pregledanje i analiza varijabli ne treba nikada biti ostavljena samo na tome da se primjene statističke procedure’’.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

ALATI : INTERAKTIVNI SOFTVER

Posao na dobijanju modela putem identifikacije je karakteriziran slijedećom sekvencom:

1. Specificiranje strukture modela
2. Računarski program nalazi najbolji model unutar ove izabrane strukture.
3. Evaluacija osobina ovog modela
4. Testiranje nove strukture modela i povrtak na korak 1.

Ovo je prikazano i organigramom na narednoj slici br. 17.1. Prva stvar koja zahtjeva pomoć je kako izračunati model i evaluirati njegove osobine. Postoji dosta softverskih programa za identifikaciju sistema koji nude ovu pomoć. Ovi programi tipično sadrže slijedeće rutine:

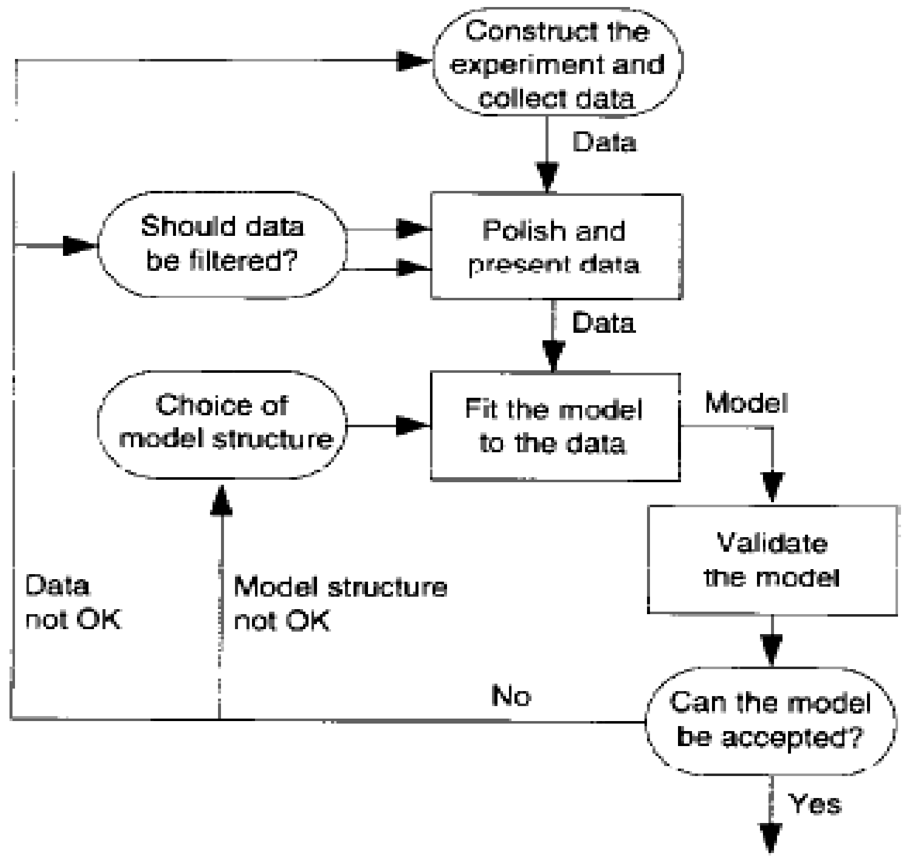
A. *Manipulacija sa podacima, plotiranje i slićne operacije*

Ovo uključuje fitriranje podataka, otklanjanje drifta, izbor segmenata podataka u zapisu, itd.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

B. *Neparametarski identifikacioni metodi*

Procjenjivanje kovariansi, Fourrier-ove transformacije, koralacija i spektralna analiza itd.



Slika br. 17.1 Identfikacioni ciklus i organigram soft. programa

17. IDENTIFIKACIJA SISTEMA U PRAKSI

C. *Metode parametarskih procjena*

Izračunavanje parametarskih estimacija u različitim strukturama modela.

D. *Prezentacija modela*

Simulacija modela, estimacija i crtanje polova i nula, računanje frekventnih funkcija i plotiranje Bode-ovih dijagrama, itd.

E. *Validacija modela*

Računanje i analiza reziduala. $\{\varepsilon(t,\hat{\theta}_N)\}$: ; poredjenje izmedju različitih osobina modela i slićno.

Različiti programski paketi se uglavnom razlikuju u izgledu i obliku korisničkog interfejsa (GUI) i različitim opcijama u izboru struktura modela.

Jedan od najpoznatijih paketa je MathWorks-ov paket System identification toolbox (SIT). koji je jedan od tollboksova Matlaba, razvijen od strane Ljunga.

Komandna struktura je data u programskom okruženju Matlaba sa konceptom radnog prostora (WORKSPACE) i MACRO mogućnostima u formi m-fajlova. SIT paket omogućava korištenje svih struktura modela tipa crne kutije (black-box) sa proizvoljnim brojem ulaza. ARX modeli i modeli u prostoru stanja sa proizvoljnim brojem ulaza i izlaza su takodjer mogući. Nadalje, korisnik može definisati proizvoljne korisnički kreirane linearne modele u prostoru stanja u diskretnom i kontinualnom vremenu. GUI pomaže korisniku da arhivira sve identificirane modele kao i da mu pomogne u korištenju programa.

Drugi softverski paketi opšteg tipa su PIM (autor Landau) , Dynamod(Midé), ESTIMA (DLR), Identifikacioni modul ISIM u okviru softverskog paketa MatrixX od NI, Frequency domain Identification toolbox u Matlabu.

Praktična strana identifikacije sistema

Kako je više puta do sada rečeno, najvažniji element u postupku identifikacije, nakon što su podaci sa procesa prikupljeni, je da se probaju razne strukture modela, izračuna najbolji model u tim strukturama, a zatim da se validira taj model. Tipično, ovo treba biti ponovljeno za nekoliko različitih struktura modela prije nego što se nadje zadovoljavajući model. Poteškoće ovog procesa ne treba podcjeniti, i zahtjevaju dosta iskustva prije nego što se njima potpuno ovlada.

Procedura koju preporučuje autor Ljung je slijedeća:

Korak 1. *Posmatranje podataka*

Iscrtati podatke. Pažljivo ih posmatrati. Pokušati uočiti dinamiku u njima. Možemo li uočiti efekte u izlazima od promjena u ulazima? Da li nelinearni efekti se mogu uočiti, kao različiti odzivi na različitim nivoima signala, ili različiti oblici odziva na porast u step signalu i smanjenje (negativni) step signal? Da li ima dijelova zapisa signala koji izgledaju ‘’razmazano’’ (messy) ili pak ne nose nikavu informaciju. Iskoristiti ovu inspekciju podataka i njihovu vizuelnu analizu i da se izaberu dijelovi podataka za namjene estimacije i validacije.

Da li fizički nivoi signala igraju neku ulogu u modelu? Ako ne, potrebno je detrendirati podatke otklanjanjem iz njih srednje vrijednosti. Modeli će nakon toga opisivati kako promjene u ulazu daju promjene na izlazu, ali neće objasniti stvarne nivoe signala. Ovo je normalna situacija

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Defaultna situacija, sa dobrim podacima je da se detrendira sa otklanjanjem srednje vrijednosti, a zatim izaberu prve dvije trečine podataka za namjene estimacije a korištenje ostatka podataka za validaciju.

Ovaj defaultni postupak se automatski primjenjuje na podacima ako se izabere opcija ‘’Data Quickstart’’ u SIT programu Matlaba.

Korak 2. *Razvoj osjećaja za poteškoće*

Izračunati i prikazati frekventni odziv procjene spektralne analize, impulsni odziv procjene korelacione analize, kao i četvrti red ARX modela sa kašnjenjem koje je procjenjeno iz korelacione analize, i defaultni red modela u prostoru stanja koji je sračunat pomoću metoda podprostora. Sve ovo korespondira sa komandom ‘’Estimate Quickstart’’ u SIT boksu.

Tražiti slaganje izmedju:

- Procjene spektralne analize i frekventnih funikcija ARX i modela u prostoru stanja

17. IDENTIFIKACIJA SISTEMA U PRAKSI

- Procjene korelacione analize i tranzijentnih odziva ARX i modela u prostoru stanja
- Mjerenih validacionih izlaznih podataka i ARX simuliranih izlaza iz ARX i modela u prostoru stanja. Ovo se naziva plotom izlaza modela (Model output plot)

Ako su ova slaganja rezonska, problem nije tako težak i relativno jednostavni linearni model će uraditi vrlo dobar posao. Malo podešavanje reda modela i modela šuma će možda biti potrebno , i nakon toga možemo nastaviti na korak 4. Ukoliko ovo nije slučaj, trebamo nastaviti sa korakom 3.

Korak 3. *Ispitivanje poteškoća*

Može postojati nekoliko razloga zašto poredjenja u koraku 2 nisu dobro izgledala i nisu uspjela. Ovaj korak diskutuje najčešće od tih razloga i kako ih prevazići:

- **Model je nestabilan**.   ARX ili model u prostoru stanja se mogu pokazati da su nestabilni, ali mogu još uvjek biti korisni za namjene upravljanja. U tom slučaju treba preći na

5 ili 10 koraka unaprijed predikciju umjesto simulacije, kada se posmatra slaganje izmedju mjerenih izlaza i izlaza iz modela.

- **Feedback u podacima** :

Ako postoji feedback sa izlaza na ulaz , zbog postojanja nekog regulatora, tada procjene spektralne i korelacione analize, kao i model u prostoru stanja nisu pouzdani. Neslaganja izmedju ovih procjena i ARX modela se mogu u ovom slučaju zanemariti. U analizi reziduala kod parametarskih modela, feedback u podacima može takodjer biti vidljiv kao korelacija izmedju reziduala i ulaza za negativna kašnjenja.

- **Model šuma**:

Ako model u prostoru stanja je vidno bolji nego ARX model u reprodukciji mjerenih izlaza, ovo je indikacija da smetnje imaju značajan uticaj, i da je nužno da se pažljivo modeliraju.

- **Red modela**:

Ako model četvrtog reda ne daje dobar plot modela izlaza, treba pokušati sa osmim redom. Ako se uklapanje značajno poboljšava, slijedi da modeli većeg reda su potrebni, ali i da su linearni modeli dovoljni.

- **Dodatni ulazi:**

Ako uklapanje izlaza modela se nije značajnije popravilo sa testovima do sada, potrebno je razmisliti o fizikalnosti aplikacije čiji se model identificira. Da li postoji više signala nego što je bilo, ili je moglo biti mjereno, koji mogu da utiču na izlaz? Ako je tako, uključiti i ove u ulaze i pokušati ponovno sa ARX modelom četvrtog reda od svih ulaza. Primjetimo da pri tome, svi ulazi ne moraju biti kontrolni signali, bilo što što je mjerljivo, uključujući i smetnje, treba biti tretirano kao ulaz.

- **Nelinearni efekti**:

Ako uklapanje izmedju mjerenih vrijednosti izlaza i izlaza iz modela je još uvjek loše, razmatrati ponovno fizikalnost aplikacije. Da li ima nelinearnih efekata u sistemu? U tom slučaju ako ima, formirati nelinearnosti iz mjerenih podataka. Ovo može biti jednostavno kao formiranje proizvoda mjerenja struja i napona ako je to električna snaga kao ulaz u sistem kod recimo procesa grijanja, gdje je temperatura izlazna varijabla. Ovo je dakle zavisno od same aplikacije.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Ne košta mnogo da se formira još čitav niz ulaza putem nekih razumnih nelinearnih transformacija mjerenih signala, i onda provjeriti da li ovakvi nelinearno formirani ulazi poboljšavaju uklapanje.

- **Opšte nelinearno mapiranje**:

U nekim aplikacijama fizikalni uvid može nedostajati, tako da je teško da se dodje do strukturnih nelinearnosti na fizikalnim osnovama. U takvim slučajevima, nelinearni modeli crne kutije (nonlinear black box) mogu biti rješenje.

- **Još uvjek ima problema** :

Ako nijedan od ovih test metoda nije doveo do modela koji je u stanju da razumno dobro reprodukuje validacione podatke, zaključak može biti da se dovoljno dobar model ne može ni proizvesti iz podataka. Za ovo mogu postojati mnogi razlozi. Jedan od najvažnijih je da podaci jednostavno ne sadrže dovoljno informacija, napr. zbog lošeg odnosa signala prema šumu, velikih i nestacionarnih smetnji, varirajućih osobina sistema, itd.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

U slučaju uspjeha jedne od ovih preporučenih metoda, iskoristiti uvid koji je ostvaren da se vidi koje ulaze koristiti ili koji red modela, i produžiti na korak 4.

Korak 4 : *Fino podešavanje reda i struktura šuma*

Za realne podatke ne postoji nešto što bi nazvali ‘’ korektna struktura modela’’. Medjutim, različite strukture mogu dati vrlo različite kvalitete modela. Jedini način da ovo nadjemo je da probamo niz različitih struktura modela i poredimo osobine dobijenih modela. Postoji nekoliko stvari na koje treba obratiti pažnju kod ovih poredjenja:

- **Uklapanje izmedju simuliranih i mjerenih izlaza:**

Gledati za uklapanje (fitovanje) izmedju simuliranih izlaza modela i mjerenih izlaza u okviru validacionih podataka. Formalno, izabrati onaj model, kod kojeg je ovaj broj najveći. U praksi je bolje biti pragmatičan i uzeti u obzir pored ovog numeričkog podatka i kompleksnost modela, kao i da li su važne osobine u odzivu sistema ‘’uhvaćene’’ od strane modela.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

• **Test analize reziduala**:

Za dobar model, kros korelaciona funkcija izmedju reziduala i ulaza ne izlazi značajno van regiona pouzdanosti. Direktan pogled u kašnjenje k koraka, pokazuje da efekat sa ulaza u(t-k) na y(t) nije korektno opisan. Pravilo ‘’od oka’’ je da sporo varirajuća kroskorelaciona funkcija van regiona povjerenja je indikacija o nedovoljnom broju polova, dok oštri vrhovi indiciraju nedovoljan broj nula, ili pogrešna vremena čistog kašnjenja.

Za model koji će se koristiti za potrebe analize i sinteze sistema upravljanja, vrlo je vrijedno prikazati rezultate rezidualne analize u frekventnom domenu.

• **Poništavanje polova i nula** :

Ako plot polova i nula (uključujući intervale povjerenja) indicira poništavanje polova i nula u dinamici, onda ovo sugeriše da se mogu koristiti modeli nižeg reda. Naročito, ako se pokaže da red ARX modela se treba povećati da bi se dobilo dobro uklapanje, ali su indicirana i poništenja polova i nula, onda su dodatni polovi uvedeni zato da bi se opisao šum. Tada treba pokušati sa ARMAX, OE ili BJ strukturama modela, sa A i F polinomima reda jednakog broju neponištenih polova.

Koje strukture modela trebaju biti testirane?

Često je potrebno samo nekoliko sekundi da se izračunaju i evaluiraju modeli u nekoj strukturi modela, tako da treba biti otvoren i blagonaklon prema ovim raznim testiranjima. Ipak, iskustvo pokazuje da kada su osnovne karakteristike ponašanja sistema ‘’ulovljene’’ , nema mnogo smisla ići sa finim podešavanjem redova sistema do beskonačnosti, samo da bi se uklapanje poboljšalo za neki dio procenta. Za ARX modele i modele u prostoru stanja, procjenjene preko metoda podprostora, postoje takodjer efikasni algoritmi da se može paralelno manipulisati sa mnogo struktura modela.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Multivarijabilni sistemi

Multivarijabilni sistemi su često mnogo veći izazov za modeliranje putem identifikacije. Naročito, sistemi sa nekoliko izlaza mogu biti vrlo teški. Osnovni razlog za ovo je da kuplovanja izmedju nekoliko ulaza i izlaza vode ka vrlo kompleksnim modelima, i strukture koje se javljaju su kompleksne i biće potrebno odrediti mnogo više parametara da se dobije dobro uklapanje.

Općenito govoreći, preferira se raditi sa modelima u prostoru stanja u multivarijabilnom slučaju, jer je lakše se nositi sa kompleksnošću strukture modela. U tom slučaju u suštini se radi o izboru reda modela.

Rad sa podskupovima ulazno-izlaznih kanala

U procesu identifikacije dobrih modela sistema, često je korisno izabrati podskupove ulaznih i izlaznih kanala. Parcijalni modeli ponašanja sistema će se onda konstruisati. Može biti naprimjer nejasno, da li svi mjereni ulazi imaju značajan uticaj na izlaze. Ovo se najlakše testira na taj način da se ukloni jedan ulazni kanal iz podataka, gradeći model kako izlaz (izlazi) zavise od preostalih ulaznih kanala, i provjeravajući da li postoji značajnija degradacija u uklapanju (fitovanju) izlaza iz modela sa mjerenim izlazima (vidjeti i raniju diskusiju u okviru koraka 3).

Općenito govoreći, uklapanje će postati bolje kada se više ulaza uključi a lošije kada se uključi više izlaza. Da bi razumjeli zašto je to tako kod povećanja broja izlaza, treba shvatiti da model koji treba da objasni ponašanje nekoliko izlaza ima teži zadatak nego onaj koji treba da brine samo za uklapanje jednog izlaza.

Ako ima poteškoća da se dobiju dobri modeli za višeizlazni sistem , tada može biti pametno da se modelira po jedan izlaz po modelu, da bi se vidjelo koji su izlazi teški za uklapanje i analizu. Modeli koji će se koristiti samo za svrhe simulacije, mogu biti vrlo dobro izgradjeni samo od pojedinačnih modela sa jednim izlazom. Medjutim, modeli za predikcije i upravljanja bi bili u stanju da daju bolje rezultate kada bi bili tako konstruirani da daju sve izlaze simultano tj istovremeno. Ovo slijedi iz činjenice da, znajući set svih prethodnih izlaznih kanala, daje bolju osnovu za procjenu, nego kada se samo znaju prošli izlazi u jednom kanalu.

Korak 5 : *Prihvatanje modela*.

Finalni korak je da se prihvati, barem privremeno, model koji će se koristiti za namjeravanu aplikaciju. Primjetimo slijedeće:

Bez obzira kako dobro izgleda procjenjeni model na ekranu PC sa SIT paketom, on je samo ‘’ulovio’’ jednostavan odraz realnosti. Medjutim, iznenadjujuće dobro, ovo je dovoljno za racionalno donošenje odluka i dobro odlučivanje u praksi analize i sinteze sistema automatskog upravljaja (SAU) procesima i poslovnim dogadjajima.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Šta identifikacija sistema može da ponudi

Tehnike identifikaciije sistema formiraju raznovrsne alate za mnoge probleme u nauci i inženjerstvu. Ove tehnike su sa svoje strane zavisne od aplikacija. Vrijednost razvijenih alata se dokazala u mnogobrojnim primjerima.

Medjutim, ipak postoje neka ograničenja koja su udružena sa ovim tehnikama o kojima je potrebno dati neke komentare.

Da li su metode adaptivnog i robusnog dizajna SAU učinile modeliranje suvišnim?

Kao što je to ranije rečeno više puta, modeliranje dinamičkih sistema i identifikacija ovih modela, je korisno u mnogim prilikama i za mnoge namjene kao : predikcije, upravljanja, simulacije, dizajn filtera, rekonstrukcija mjerenih podataka iz šuma, itd.

Ponekad se pojavi u naučnoj javnosti izjava da se potreba za modelom može prevazići na taj način da se razvije mnogo elaboriranije rješenje sistema upravljanja kao:

17. IDENTIFIKACIJA SISTEMA U PRAKSI

- adaptivni mehanizmi gdje parametri odlučivanja se direktno podešavaju,
- robusni dizajn koji je neosjetljiv na korektnost modela kojim taj sistem upravlja. Treba ipak primjetiti, da se adaptivne šeme tipično mogu interpretirati kao rekurzivni identifikacioni algoritmi koji su primjenjeni na specifičnu strukturu modela (tj. na model koji je parametriziran po članovima odgovarajučeg optimalnog regulatora).

Time je karakteristika gradnje modela vrlo mnogo prisutna i u adaptivnim mehanizmima.

Robusni dizajn je baziran na nominalnom modelu i odredjen je tako da je dobar rad osiguran čak ako i stvarni sistem odstupa od nominalnog modela. Obićno, okolina oko nominalnog modela se može specificirati unutar koje je prihvatljiva degradacija performanse sistema. Tada je vrlo korisna činjenica, da modeli dobijeni sa identifikacijom sistema se mogu isporučiti sa tagom kvaliteta, kao procjenjena odstupanja od istinskog opisa u parametarskom domenu ili u frekventnom domenu. Takvi modeli su onda prikladni i za robusni dizajn.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Ograničenja kvaliteta podataka

Jasno je da je ograničenje korištenja tehnika identifikacije sistema vezano sa raspoloživošću dobrih podataka i dobrim strukturama modela. Bez razumno dobrog zapisa podataka ne može se mnogo uraditi na polju identifikacije procesa sa kojeg su podaci, i postoji nekoliko razloga zašto se takvi zapisi ne mogu dobiti u nekim aplikacijama. Prvi i vrlo razumljiv razlog je da je vremenska skala procesa vrlo spora tako da će bilo koji informativni zapis podataka sa takvih procesa biti kratak i nedovoljan. Ovo je prije svega karakteristika za ekološke i ekonomske sisteme.

Drugi razlog je da ulaz možda neće biti raspoloživ za manipulacije, ili zbog same svoje prirode, ili zbog razloga sigurnosti ili samog odvijanja procesa proizvodnje. Nadalje odnos signala prema šumu može takodjer da bude loš, i identifikabilnost (tj. raspoloživost informativnog seta podataka) se možda ne može garantirati. Loš odnos korisnog signala prema šumu se može, barem teoretski, kompenzirati da se uzimaju duži zapisi podataka. Čak ako postrojenje i dozvoljava duga vremena eksperimentisanja, to ne mora uvjek biti rješenje, pošto vremenske varijacije u procesu, drift, spore smetnje i drugi uzroci ovo dodatno opterečuju.

Konačno, kada nam je i dozvoljeno da manipulišemo sa ulazima, možemo provoditi eksperimente i duži period vremena i imamo dobar odnos izmedju signala i šuma, još uvjek može biti teško dobiti dobre zapise podataka. Primarni razlog za ovo je prisutvo nemjerljivih smetnji koje se ne uklapaju u standardnu sliku ‘’stacionarnih stohastičkih procesa’’. Ovo se može djelomično kompenzirati sa korištenjem robusnih normi, i ta mjera često može biti uspješna. Medjutim i dalje ostaje činjenica: kvalitet podataka mora biti od primarne važnosti kod svake identifikacije. Ovo ujedno odredjuje i cijenu svakog identifikacionog poduhvata, jer dobijanje kvalitetnih zapisa podataka iz dobro organizovanog i vodjenog eksperimenta može biti i značajno skup poduhvat.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Ograničenja : Strukture modela

Trivijalna je činjenica da loša struktura modela ne može ponuditi dobar model, bez obzira na raspoloživost dobrih podataka. Na primjer, ARX struktura modela ne može nikada obezbjediti dobar opis sistema iako se mogu imati podaci koji su vrlo kvalitetni i skupljani dug period vremena, ako u sistemu imamo prisutnu neku statičku nelinearnost. Krucijalni nelinearni mehanizmi moraju biti ugradjeni u model, a ovo zahtjeva da se za takav proces ostvari i neki fizikalni uvid i razmatranje.

Dakle prvi problem je da li proces (oko radne tačke od interesa) dozvoljava standardni, linearni, black-box opis modela sistema, ili se mora konstruisati neki specifičan – kastomizirani model za taj proces. U prvom slučaju, naše šanse za uspjeh su dobre; u drugom, mi moramo da ostvarimo odredjeni fizikalni uvid prije nego što se model može procjeniti, ili da se nadamo da se nelinearna dinamika može ‘’uloviti’’ pomoću strukture nelinearne crne kutije. Očito je da je ovaj problem zavisan od konkretne aplikacije i o njemu nema mnogo diskusija u literaturi o identifikaciji.

17. IDENTIFIKACIJA SISTEMA U PRAKSI

Zbog toga moramo još jednom naglasiti da : razmišljanje, intuicija, i ostvarenje uvida u srž problema, nikad ne mogu postati suvušni i biti zamjenjeni nekim ‘’super inteligentnim algoritmima i metodama automatske konstrukcije modela procesa’’ koji se identificira.