

Identifikacija

Univerza v Ljubljani, Fakulteta za elektrotehniko
Podiplomski študijski program 2. stopnje, 1. letnik
Izbirni predmet Modula A
Poletni semester 2013/2014
Sašo Blažič, Drago Matko



Modeli kot abstrakcija realnosti

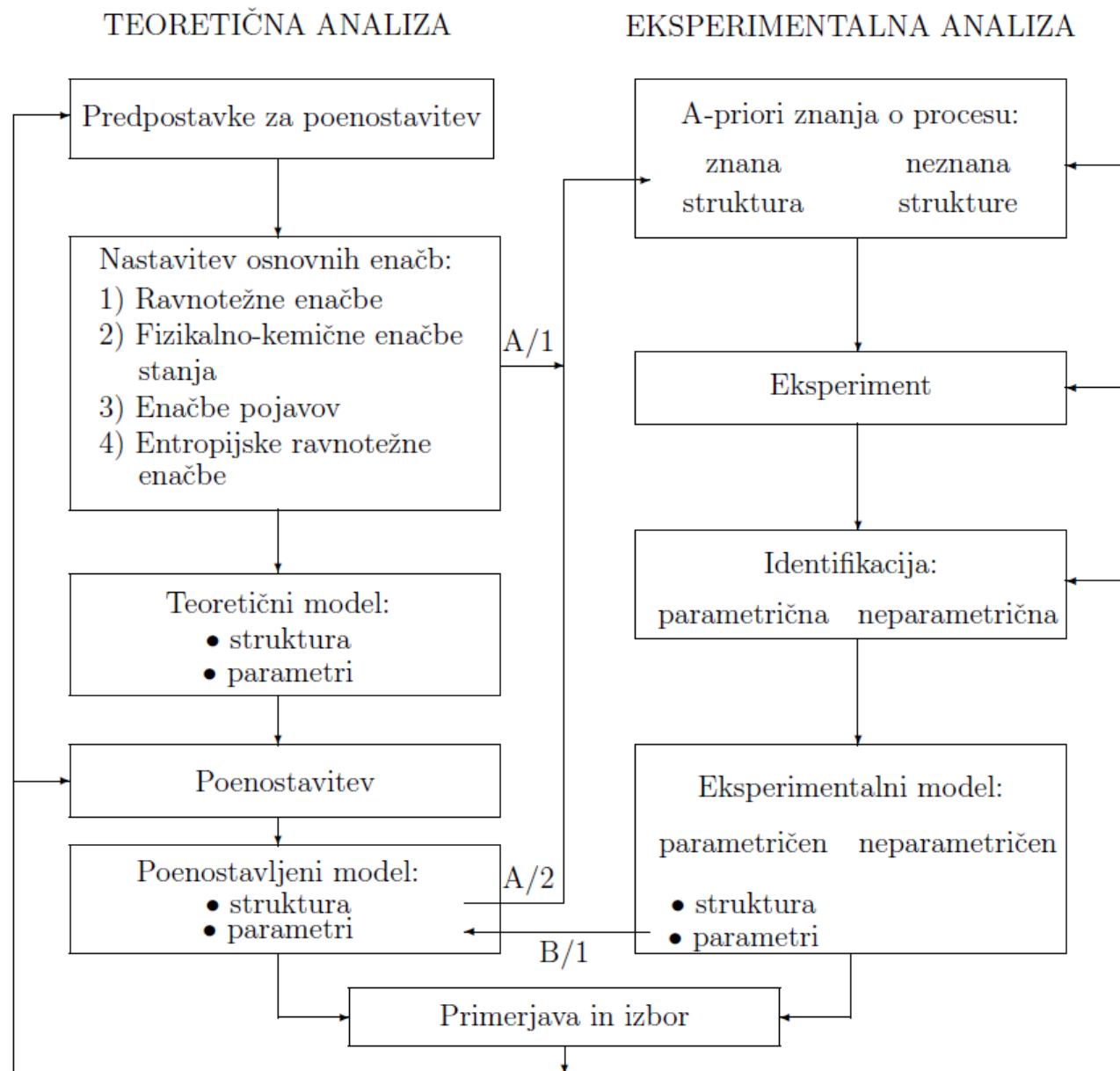


- Gradnja modelov s pomočjo opazovanja in študij njihovih lastnosti sta bistveni sestavini modernih znanosti
- Modeli imajo lahko bolj ali manj formalni značaj, vendar je njihova osnovna lastnost, da skušajo združiti opazovanja v nek vzorec (zgled), ki ima enake lastnosti kot opazovani sistem
- Sistem je omejena ureditev procesov, ki vplivajo eden na drugega, proces pa označuje pretvorbo in/ali transport (prenos) materije, energije in/ali informacije (DIN 66201)
- Obravnavali bomo le *matematične modele*, ki omogočajo opis časovnega (statičnega in dinamičnega) obnašanja poljubnih sistemov (tehničnih, bioloških, ekonomskih) in predstavljajo bistveno sestavino moderne *sistemske teorije*
- Matematični model lahko dobimo s postopki teoretičnega modeliranja



- Alternativni način modeliranja je gradnja modela na podlagi analize merjenih signalov procesa – ta postopek se zato imenuje **eksperimentalno modeliranje** ali **identifikacija**
- Cilj identifikacije je torej določiti model sistema na osnovi podatkov iz sistema
- Pristop je zelo privlačen, saj lahko do modela pridemo (vsaj v teoriji) brez poznavanja fizikalnih lastnosti sistema
- Žal pa so merjeni podatki (skoraj) vedno nenatančni in netočni, saj so vsi senzorji podvrženi določenim napakam
- Druga težava, ki nastopi, je, da so modeli vedno nepopolni – realni sistemi so običajno bistveno bolj kompleksni kot njihovi poenostavljeni modeli, zato je iskanje idealne slike realnega sistema nesmiselno – ključno je, da dobimo model, ki je **uporaben** oz. zadosti svojemu **namenu**

Teoretično modeliranje vs. identifikacija



Teoretično modeliranje vs. identifikacija



- Rezultat teoretičnega modeliranja je sistem navadnih in/ali parcialnih diferencialnih enačb; če ga uspemo rešiti, dobimo teoretični model procesa z določeno strukturo in parametri
- Pogosto je ta model obsežen in kompliciran in ga moramo poenostaviti pred nadaljnjo uporabo
- Tudi če ne moremo eksplicitno rešiti sistema enačb, dajejo posamezne enačbe pomembne napotke o strukturi modela (n.pr. ravnotežne enačbe so vedno linearne in enačbe pojavov linearne na širokih območjih, fizikalno-kemične enačbe stanj pa pogosto uvajajo nelinearne odvisnosti)
- Pri identifikaciji dobimo matematični model procesa iz meritev (predhodna znanja o procesu, ki smo jih pridobili n.pr. s teoretično analizo ali z analizo meritev, so bistvena za določitev strukture – pri linearnih procesov red)
- Naslednji korak je zajem meritev iz procesa in ovrednotenje signalov s pomočjo identifikacijskega postopka, s čimer izrazimo povezavo vhodnega in izhodnega signala z nekim matematičnim modelom
- Vhodni signali so lahko naravni, v procesu nastopajoči delovni signali ali umetno uvedeni preskusni signali
- Z ozirom na namen uporabe izvedemo identifikacijo parametričnih ali neparametričnih modelov

Teoretično modeliranje vs. identifikacija



- Teoretični model vsebuje funkcionalno povezavo med fizikalnimi podatki o procesu in njegovimi parametri
- V eksperimentalnem modelu nasprotno nadomeščajo parametre le številčne vrednosti, katerih funkcijska zveza s fizikalnimi podatki o procesu ostane neznana (običajno pa točneje opišemo trenutno dinamično obnašanje procesa ali to izpeljemo z manjšimi stroški)
- Teoretično in eksperimentalno modeliranje se dopolnjujeta
- S primerjavo modelov nastane povratna zveza v postopku
- Modeliranje je v splošnem iterativni postopek
- Idealno je kombinirati oba pristopa: običajno s teoretičnim modeliranjem določimo strukturo, z identifikacijo pa parametre



Teoretično modeliranje vs. identifikacija

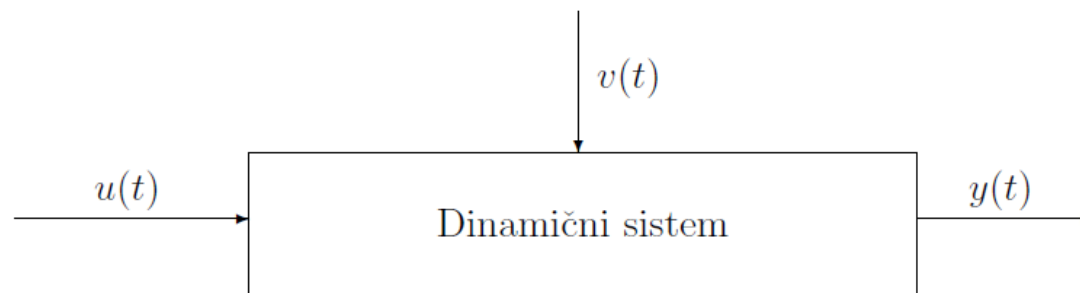
Teoretično modeliranje	Identifikacija
Struktura modela je posledica naravnih zakonov	Strukturo modela moramo privzeti
Opis obnašanja notranjih spremenljivk stanja in vhodno/izhodnega obnašanja	Identificiramo le vhodno/izhodno obnašanje
Parametri modela so podani kot funkcije procesnih veličin	Parametri modela so čiste številčne vrednosti, v katerih ne vidimo povezave s fizikalnimi procesnimi veličinami
Model velja za ves razred tipa procesov pri različnih obratovalnih stanjih (mnoge procesne veličine so pogosto le nedoločeno poznane)	Model velja le za raziskovalni proces in določeno stanje obratovanja (zato lahko to obnašanje opišemo relativno točno)
Model lahko zgradimo tudi za nek proces, ki ne obstoja	Model lahko identificiramo le za nek obstoječ proces
Bistvene notranje poteke procesov moramo poznati z možnostjo matematičnega opisa	Ni potrebno poznati notranjih potekov procesov
	Ker so postopki identifikacije neodvisni od posameznih procesov, lahko enkrat razvito programsko opremo uporabimo pri mnogih različnih procesih
Večinoma zahteva veliko časa	Večinoma zahteva relativno manj časa



- *Identifikacija je eksperimentalna določitev časovnega obnašanja procesa z uporabo merjenih signalov. Časovno obnašanje je definirano znotraj nekega razreda matematičnih modelov in ponazarja identificirani proces tako, da so razlike med dejanskim procesom in njegovim matematičnim modelom čim manjše.*
- Ta definicija se naslanja na definicijo, ki jo je postavil Zadeh in vsebuje tri bistvene pojme:
 - Merjeni signali
 - Razred matematičnih modelov
 - Razlika (pogrešek) med dejanskim procesom in njegovim matematičnim modelom
- Vsi trije pojmi so bistveni pri delitvi identifikacijskih postopkov, ki jih bomo obravnavali v nadaljevanju



- Identificirani dinamični sistem lahko konceptualno opišemo tako, kot to prikazuje slika



- Na sistem vplivajo vhodne spremenljivke $u(t)$ in motnje $v(t)$
- Uporabnik lahko vpliva na vhodne signale, ne more pa vplivati na motnje
- Motnje delimo na takšne, ki jih lahko direktno merimo in na takšne, ki jih ne moremo direktno meriti, je pa njihov vpliv seveda viden posredno na izhodu $y(t)$



- Identifikacija statičnega modela 
 - Če nimamo prave strukture, se zdi, da vsaka meritev potrebuje svoj model
 - Če je množica modelov le posledica stohastične narave procesa, lahko poiščemo model, ki se najbolj prilega veliki množici meritev
- Identifikacija realnega procesa – kvadrokopterja 
 - Izredno pomemben korak je obdelava meritev pred identifikacijo
 - Določitev strukture se je spet izkazala za ključno
 - Tudi postopek iskanja modela je lahko problematičen – optimizacija obtiči v lokalnih minimumih ipd.
- Identifikacija pošumljenega statičnega modela 
 - Zaradi šuma ocena parametrov več ni enaka dejanski vrednosti
 - Pomembno je, da ocena z večanjem števila opazovanj konvergira k pravi vrednosti (lastnosti metod se v tem pogledu zelo razlikujejo)



- Vpliv porazdelitve šuma na identifikacijo 
 - Če lahko motnje opišemo z normalno porazdelitvijo, je varianca oz. standardna deviacija ocene parametra ustrezna mera zaupanja
 - Če imamo drugačno porazdelitev šuma, je potrebno podati porazdelitev ocene parametra
 - Zaradi centralnega limitnega teorema z večanjem števila vzorcev gostota verjetnosti ocene pogosto konvergira k normalni porazdelitvi
- Šum na vhodu (ki v tem primeru vstopa v regresor) 
 - Regresijska enačba povezuje regresor ψ , parametre θ in izhod y :
$$y = \psi^T \theta$$
 - Če je šum na regresorju (v tem primeru na toku), to povzroči sistematično napako ocene, ocena je namreč pristranska
 - Pristranskost je odvisna od razmerja signal/šum na regresorju
 - Pristranskost ni odvisna od šuma na izhodu (če sta motnji nekorelirani) – smiselno zamenjati izhode in regresorje, če možno



- Preveliko število ocenjevanih parametrov 
 - Če ocenjujemo preveliko število parametrov, so ocenjeni parametri med seboj korelirani in imajo veliko varianco
 - Ekstrapolacija modela je vedno problematična, v primeru preparametrizacije pa je problem še hujši



- Merjeni signali vsebujejo motnje
- Čas za izvedbo identifikacijskega postopka je omejen
- Velikosti vhodnega in izhodnega signala sta omejeni
- Oblika vhodnega signala je omejena z uporabljenim izvršnim členom, ti imajo predvsem naslednje prenosno obnašanje:
 - Proporcionalni pogoni izvršnih členov (pnevmatski in hidravlični)
 - Integrirni pogoni izvršnih členov s spremenljivo hitrostjo (hitrostno krmiljeni enosmerni motorji)
 - Integrirni pogoni izvršnih členov s konstantno hitrostjo (izmenični motorji s tropoložajnim stikalom)
 - Pogoni izvršnih členov s kvantizacijo (električni koračni motorji)
 - Pogoni izvršnih členov z dvopoložajnim stikalom (npr. grelne naprave)



1. **Definicija uporabe modela** – uporabnost modela je bistvena za njegovo ovrednotenje
2. **Zbiranje apriornih znanj** – temelji na dobrem poznavanju procesa in fizikalnih zakonitosti, na izkušnjah pri preteklem delu s procesom oziroma s sorodnimi procesi, na zbiranju statističnih podatkov o signalih (predvsem motilnih) → struktura modela
3. **Načrtovanje meritev**, kjer izberemo:
 - vhodne signale (naravno obratovalne, umetno preskusne, njihovo obliko, amplitudo in frekvenčne spektre),
 - čas vzorčenja in merilni čas,
 - meritev v odprti ali zaprti regulacijski zanki,
 - nesprotno ali sprotno identifikacijo,
 - aparaturno opremo,
 - filtriranje motenj



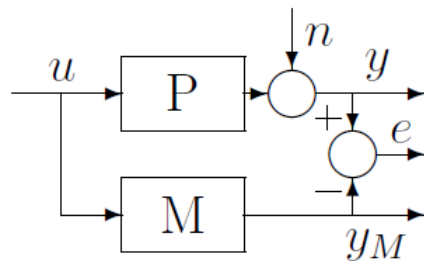
4. **Izvedba meritev**
 5. **Identifikacijski algoritem**, ki na podlagi meritev določi identifikacijski model
 6. **Preskus identificiranega modela** na signalih, ki niso bili uporabljeni pri identifikaciji (vrednotenje ali validacija)
 7. Če je vrednotenje negativno, je potrebno **ponoviti** nekatere zgornjih korakov
- Model identificiranega procesa dobimo le redko v eni potezi
 - Pogosto moramo npr. izvesti predhodne meritve in šele nato po analizi pristopiti k dejanski identifikaciji
 - Postopek identifikacije je zato v splošnem iterativen



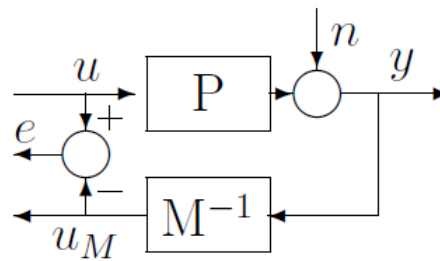
- Glede na razred matematičnih modelov:
 - Parametrične identifikacijske metode dajo kot rezultat parametrične modele, to je modele z eksplicitno izraženimi parametri (npr. diferencialne enačbe, prenosne funkcije).
 - Neparametrične identifikacijske metode dajejo kot rezultat neparametrične modele, ki podajajo vhodno-izhodno obnašanje v obliki tabel vrednosti oz. krivulje (n.pr. utežne funkcije, frekvenčni odzivi, odzivi na stopnico v tabelarični ali grafični predstavitvi)
- Glede na razred uporabljenih signalov:
 - Delitev glede na zveznost neodvisne spremenljivke:
 - zvezni
 - diskretni
 - Delitev glede na naravo signalov:
 - določeni (deterministični)
 - naključni
 - psevdonaključni (podobne lastnosti kot naključni, a umetno tvorjeni)



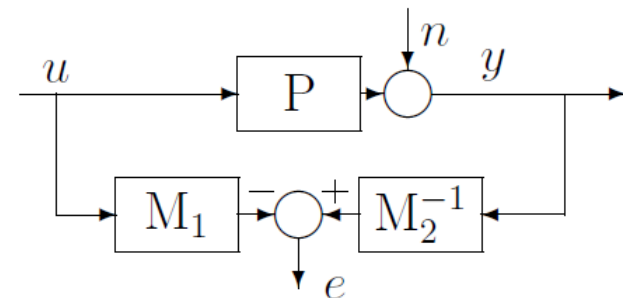
- Glede na pogrešek med procesom in modelom:
 - Izhodni (razlika med izhodom procesa in izhodom modela; model in proces sta priključena na isti vhodni signal)
 - Vhodni (razlika med vhodom procesa in izhodom inverznega modela, ki ima kot vhod izhod identificiranega procesa)
 - Posplošeni (model je razdeljen v dva dela; prvi del modela ima isti vhod kot identificirani proces; inverz drugega dela modela ima kot vhod izhod identificiranega procesa; posplošeni pogrešek je razlika med izhodoma obeh delov modela)



Izhodni pogrešek



Vhodni pogrešek



Posplošeni pogrešek



- Glede na sočasnost meritev in vrednotenja:
 - Nesprotne metode, kjer posnamemo signale na ustrezen medij in jih šele kasneje obdelamo in sicer vse podatke naenkrat.
 - Sprotne metode, kjer poteka obdelava sočasno s pojavljanjem novih parov vhodno-izhodnih signalov. Če dobivamo signale neposredno z meritvijo vhoda in izhoda identificiranega procesa, mora teči obdelava v **realnem času**. Obdelava v realnem času je torej sprotna obdelava, ni pa nujno, da je vsaka sprotna obdelava tudi obdelava v realnem času, saj lahko sprotno obdelujemo tudi signale, ki smo jih predhodno spravili na ustrezni medij in jih kasneje reproduciramo in obdelujemo na enak način kot če bi jih dobivali neposredno z meritvijo vhoda in izhoda procesa.
- Glede na uporabljeni postopek obdelave podatkov:
 - Nerekurzivne metode
 - Rekurzivne metode



- Parametrični modeli:
 - Zvezni modeli:
 - Diferencialne enačbe
 - Zvezne prenosne funkcije, ki jih dobimo iz diferencialnih enačb z Laplaceovo transformacijo ob predpostavki, da sistem ne vsebuje začetne energije
 - Zapis v prostoru stanj itn.
 - Diskretni modeli:
 - Diferenčne enačbe
 - Diskretne prenosne funkcije, ki jih dobimo iz diferenčnih enačb z z-transformacijo ob predpostavki, da sistem ne vsebuje začetne energije
 - Zapis v prostoru stanj itn.
- Neparametrični modeli:
 - Frekvenčni odziv
 - Odziv na stopnico, impulz itn.

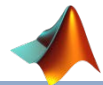


- Linearne časovno nespremenljive procese s koncentriranimi parametri opišemo z navadnimi linearnimi diferencialnimi enačbami s konstantnimi koeficienti v naslednji obliki
$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} \dot{y} + a_n y = b_0 u^{(n)} + b_1 u^{(n-1)} + \dots + b_{(n-1)} \dot{u} + b_n u$$
 - $u(t)$ vhod v proces, $y(t)$ njegov izhod, b_i in a_i so parametri procesa
- S pomočjo Laplaceove transformacije dobimo pri pogoju, da proces ne vsebuje začetne energije (vsi začetni pogoji v zgornji enačbi so enaki nič), prenosno funkcijo procesa:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \dots + b_{n-1} s + b_n}{s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n}$$

- Laplaceova transformacija izhoda $Y(s)$ je produkt prenosne funkcije $G(s)$ in Laplaceove transformacije vhoda $U(s)$
- Produktu v frekvenčnem prostoru ustreza konvolucija v časovnem:

$$y(t) = \int_0^\infty g(\tau) u(t - \tau) d\tau = \int_0^\infty u(\tau) g(t - \tau) d\tau$$





- Parametrični zvezni model je model zveznega procesa izražen s parametri, npr. parametri b_i in a_i prenosne funkcije
- Glede na konvolucijsko enačbo je možen neparametrični model odziv na enotin impulz $g(t)$, ki ga imenujemo impulzni odziv ali utežna funkcija
- Drugi neparametrični model je frekvenčni odziv $G(j\omega)$:
 - Interpretacija s harmoničnim vzbujanjem: Če vzbujamo stabilni linearni časovno nespremenljivi sistem s sinusnim signalom:
$$u(t) = U_0 \sin(\omega_0 t + \varphi_u)$$
se sistem odzove s harmoničnim nihanjem enake frekvence:
$$y(t) = Y_0 \sin(\omega_0 t + \varphi_y)$$
 - Razmerje amplitud je določeno z amplitudnim odzivom $\frac{Y_0}{U_0} = |G(j\omega_0)|$
 - Fazna razlika je določena s faznim odzivom $\varphi_y - \varphi_u = \angle[G(j\omega_0)]$



- Povezava med impulznim odzivom $g(t)$ in prenosno funkcijo $G(s)$:

$$G(s) = \mathcal{L}\{g(t)\} \quad g(t) = \mathcal{L}^{-1}\{G(s)\}$$

- Povezava med frekvenčnim odzivom $G(j\omega)$ in pren. funkcijo $G(s)$:

$$G(j\omega) = \lim_{s \rightarrow j\omega} G(s)$$

- Povezava med impulznim odzivom $g(t)$ in frekvenčnim odzivom $G(j\omega)$:

$$G(j\omega) = \mathcal{F}\{g(t)\} \quad g(t) = \mathcal{F}^{-1}\{G(j\omega)\}$$

- Operator $\mathcal{F}\{\cdot\}$ je Fourierova transformacija:

$$X(\omega) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt = \lim_{s \rightarrow j\omega} X(s)$$

- $x(t)$ je zvezni signal, $X(\omega)$ njegova Fourierova transformiranka

- Povezava med frekvenčnim odzivom in Fourierovimi transformirankami:

$$G(j\omega) = \frac{Y(\omega)}{U(\omega)}$$



- Strejceva metoda identifikacije:
 - Temelji na identifikaciji sistema pri vzbujujanju s stopničastim preizkusnim signalom
 - Je zelo enostavna, a je primerna za majhen nabor nezahtevnih procesov, ki pa so v praksi dokaj pogosti
- Metoda s prilagajanjem modela:
 - Je zelo intuitivna
 - Temelji na optimizaciji, zato je računsko precej zahtevna
 - Primerna za širok nabor procesov
 - Z ustreznimi prilagoditvami jo lahko uporabimo tudi za identifikacijo diskretnih, nelinearnih, časovno spremenljivih procesov

Strejceva metoda identifikacije



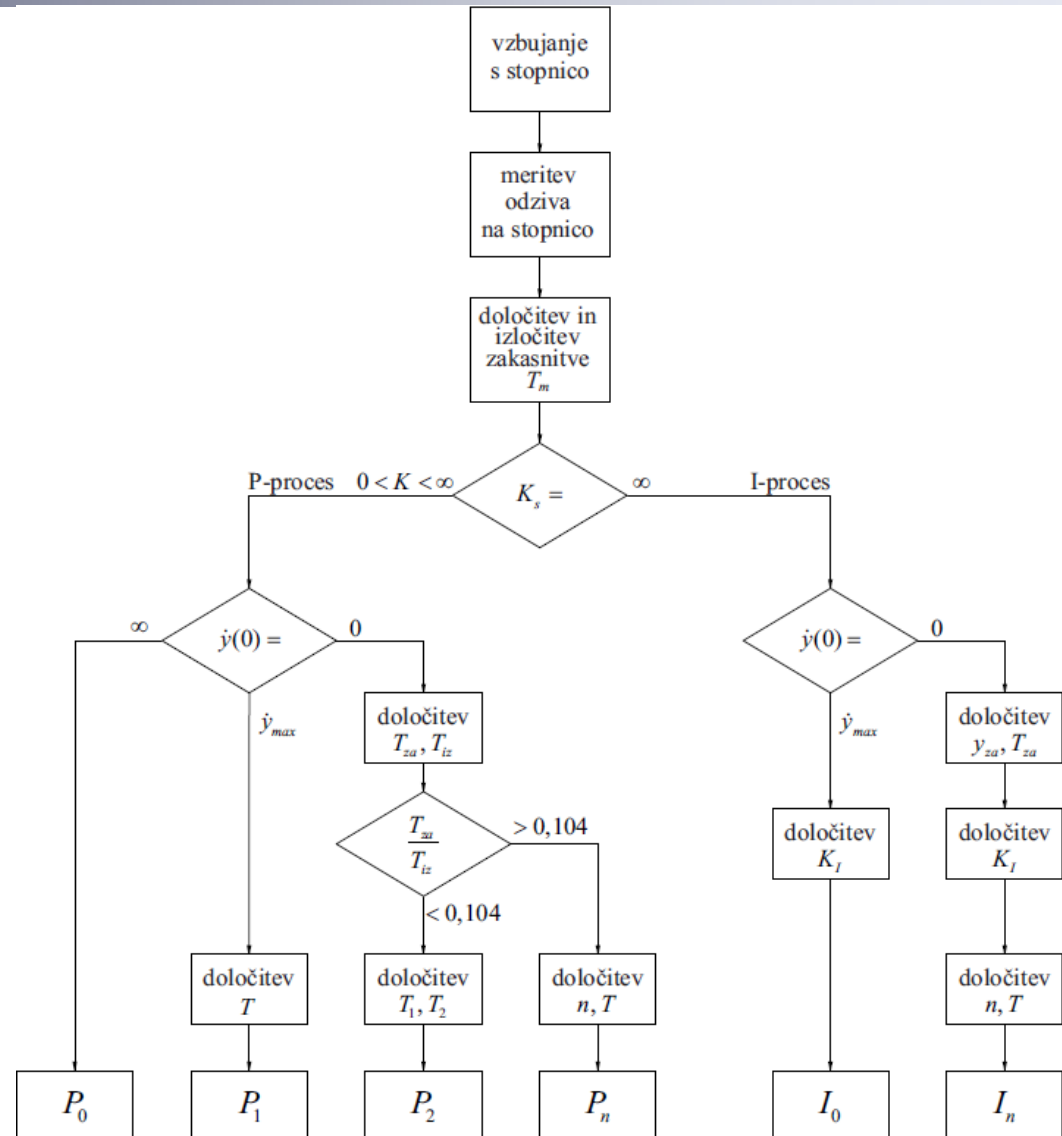
- Najpogostejše imamo predvsem za namene načrtovanja vodenja opravka s proporcionalnimi ali integrirnimi procesi
- Proporcionalni sistemi se pri vzbujujanju s stopničastim vzbujujalnim signalom odzovejo s končno spremembo izhoda
- Integrirni sistemi se na stopničasto vzbujujanje odzovejo z neomejenim odzivom, pri čemer je asimptota polinomska funkcija (pri sistemu 1. vrste premica)

PROCES	ODZIV NA STOPNICO U_H (poenostavljeni)	KARAKTERISTIČNI PARAMETRI		
		K_S	$\dot{Y}_{max}$	T_{mn}
P_0		$K_{PS} = \frac{Y_H}{U_H}$	∞	0
P_1		$K_{PS} = \frac{Y_H}{U_H}$	$\frac{Y_H}{T_S}$	0
P_n (aperiodični) $n \geq 2$		$K_{PS} = \frac{Y_H}{U_H}$	$\frac{Y_H}{T_{iz}}$	T_{za} (T_{mn})
T_m		$K_{PS} = \frac{Y_H}{U_H}$	∞	T_m
I_0		∞	$K_I U_H$	0
I_n $n \geq 1$		∞	$K_I U_H$	T_{za} (T_{mn})

Strejceva metoda identifikacije



- Diagram poteka za izbor enega od enostavnih modelov



Proporcionalni sistemi drugega reda – P_2



- Model izberemo, če:

$$\frac{T_{za}}{T_{iz}} < 0,104$$

- Model procesa:

$$G(s) = \frac{K_s}{(T_1s + 1)(T_2s + 1)}$$

- Za določitev potrebujemo:

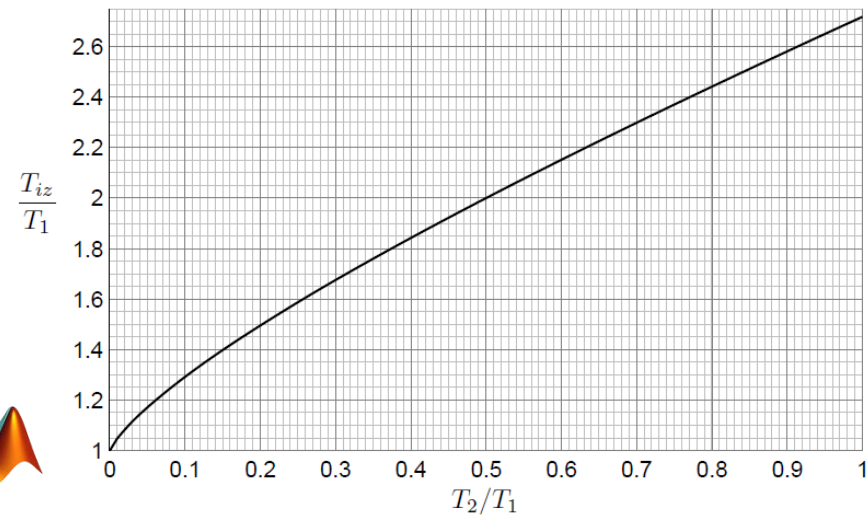
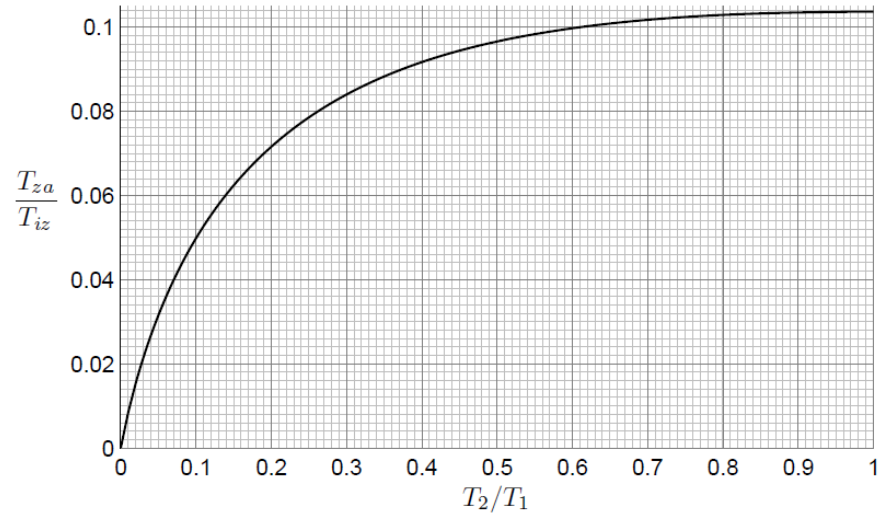
- T_{za}

- T_{iz}

- $K_s = \frac{\Delta Y}{\Delta U}$

- Postopek:

$$\frac{T_{za}}{T_{iz}} \rightarrow \frac{T_2}{T_1} \rightarrow \frac{T_{iz}}{T_1} \rightarrow T_1 \rightarrow T_2$$



Proporcionalni sistemi višjega reda – P_n



- Model izberemo, če:

$$\frac{T_{za}}{T_{iz}} > 0,104$$

- Model procesa:

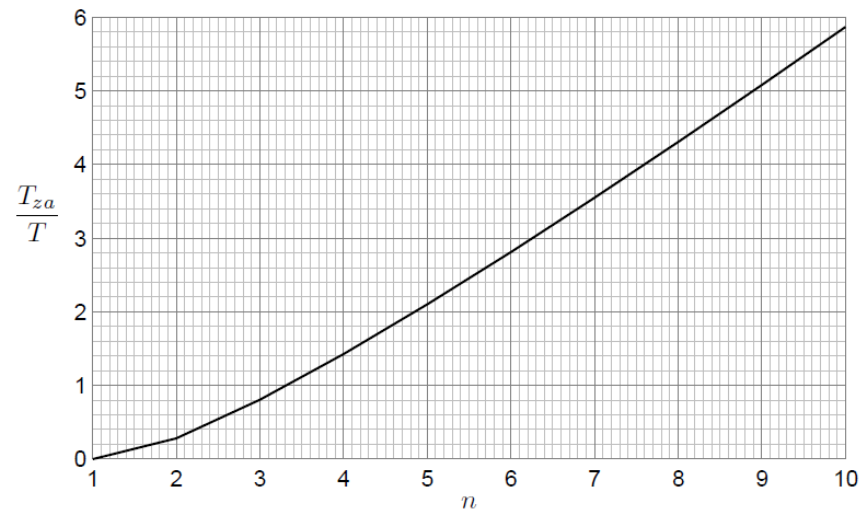
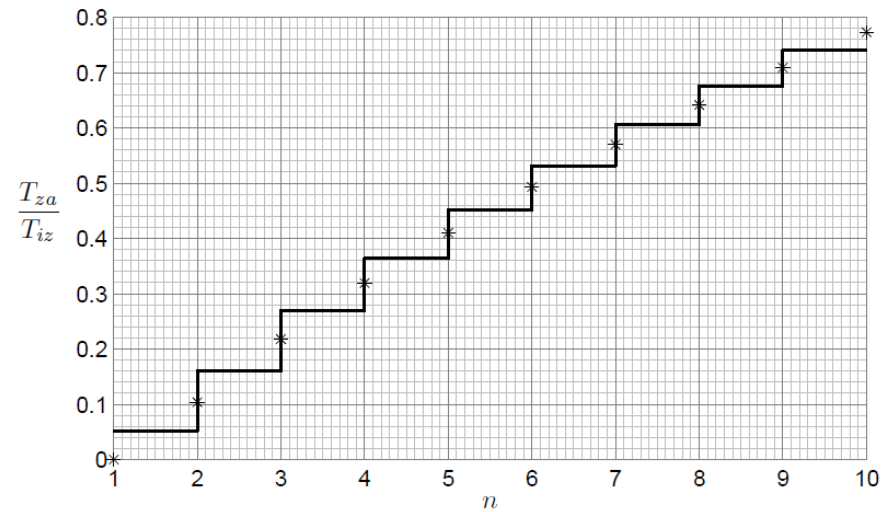
$$G(s) = \frac{K_s}{(Ts + 1)^n}$$

- Za določitev potrebujemo:

- T_{za}
- T_{iz}
- $K_s = \frac{\Delta Y}{\Delta U}$

- Postopek:

$$\frac{T_{za}}{T_{iz}} \rightarrow n \rightarrow \frac{T_{za}}{T} \rightarrow T$$



Integrirni sistemi prve vrste



- Integrirni sistem ima pole pri $s = 0$
- Vrsta – število polov pri $s = 0$
- Model procesa torej:

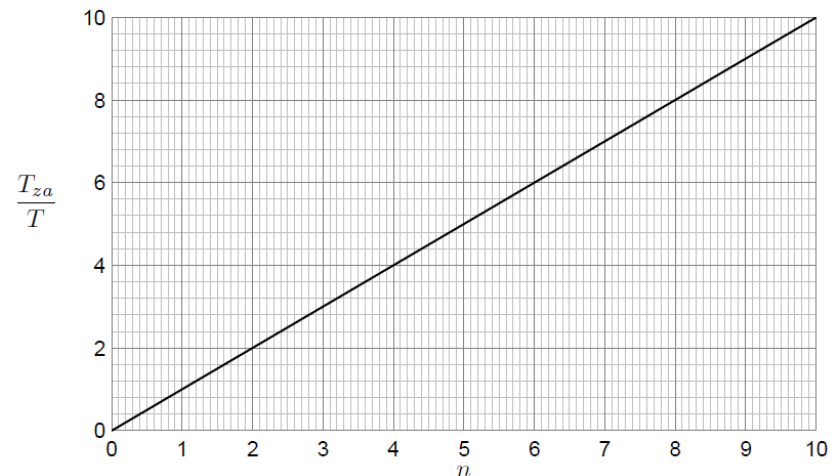
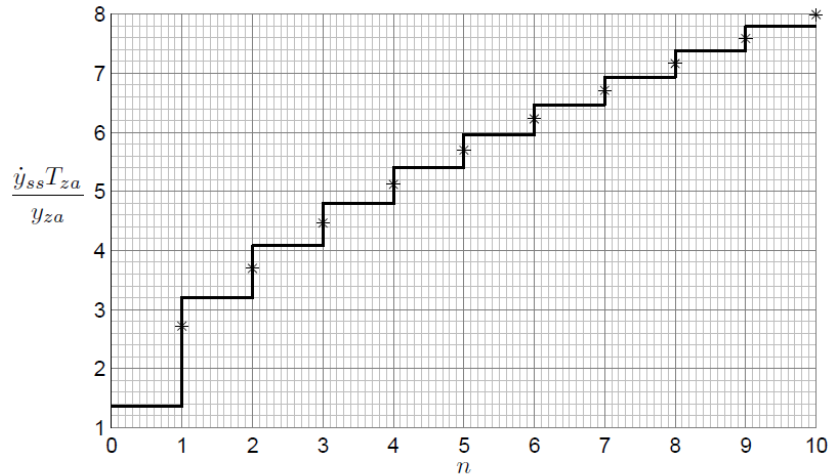
$$G(s) = \frac{K_I}{s(Ts + 1)^n}$$

- Za določitev potrebujemo:

- T_{za}
- $y_{za} = y(t)|_{t=T_{za}}$
- $K_I = \frac{\dot{y}_{ss}}{\Delta U}$

- Postopek:

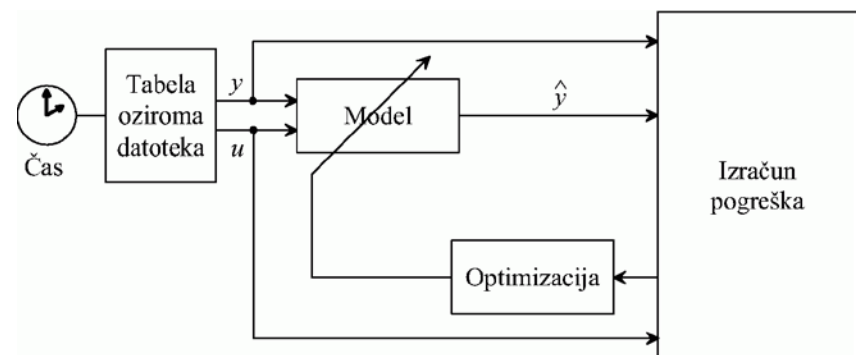
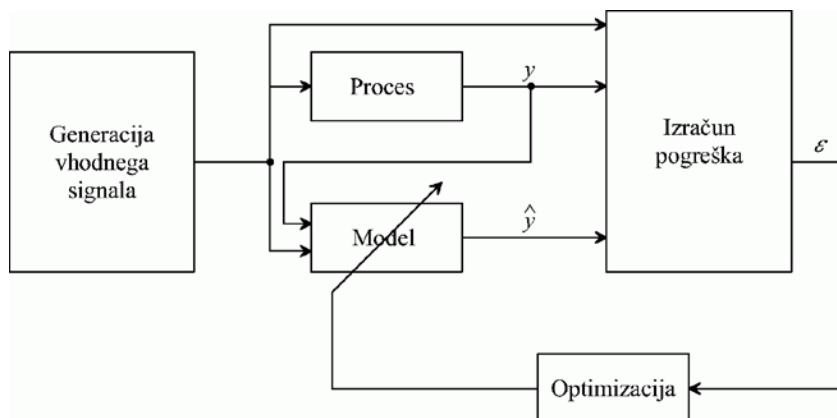
$$\frac{\dot{y}_{ss} T_{za}}{y_{za}} \rightarrow n \rightarrow \frac{T_{za}}{T} \rightarrow T$$



Metoda s prilagajanjem modela



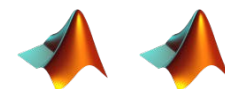
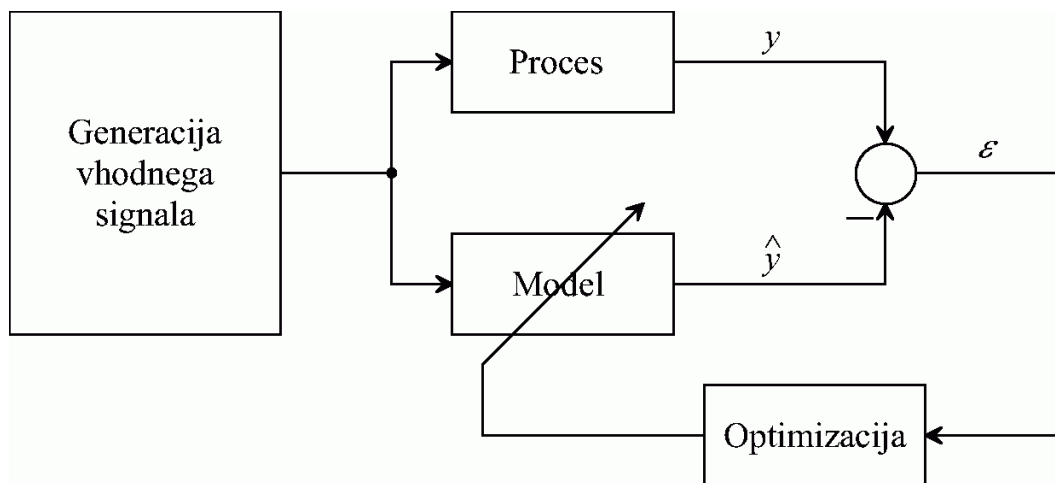
- Proces in model vzbujaemo z istim signalom, nato izračunamo pogrešek med izhodom procesa in izhodom modela
- Tri vrste pogreškov, zato ima metoda tri variante
- Skupno vsem je vzbujanje procesa in modela ter **optimizacija** modela, ki zajema tako optimizacijo strukture kakor tudi optimizacijo parametrov modela 
- Uporabna predvsem v nesprotni (off-line) izvedbi, ko je odziv procesa posnet na datoteki



Metoda z izhodnim pogreškom



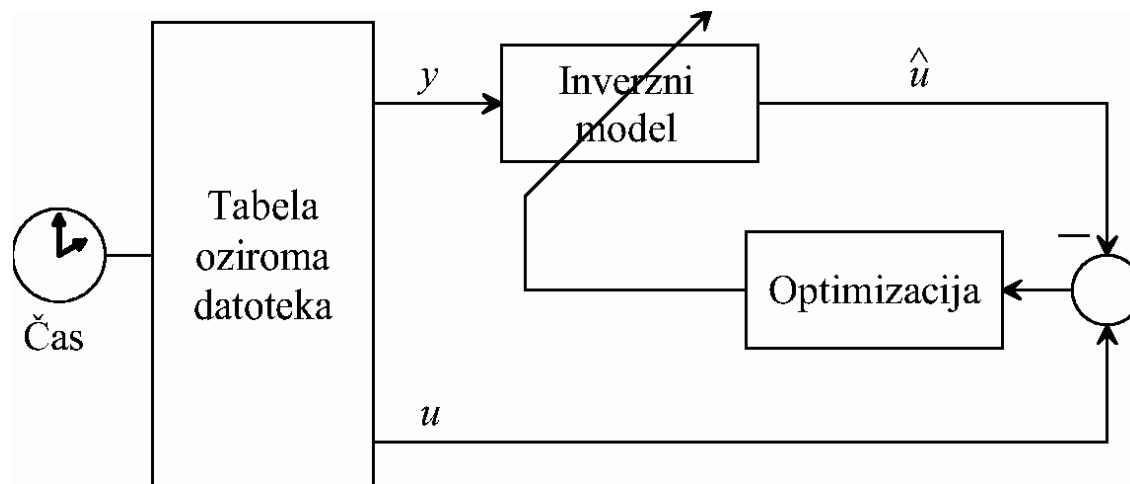
- Bistvo metode je v paralelni povezavi modela in procesa
- Model mora biti zato realiziran v rekurzivni obliki, ko ima prenosna funkcija števec in imenovalec, ki sta polinoma
- Metoda uporabna za zvezne in diskretne procese
- Prikazana izvedba, ki zajema signale iz procesa, možna pa je tudi izvedba, kjer so signali procesa posneti na datoteki



Metoda z vhodnim pogreškom



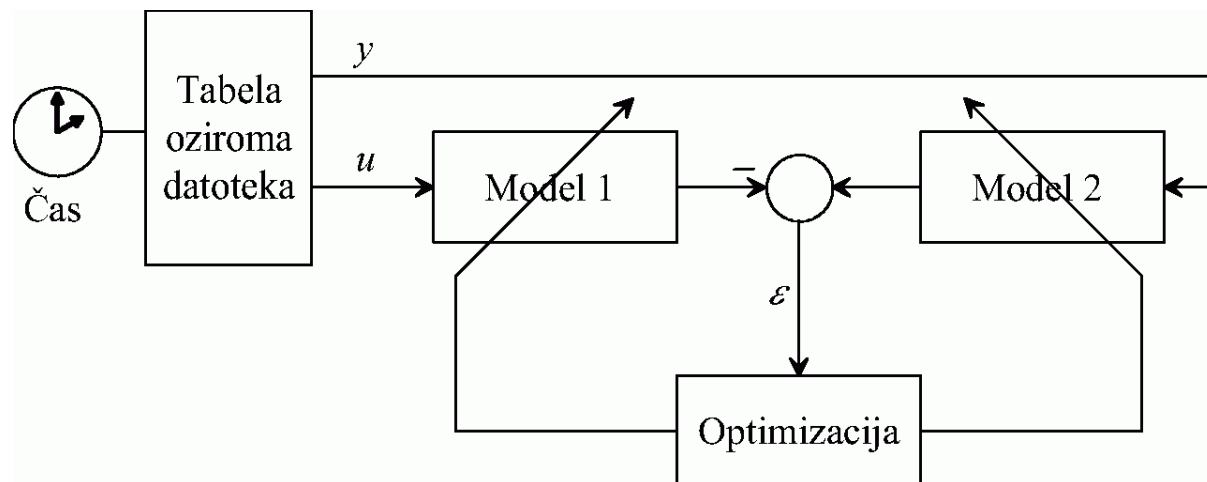
- Prikazana izvedba s signali, ki so posneti na datoteki
- Metoda le omejeno uporabna, saj bi inverz modela s časovno zakasnitvijo potreboval predikcijo; v praksi to rešujemo tako, da namesto signala y vodimo v inverzni model njegovo predikcijo (ki jo seveda imamo, saj so signali posneti na datoteki)



Metoda s posplošenim pogreškom



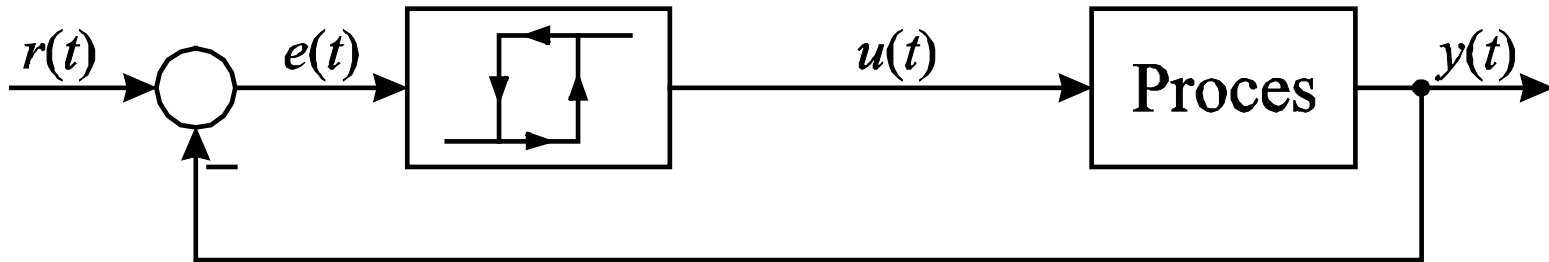
- Metoda s posplošenim pogreškom izmed vseh metod najbolj uporabna, predvsem pri diskretnih sistemih, ker daje linearni sistem enačb za ocenjevanje parametrov
- Model 1 predstavlja števec prenosne funkcije, model 2 imenovalec





- Obravnavali bomo le merjenje frekvenčnega odziva:
 - Na vhod procesa damo sinusni signal določene frekvence; ker (predpostavka) proces linearen, tudi izhod procesa (po izteku prehodnih pojavov) sinusne oblike (drugačna amplituda in faza)
 - Razmerje amplitud signalov imenujemo amplitudni odziv, razliko faz obeh signalov pa fazni odziv
 - Obe veličini skupaj (kompleksno število) tvorita frekvenčni odziv
 - Postopek ponovimo pri vseh frekvencah; v praksi pri zanimivih
 - Merjenje razmerja amplitud in razlike faz na različne načine:
 - Včasih v ta namen uporabljali kompenzacijski princip
 - Danes predvsem korelacijske metode – prednost izločanje šuma
 - Dve možnosti vzbujanja procesa:
 - Generiramo sinusni vhodni signal (računalnik, signalni generator); prednost metode znana frekvenca, slabost proces v odprti zanki
 - **Relejska povratna zanka**

Relejska povratna zanka



- V večini primerov sistem zaniha v okolici delovne točke; prednost – sistem s povratno zanko ostane v bližini DT, slabost – frekvenco potrebno izmeriti (problem, ker perioda običajno ni mnogokratnik časa vzorčenja)
- Signal $u(t)$ v primeru nihanja niz pravokotnih pulzov, $y(t)$ pa blizu sinusnemu signalu; oba približno nasprotne faze
- Åstrom in Hägglund uporabila za nastavljanje PID
- Analiza sistema je možna s pomočjo opisnih funkcij, tu bomo zadevo obravnavali le simulacijsko





- V praksi so merjeni podatki vedno moteni s šumom, zato moramo uporabljati teorijo stohastičnih signalov
- Če so (stohastični) signali med seboj odvisni, lahko povezavo med njimi ovrednotimo s pomočjo korelacijskih funkcij
- Avtokorelacijska funkcija podaja odvisnost med trenutno in za argument τ premaknjeno vrednostjo signala:

$$\phi_{uu}(\tau) = E\{u(t)u(t + \tau)\}$$

- Zaradi ergodičnosti lahko operator matematičnega upanja $E\{\cdot\}$ zamenjamo s povprečenjem po času:

$$\phi_{uu}(\tau) = E\{u(t)u(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)u(t + \tau) dt$$



- Podobno velja za križnokorelacijske funkcije:

$$\phi_{uy}(\tau) = E\{u(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t)y(t + \tau)dt$$

- Če opazujemo signale le od časa nič naprej:

$$\phi_{uu}(\tau) = E\{u(t)u(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t)u(t + \tau)dt$$

$$\phi_{uy}(\tau) = E\{u(t)y(t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(t)y(t + \tau)dt$$



- Povezava med avto- in križnokorelacijskimi funkcijami:

$$\begin{aligned}\phi_{uy}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) \int_0^\infty g(\sigma) u(t + \tau - \sigma) d\sigma dt \\ &= \int_0^\infty g(\sigma) \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) u(t + \tau - \sigma) dt d\sigma \\ &= \int_0^\infty g(\sigma) \phi_{uu}(\tau - \sigma) d\sigma = \int_0^\infty g(t) \phi_{uu}(\tau - t) dt\end{aligned}$$

- Križnokorelacijska funkcija med vhodnim in izhodnim signalom je torej konvolucija impulznega odziva in avtokorelacijske funkcije vhodnega signala → impulzni odziv lahko tolmačimo tudi kot dekonvolucijo križnokorelacijske funkcije $\phi_{uy}(\tau)$ in avtokorelacijske funkcije $\phi_{uu}(\tau)$



- Podobno velja za relacijo med $\phi_{yu}(\tau)$ in $\phi_{yy}(\tau)$:

$$\begin{aligned}\phi_{yy}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} y(t) \int_0^\infty g(\sigma) u(t + \tau - \sigma) d\sigma dt \\ &= \int_0^\infty g(t) \phi_{yu}(\tau - t) dt\end{aligned}$$

- Pozor: $\phi_{yu}(\tau)$ in $\phi_{uy}(\tau)$ sta dve različni križnokorelacijski funkciji, a je povezava enostavna $\rightarrow \phi_{yu}(\tau) = -\phi_{uy}(-\tau)$
- Fourierovi transformi korelacijskih funkcij so funkcije spektralne močnostne gostote:

$$\Phi_{uu}(\omega) = \mathcal{F}\{\phi_{uu}(\tau)\}$$

$$\Phi_{uy}(\omega) = \mathcal{F}\{\phi_{uy}(\tau)\}$$

$$\Phi_{yy}(\omega) = \mathcal{F}\{\phi_{yy}(\tau)\}$$



- Povezava med spektralnimi močnostimi gostotami:

$$\begin{aligned}\Phi_{uy}(\omega) &= \int_{-\infty}^{\infty} \phi_{uy}(\tau) e^{-j\omega\tau} d\tau \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} g(t) \phi_{uu}(\tau - t) dt e^{-j\omega\tau} d\tau \\ &= \int_0^{\infty} g(t) e^{-j\omega t} \int_{-\infty}^{\infty} \phi_{uu}(\tau - t) e^{-j\omega(\tau - t)} d\tau dt = G(j\omega) \Phi_{uu}(\omega)\end{aligned}$$

- Podobno dobimo še:

$$\Phi_{yy}(\omega) = G(j\omega) \Phi_{yu}(\omega)$$

- Povezava med obema enačbama:

$$\begin{aligned}\Phi_{yy}(\omega) &= G(j\omega) \Phi_{yu}(\omega) = G(j\omega) \Phi_{uy}(-\omega) = \\ &= G(j\omega) G(-j\omega) \Phi_{uu}(-\omega) = |G(j\omega)|^2 \Phi_{uu}(\omega)\end{aligned}$$

- upoštevali smo $\Phi_{yu}(\omega) = \Phi_{uy}(-\omega)$ in $\Phi_{uu}(-\omega) = \Phi_{uu}(\omega)$



- Relacija med Fourierovo transformacijo signala in njegovo spektralno močnostno gostoto:

$$\begin{aligned}\Phi_{uy}(\omega) &= \int_{-\infty}^{\infty} \phi_{uy}(\tau) e^{-j\omega\tau} d\tau = \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) y(t + \tau) dt e^{-j\omega\tau} d\tau \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) e^{j\omega t} \int_{-\infty}^{\infty} y(t + \tau) e^{-j\omega(t+\tau)} d\tau dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u(t) e^{j\omega t} Y(\omega) dt = \lim_{T \rightarrow \infty} \frac{1}{T} U(-\omega) Y(\omega)\end{aligned}$$

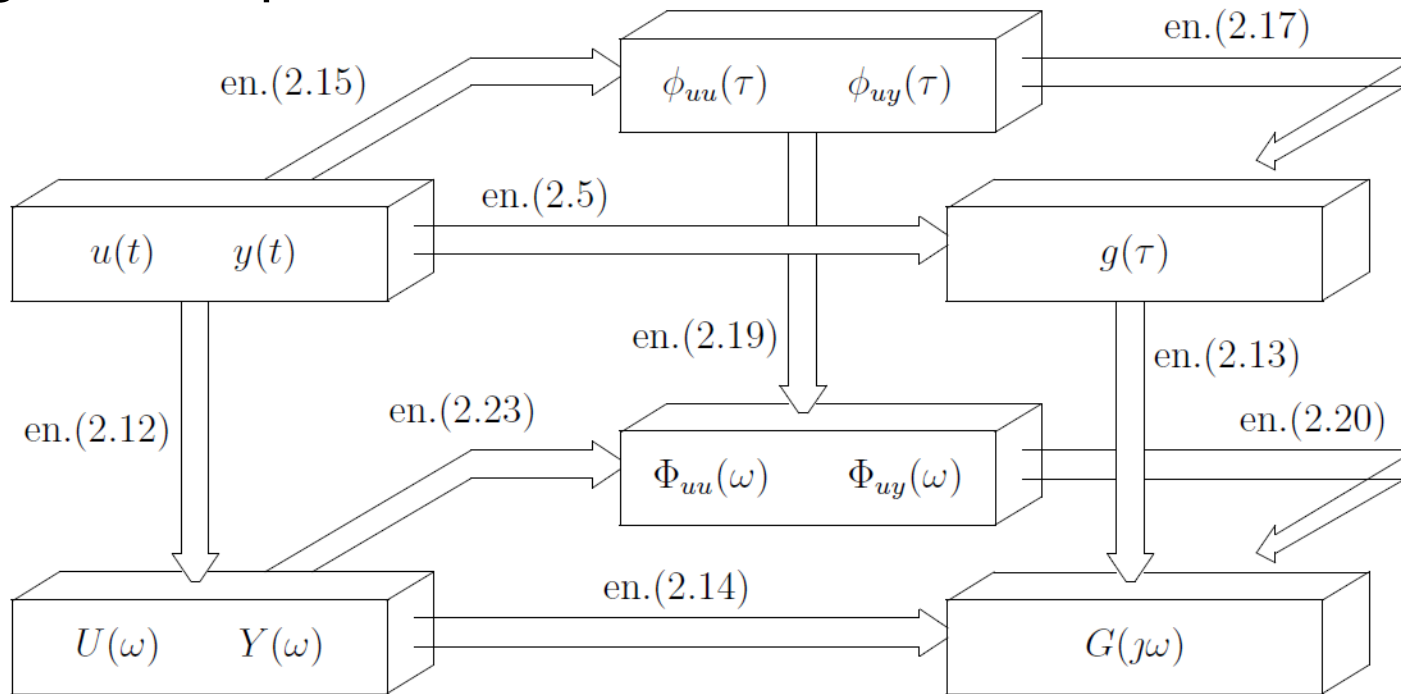
- Posebej zanimiva je zgornja relacija, če gre za isti signal:

$$\Phi_{uu}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} U(-\omega) U(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} |U(\omega)|^2$$

Korelacijske funkcije in spektri



- Odnosi med vhodno-izhodnimi signali in neparametričnimi modeli v grafičnem prikazu



- Sklici na enačbe se nanašajo na učbenik D. Matko: Identifikacije (Založba FE, 1999)



- Linearne časovno nespremenljive procese s koncentriranimi parametri opišemo z navadnimi linearnimi diferenčnimi enačbami s konstantnimi koeficienti v naslednji obliki

$$\begin{aligned} y(k) + a_1 y(k-1) + \dots + a_{n-1} y(k-n+1) + a_n y(k-n) = \\ = b_0 u(k) + b_1 u(k-1) + \dots + b_{n-1} u(k-n+1) + b_n u(k-n) \end{aligned}$$

- $u(k)$ vhod v proces, $y(k)$ njegov izhod, b_i in a_i so parametri procesa
- S pomočjo z-transformacije dobimo pri pogoju, da proces ne vsebuje začetne energije (vsi začetni pogoji v zgornji enačbi so enaki nič), prenosno funkcijo procesa:

$$G(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_{n-1} z^{-n+1} + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_{n-1} z^{-n+1} + a_n z^{-n}}$$

- Z-transformacija izhoda $Y(z)$ je produkt pren. f. $G(z)$ in $U(z)$
- Produktu v frekvenčnem prostoru ustreza konvolucija v časovnem:

$$y(k) = \sum_{\tau=0}^{\infty} g(\tau) u(k-\tau)$$



- Parametrični diskretni model je model diskretnega procesa izražen s parametri, npr. parametri b_i in a_i prenosne funkcije
- Glede na konvolucijsko enačbo je možen neparametrični model impulzni odziv ali utežna funkcija $g(k) \rightarrow$ to je odziv sistema na enotni impulz $\delta(k) = \begin{cases} 1, k = 0 \\ 0, k \neq 0 \end{cases} \rightarrow \delta(k)$ je drugačne oblike kot $\delta(t)$
- Drugi neparametrični model je frekvenčni odziv $G(e^{j\omega T})$:
 - Interpretacija s harmoničnim vzbujanjem: Če vzbujamo stabilni linearni časovno nespremenljivi sistem s sinusom:
$$u(k) = U_0 \sin(\omega_0 kT + \varphi_u)$$
 - se sistem odzove s harmoničnim nihanjem enake frekvence:
$$y(k) = Y_0 \sin(\omega_0 kT + \varphi_y)$$
 - Razmerje amplitud je določeno z amplitudnim odzivom $\frac{Y_0}{U_0} = |G(e^{j\omega_0 T})|$
 - Fazna razlika je določena s faznim odzivom $\varphi_y - \varphi_u = \angle[G(e^{j\omega_0 T})]$



- Povezava med impulznim odzivom $g(k)$ in prenosno funkcijo $G(z)$:
$$G(z) = \mathcal{Z}\{g(k)\} \quad g(k) = \mathcal{Z}^{-1}\{G(z)\}$$
- Povezava med frekvenčnim odzivom $G(j\omega)$ in pren. funkcijo $G(s)$:
$$G(e^{j\omega T}) = \lim_{z \rightarrow e^{j\omega T}} G(z)$$
- Povezava med impulznim odzivom $g(k)$ in frekv. odzivom $G(e^{j\omega T})$:
 - Običajno je smiselna le uporaba diskretne Fourierove transformacije oz. njene hitre različice, zaradi česar postane spekter signala diskreten, kar bomo obravnavali v nadaljevanju
- Ugotovitve glede povezav med korelacijskimi funkcijami in spektri veljajo tudi za diskretne signale:
 - $\phi_{uy}(\tau) = \sum_{k=0}^{\infty} g(k)\phi_{uu}(\tau - k)$
 - $\Phi_{uy}(\omega) = G(e^{j\omega T})\Phi_{uu}(\omega)$
 - $\Phi_{yy}(\omega) = |G(e^{j\omega T})|^2\Phi_{uu}(\omega)$

Povezava med zveznimi in diskretnimi signali



- Signali v praksi so običajno zvezni in neperiodični
- Za obravnavo na digitalnih računalnikih pa uporabljamo predvsem Diskretno Fourierovo transformacijo, ki predpostavlja časovno diskretne in periodične signale
- Označevanje:
 - Časovno zvezni signali z majhno črko, npr. $x(t)$
 - Njihov Fourierov transform z veliko črko, npr. $X(\omega)$
 - Vzorčeni signali z zvezdico, npr. $x^*(t)$, $X^*(\omega)$
 - Periodični signali s p , npr. časovni $x^p(t)$, Fourierov transform oz. koeficient Fourierove vrste $X^p(m\frac{2\pi}{T})$
 - Časovno diskretni in periodični signali z oznako d , npr. $x^d(k)$ za časovni signal, $X^d(m)$ za njegov Diskretni Fourierov transform
 - Čas vzorčenja pri vzorčenih signalih bo T , perioda pa t_p
 - Frekvenčni prirastek bo F , perioda (frekvenca vzorčenja) pa f_s



- Predpostavimo, da je originalni časovno zvezni in neperiodični signal različen od 0 le na intervalu med 0 in t_p , izven tega območja pa je enak 0 – v praksi je to običajno območje, v katerem smo merili signal in je vedno končne dolžine
- Poleg tega bomo predpostavili, da je vrednost signala $x(t)$ pri zgornji meji enaka 0, torej:

$$x(t) = 0 \quad t < 0 \text{ ali } t \geq t_p$$

- Vzorčeni signal dobimo tako, da pomnožimo originalni signal $x(t)$ z vlakom delta impulzov:

$$x^*(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT)$$

- Ker je vrednost signala $x(t)$ različna od 0 le v območju med 0 in $t_p = NT$, meji v vsoti zgornjega izraza zamenjamo z 0 oz. $N - 1$

$$x^*(t) = x(t) \sum_{k=0}^{N-1} \delta(t - kT) = \sum_{k=0}^{N-1} x(kT) \delta(t - kT)$$



- Pri obdelavi na digitalnih računalnikih moramo uporabljati vzorčene signale – tako časovne kot tudi frekvenčne
- Vzorčenost v enem prostoru ima za posledico periodičnost v drugem, zato pravzaprav potrebujemo periodične signale
- Predpostavili bomo, da dobimo periodični signal iz originalnega signala tako, da le tega enostavno prestavimo za $\pm t_p, \pm 2t_p, \dots$

$$x^p(t \pm lt_p) = x(t) \quad l = 0, 1, 2, \dots$$

- Digitalni signal $x^d(k)$, ki pripada originalnemu signalu $x(t)$, je definiran le za celoštevilčne vrednosti argumenta:

$$x^d(k) = x(t) \Big|_{t=kT}$$

- Uporabljali bomo le periodične digitalne signale, zato:

$$x^d(k) = \begin{cases} x^p(kT) = x(kT) & k = 0, 1, 2, \dots, N-1 \\ x^d(k \pm lN), l \in \mathbb{N} & \text{sicer} \end{cases}$$



- Fourierova transformacija zveznega signala je

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt = \int_0^{t_p} x(t)e^{-j\omega t} dt$$

- Frekvenčni spekter takšnega signala je zvezen in neperiodičen
- Če predpostavimo periodični signal, moramo namesto Fourierove transformacije uporabiti Fourierovo vrsto, frekvenčni spekter takšnega signala pa je diskreten, to se pravi, da vsebuje le enosmerno komponento (frekvenco 0), osnovno harmonsko komponento ter njene mnogokratnike (frekvence $0, \pm\frac{1}{t_p}, \pm\frac{2}{t_p}, \dots$)

- Koeficient Fourierove vrste je

$$X^p(m\frac{2\pi}{t_p}) = \frac{1}{t_p} \int_0^{t_p} x^p(t)e^{-jm\frac{2\pi}{t_p}t} dt = \frac{1}{t_p} \int_0^{t_p} x(t)e^{-jm\frac{2\pi}{t_p}t} dt$$

- Primerjava: $X(\omega)|_{\omega=m\frac{2\pi}{t_p}} = t_p X^p(m\frac{2\pi}{t_p})$



- Fourierovo transformacijo digitalnih (vzorčenih in periodičnih) signalov računamo na digitalnih računalnikih z diskretno Fourierovo transformacijo oziroma z njeno hitro različico FFT:

$$X^d(m) = \sum_{k=0}^{N-1} x^d(k) e^{-j\frac{2\pi}{N}mk}$$
$$x^d(k) = \frac{1}{N} \sum_{m=0}^{N-1} X^d(m) e^{j\frac{2\pi}{N}mk}$$

- Računanje členov Fourierove vrste ali vrednosti Fourierove transformacije pri večkratnikih frekvence $\frac{2\pi}{t_p}$ iz DFT:

$$X^p\left(m\frac{2\pi}{t_p}\right) = \frac{1}{N}X^d(m) \quad m < \frac{N}{2}$$

$$X(\omega)|_{\omega=m\frac{2\pi}{t_p}} = t_p X^p\left(m\frac{2\pi}{t_p}\right) = TX^d(m) \quad \omega < \frac{\pi}{T}$$



- Diskretna križnokorelacijska funkcija je definirana kot:

$$\phi_{xy}^d(\tau_i) = \frac{1}{N} \sum_{k=0}^{N-1} x^d(k) y^d(k + \tau_i)$$

- τ_i je celo število in N perioda funkcije
- Križna korelacijska funkcija vzorčenih periodičnih signalov:

$$\phi_{xy}^{*p}(\tau_i T) = \begin{cases} \frac{1}{T} \phi_{xy}^d(\tau_i) & \tau_i = 0, \pm 1, \pm 2, \dots \\ 0 & \tau_i \text{ ni celo število} \end{cases}$$

- Relacija med križnokorelacijskima funkcijama periodičnega in aperiodičnega signala:

$$\phi_{xy}^p(\tau) = \frac{1}{t_p} \phi_{xy}(\tau)$$



- Močnostni spekter:

$$\Phi_{xy}^p \left(m \frac{2\pi}{t_p} \right) = \frac{1}{N} \Phi_{xy}^d(m) \quad m < \frac{N}{2}$$

- Spektralna močnostna gostota:

$$\begin{aligned} \Phi_{xy}(\omega) \Big|_{\omega = m \frac{2\pi}{t_p}} &= t_p \Phi_{xy}^p \left(m \frac{2\pi}{t_p} \right) & \omega < \frac{\pi}{T} \\ &= T \Phi_{xy}^d(m) \end{aligned}$$

- Spektralno močnostno gostoto dobimo, če močnostni spekter porazdelimo po celotnem območju med dvema vrednostma vzorčenega močnostnega spektra, torej če ga delimo s širino prirastka $(\frac{1}{t_p})$, oziroma pomnožimo s t_p



- Zelo pomemben dejavnik pri identifikaciji procesov so motnje
- Pod tem pojmom razumemo vso dinamiko procesa, ki je matematični model ne upošteva, vse nemerljive vplive in pogreške meritev
- Če predpostavimo linearnost procesov, lahko vse motnje, razen merilnega pogreška vhodnega signala združimo v en sam člen $n(t)$, ki ga prištejemo nemotenemu izhodu procesa y_0 :

$$y(t) = y_0(t) + n(t)$$

- Motnje po definiciji nepredvidljive, v praksi nekaj tipičnih vrst:
 - visokofrekvenčni kvazistacionarni stohastični signali
 - nizkofrekvenčne nestacionarne motnje kot n.pr. signali lezenja (s tujko drift) in stopnice (nastane tudi zaradi pristranske ocene DT)
 - sinusne motnje
 - neznane motnje kot n.pr. špice (s tujko spikes) itd.



- Pri identifikaciji nastopajo šumi v dvojni vlogi:
 - motilni signali, ki vključujejo vse nemerljive vplive in pogreške meritev
 - vzbujevalni signali, ki morajo biti frekvenčno bogati
- Za opis signalov sta pomembni dve lastnosti:
 - Distribucija oz. porazdelitev signala
 - Opisuje verjetnost, da amplituda signala zavzame vrednost v določenem območju
 - Frekvenčni spekter signala
 - Podaja medsebojno odvisnost vrednosti signala v posameznih časovnih trenutkih – ta odvisnost je v tesni zvezi z vsebovanostjo posameznih frekvenc v signalu (odtod tudi izraz spekter)
 - Pri *belem šumu* so enakomerno zastopane vse frekvence (tako kot pri beli svetlobi), vsebuje torej tudi neskončno frekvenco, zato so poljubno gosto si sledeče vrednosti signala medsebojno statistično neodvisne



- Definirana s funkcijo gostote verjetnosti (angl. *probability density function, PDF*) $p(x)$, za katero velja $\int_{-\infty}^{\infty} p(x)dx = 1$. Pogoste PDF:
 - Normalna ali Gaußova porazdelitev:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$$

- $\bar{x}$ je srednja vrednost signala in σ njegova standardna deviacija
- vsota dveh signalov z Gaußovo porazdelitvijo ima Gaußovo porazdelitev

- Enakomerna porazdelitev:

$$p(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{drugje} \end{cases}$$

- funkcija gostote verjetnosti v določenem območju konstanta

- Binarna porazdelitev:

$$p(x) = \frac{1}{2}[\delta(x) + \delta(x-1)]$$

- diskretna porazdelitev, zato funkcija gostote verjetnosti iz δ -impulzovov





- Spektralna močnostna gostota belega šuma je po definiciji konstantna:

$$\Phi_{vv}(\omega) = \Phi_0$$

- Beli šum $v(t)$ je signal s statistično neodvisnimi vrednostmi
- Njegova avtokorelacijska funkcija je enotin impulz
- Običajno privzamemo, da je srednja vrednost šuma enaka nič, zato lahko zapišemo njegovo avtokorelacijsko funkcijo v naslednji obliki:

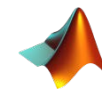
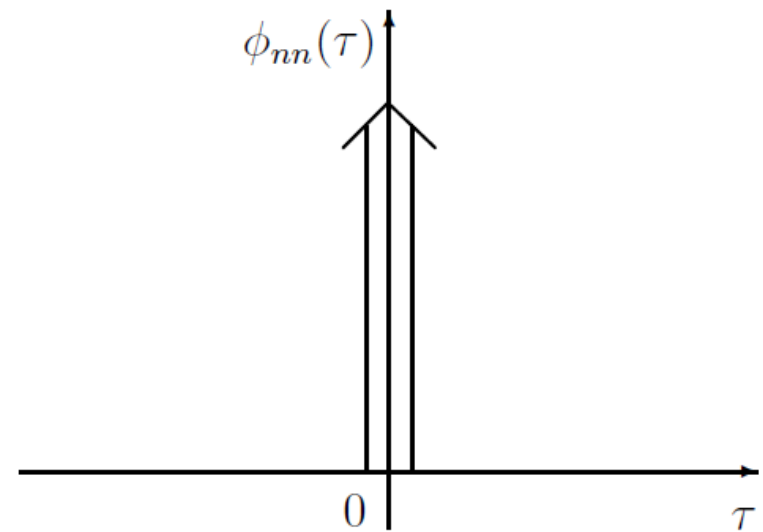
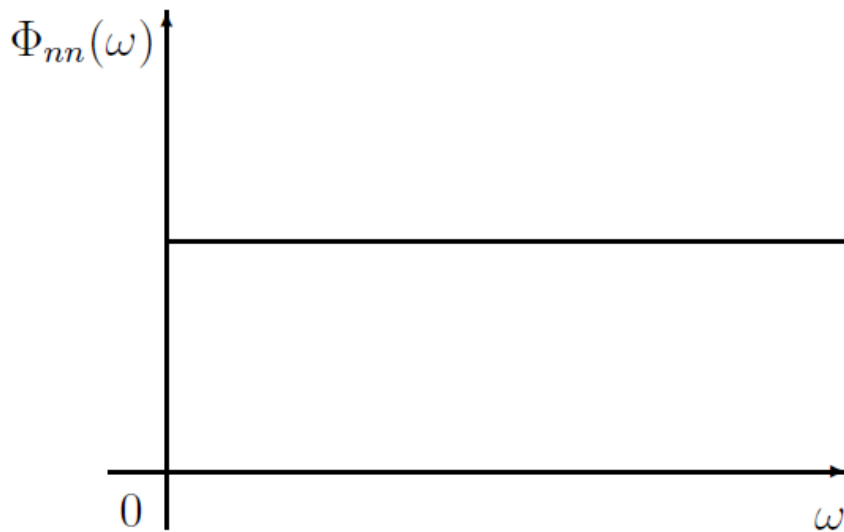
$$\phi_{vv}(\tau) = \Phi_0 \delta(\tau)$$

- Φ_0 je spektralna močnostna gostota belega šuma
- Beli šum bi bil zelo primeren signal za vzbujanje procesov, toda

$$\lim_{t_p \rightarrow \infty} \frac{1}{t_p} \int_{-\frac{t_p}{2}}^{\frac{t_p}{2}} v^2(t) dt = \phi_{vv}(0) = \infty \quad \rightarrow \textbf{neskončna moč}$$



- Spektralna močnostna gostota in avtokorelacija belega šuma

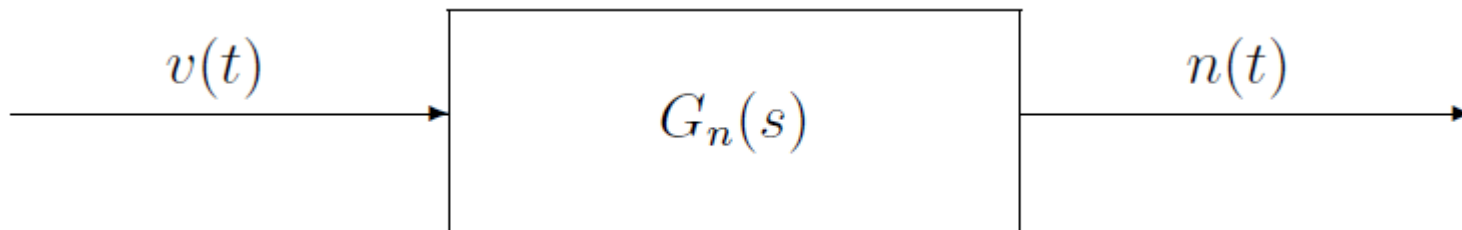




- Barvni šum ne vsebuje vseh frekvenc, njegov frekvenčni spekter (spektralna močnostna gostota) ni konstanten
- Frekvenčni spekter signalov lahko preoblikujemo s filtri:

$$\Phi_{yy}(\omega) = |G(j\omega)|^2 \Phi_{uu}(\omega)$$

- Spektralno močnostno gostoto imenujemo *racionalna*, če jo je možno dobiti iz belega šuma s filtriranjem na filtru z racionalno prenosno funkcijo, ki ga imenujemo šumni filter



- Generacija barvnega šuma na sliki poenostavi obravnavo
- To ni način njegovega generiranja!



- Spektralna močnostna gostota barvnega šuma na izhodu šumnega filtra, če imamo na vhodu filtra beli šum:
$$\Phi_{nn}(\omega) = |G_n(j\omega)|^2 \Phi_0$$
- Če je filter $G_n(s)$ stabilen, je barvni šum *stacionaren*
- Iz belega šuma je možno preko ustreznega stabilnega in fazno minimalnega filtra generirati (seveda le matematično) šum s poljubno racionalno spektralno močnostno gostoto
- To dejstvo je posledica kvadrata absolutne vrednosti frekvenčnega odziva v zgornji enačbi:
 - Lega pola oz. ničle simetrično glede na imaginarno os (pri zveznih sistemih) ali krog enote (pri diskretnih sistemih) namreč ne vpliva na izraz $|G(j\omega)|^2$, prav isto vrednost dobimo namreč, če pol oz. ničlo zrcalimo glede na imaginarno os ali krog enote



- *Širokopasovni šum* je poseben primer barvnega šuma, kjer je filter nizkoprepustni filter prvega reda z visoko mejno frekvenco

$$G_n(s) = \frac{1}{1 + sT_g}$$

- $T_g = \frac{1}{\omega_g}$ je časovna konstanta nizkopasovnega filtra
- Spektralna močnostna gostota širopasovnega šuma:

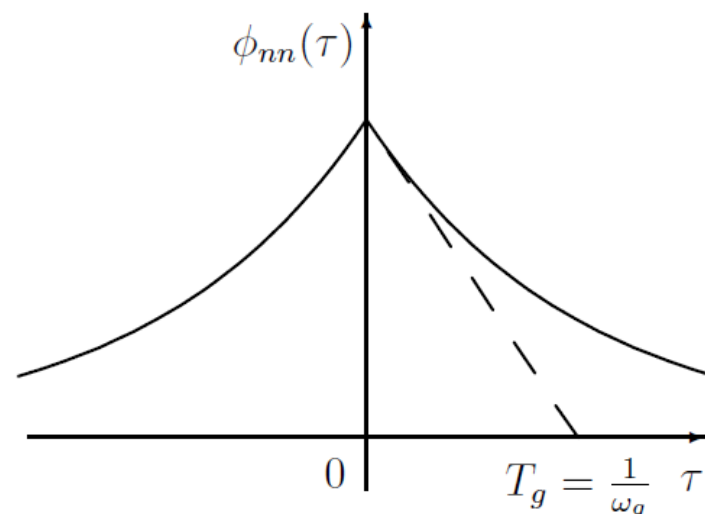
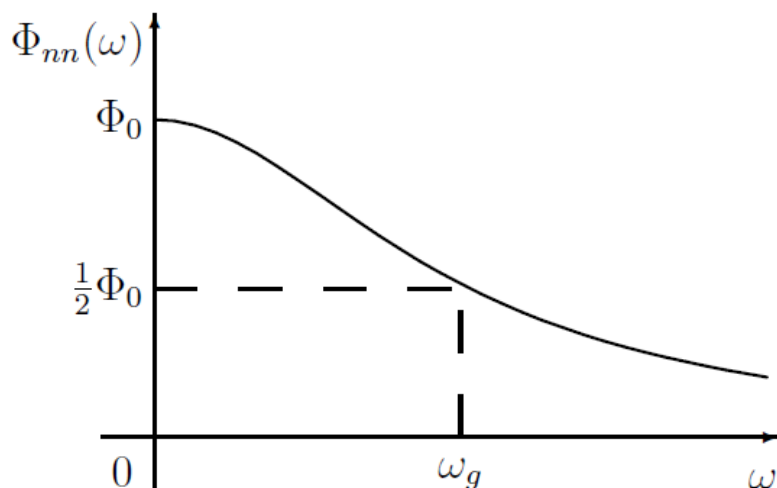
$$\Phi_{nn}(\omega) = |G_n(j\omega)|^2 \Phi_0 = \frac{\Phi_0}{1 + T_g^2 \omega^2} = \frac{\Phi_0}{1 + \left(\frac{\omega}{\omega_g}\right)^2}$$

- Avtokorelacija širopasovnega šuma:

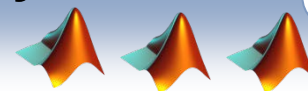
$$\phi_{nn}(\tau) = \frac{\Phi_0 e^{-\left|\frac{\tau}{T_g}\right|}}{2T_g} = \frac{\Phi_0 e^{-|\tau|\omega_g}}{2} \omega_g$$



- Spektralna močnostna gostota in avtokorelacija širokopasovnega šuma:



- Srednja moč signala (varianca) je enaka $\phi_{nn}(0) = \frac{\Phi_0}{2T_g} = \frac{\Phi_0 \omega_g}{2}$
 - Pri majhnih T_g (velikih ω_g) postane širokopasovni šum zelo podoben belemu, saj postane avtokorel. funkcija skoraj δ -impulz (ploščina pod avtokorel. f. je Φ_0), njegova spektralna močnostna gostota pa je skoraj konstantna v zelo širokem frekvenčnem območju



Zvezni naključni binarni signal



- *Zvezni naključni binarni signal* (ZNBS) je pogosto označen z angleško kratico RBS (*Random Binary Signal*)
- Ima dve stanji signala: $+a$ in $-a$
- Sprememba iz enega v drugo stanje je časovno naključna
- Imenujemo ga tudi „naključni telegrafski signal“
- V primerjavi z naključnim preskusnim signalom z zvezno amplitudno porazdelitvijo ima binarni šum naslednje prednosti:
 - RBS ima pri omejeni amplitudi največjo močnostno gostoto (dobro vzbujanje je zelo pomembno za kakovost modela – „dobro“ pomeni med drugim tudi velika moč vzbujanja)
 - Enostavno vzbujanje s krmiljenjem releja (pomembno pri omejitvah s strojno opremo)
 - Tvorjenje križnokorelacijske funkcije na enostaven način, in sicer z množenjem izhodnega signala s $+a$ ali $-a$

Avtokorelacijska funkcija ZNBS



- Predpostavimo, da je povprečje menjav predznaka v časovnem intervalu Δt podano s Poissonovo porazdelitvijo:

$$P(n) = \frac{(\mu\Delta t)^n}{n!} e^{-\mu\Delta t} \quad (\mu \text{ je povprečno število menjav predznaka v časovni enoti})$$

- To pomeni, da je verjetnost menjav predznaka signala naslednja:
 - 0 menjav $\rightarrow P(0) = e^{-\mu\Delta t}$
 - 1 menjava $\rightarrow P(1) = \mu\Delta t e^{-\mu\Delta t}$
 - 2 menjavi $\rightarrow P(2) = \frac{(\mu\Delta t)^2}{2!} e^{-\mu\Delta t} \dots$ 
- Produkt $u(t)u(t + \tau)$ ZNBS ima v nekem trenutku vrednost $+a^2$ ali $-a^2$ (ali imata obe vrednosti enak ali nasproten predznak)
- $E\{u(t)u(t + \tau)\} = \begin{cases} +a^2 & \text{v času } \tau \text{ nastopilo sodo menjav predznaka} \\ -a^2 & \text{v času } \tau \text{ nastopilo liho menjav predznaka} \end{cases}$
- $E\{u(t)u(t + \tau)\} = +a^2$ pri $\tau = 0$

Avtokorelacijska funkcija ZNBS



- Ker nastopajo menjave predznaka naključno, lahko napišemo le matematično upanje za srednjo vrednost produkta pri premiku:

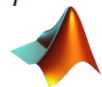
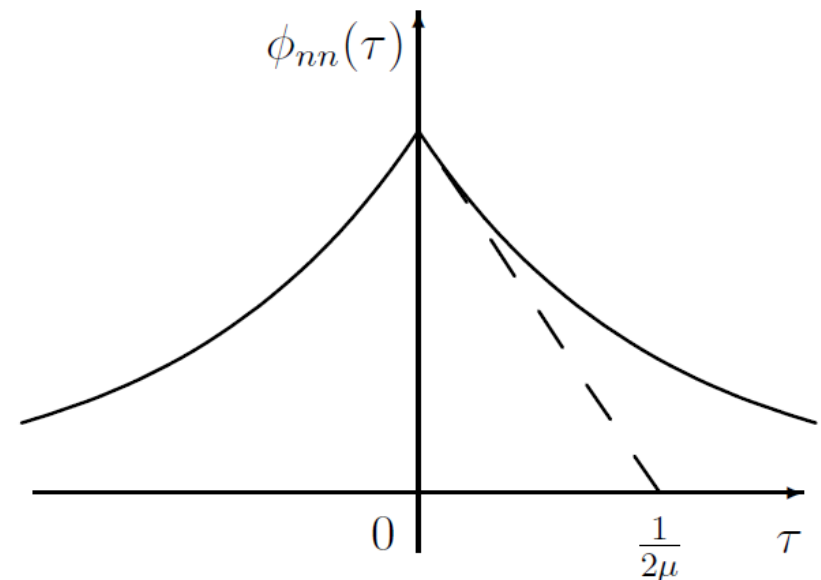
$$\begin{aligned} E\{u(t)u(t + \tau)\} &= +a^2[P(0) + P(2) + \dots] - a^2[P(1) + P(3) + \dots] \\ &= a^2 e^{-\mu|\tau|} \left[1 - \frac{\mu|\tau|}{1!} + \frac{(\mu|\tau|)^2}{2!} - \frac{(\mu|\tau|)^3}{3!} + \dots \right] = a^2 e^{-2\mu|\tau|} \end{aligned}$$

- Enak potek kot avtokorelacijska funkcija širopasovnega šuma, če:

$$- a^2 = \frac{\Phi_0}{2} \omega_g$$

$$- \mu = \frac{\omega_g}{2}$$

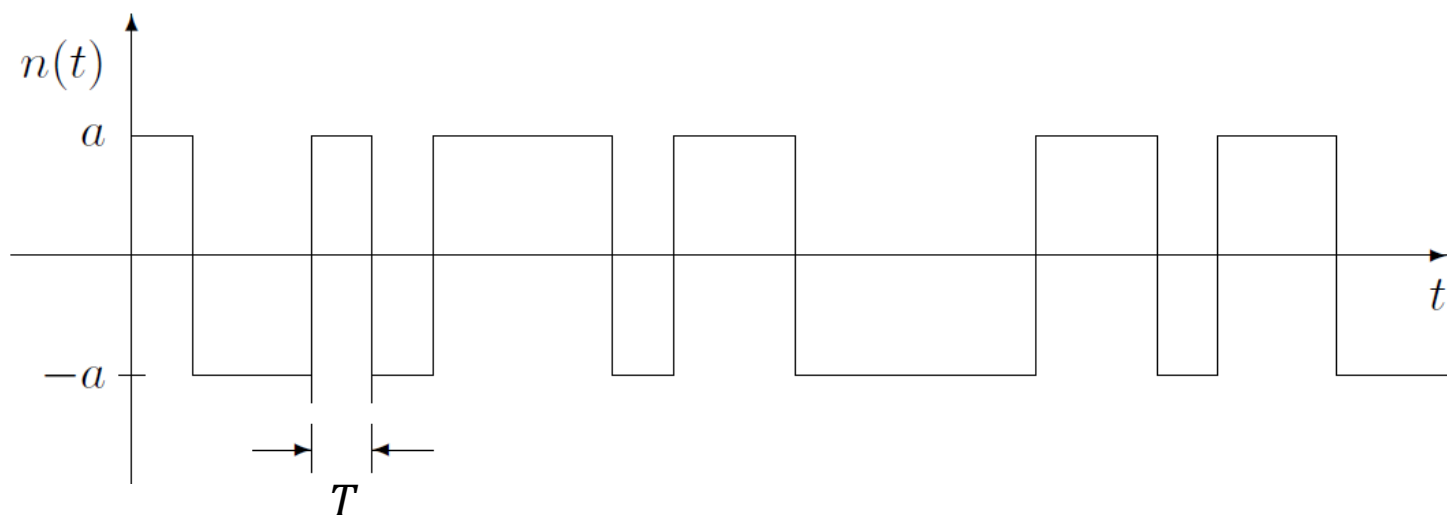
- Težava je, da so lahko časi med preklopi poljubno kratki (praktična realizacija ni mogoča!)



Diskretni naključni binarni signal (DNBS)



- Angleško ime: *Discrete Random Binary Signal (DRBS)*
- Rešuje problem (pre)kratkih časov med preklopi – le-ti nastopijo le v časovno diskretnih trenutkih $kT, k = 1, 2, 3 \dots$, kjer je T dolžina časovnega intervala, ki ga imenujemo tudi takt
- Tipičen potek:





- Avtokorelacijsko funkcijo DNBS dobimo z naslednjim premislekom:
 - za premik $\tau = 0$ dobimo samo pozitivne produkte, a^2 , zato je njihova povprečna vrednost, to je avtokorelacijska funkcija pri premiku 0:
 $\phi_{nn}(0) = a^2$
 - Pri majhnem časovnem premiku $|\tau| < T$ dobimo tudi negativne zmnožke, njihovo trajanje pa je sorazmerno premiku τ
 - Za $|\tau| \geq T$ nastopajo pozitivni in negativni zmnožki enako pogosto, tako da je $\phi_{nn}(\tau) = 0$
- Avtokorelacijska funkcija DNBS je torej enaka:

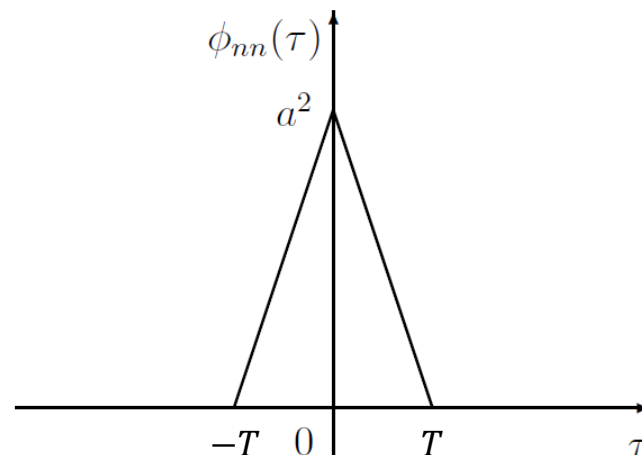
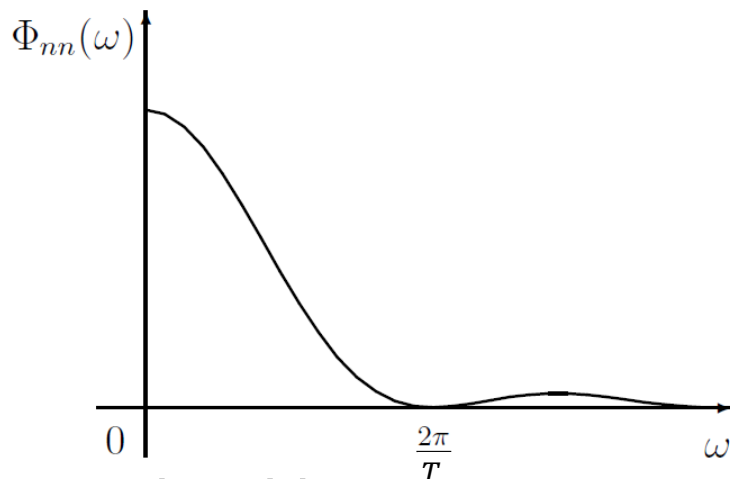
$$\phi_{nn}(\tau) = \begin{cases} a^2 \left[1 - \frac{|\tau|}{T} \right] & |\tau| < T \\ 0 & |\tau| \geq T \end{cases}$$

Diskretni naključni binarni signal (DNBS)

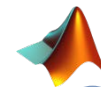


- Močnostna gostota DNBS sledi iz Fourierove transformacije avtokorelacijske funkcije:

$$\Phi_{nn}(\omega) = a^2 T \left(\frac{\sin \frac{\omega T}{2}}{\frac{\omega T}{2}} \right)^2$$



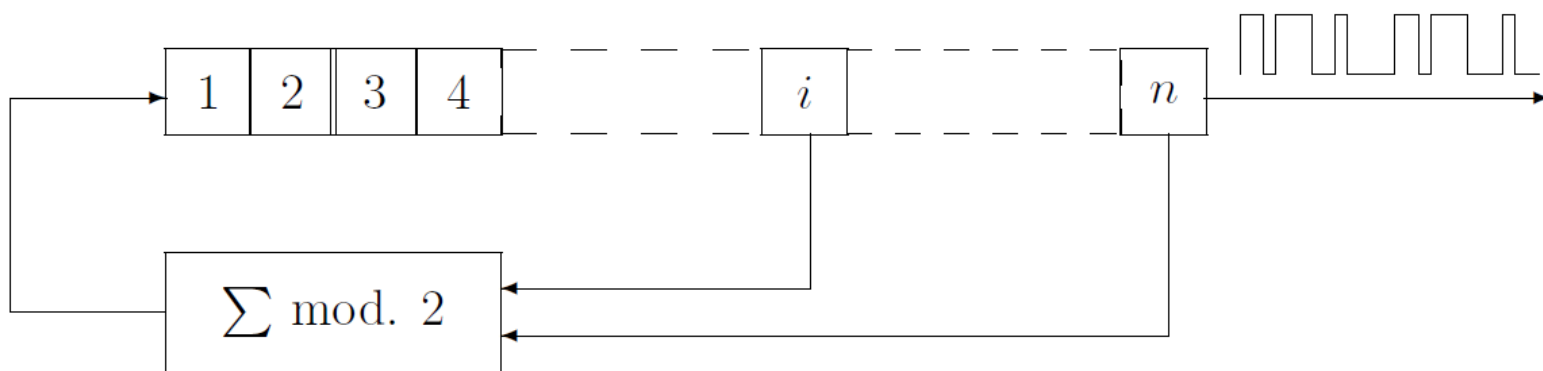
- DNBS bi lahko realizirali z neskončno dolgim premikalnim registrom s povratno vezavo (v praksi končno dolg $\rightarrow$ PNBS)



Psevdonaključni binarni signal (PNBS)



- Neskončno dolgega premikalnega registra, s katerim bi realizirali DNBS, seveda v praksi ni možno realizirati
- Če realiziramo signal s pomočjo premikalnega registra, ki ima n stopenj z binarno informacijsko vsebino (0 ali 1), dobimo *psevdonaključni binarni signal* (s tujko Pseudo-Random Binary Signal – PRBS)



- Vsebinsko posameznih stopenj premaknemo v periodičnih časovnih trenutkih (s periodo T) v naslednjo stopnjo

Psevdonaključni binarni signal (PNBS)



- Binarno funkcijo (ekvivalenco oziroma vsoto po modulu 2) zadnje in še ene vmesne stopnje vodimo nazaj v prvo stopnjo
- Pri vsakem premiku dobimo v zadnji stopnji najprej binarno zaporedje, ki je bilo shranjeno v celotnem registru (začetno stanje), nato pa še zaporedje, ki je generirano s pomočjo povratne zanke

Število stopenj n	Stopnje za povratno zanko	Trajanje periode N
2	1 in 2	3
3	1 in 3 ali 2 in 3	7
4	3 in 4 ali 1 in 4	15
5	2 in 5 ali 3 in 5	31
6	5 in 6	63
7	4 in 7	127
8	ne da polne periode	-
9	5 in 9	511
10	7 in 10	1023
11	10 in 11	2047

- Pri vsaki binarni funkciji je eno začetno stanje prepovedano (same ničle pri vsoti po mod 2)
- Pri pravilni izbiri stopenj za povr. vezavo nastane $2^n - 1$

kombinacij ničel in enic v registru, nato pa se dogajanje ponovi

Psevdonaključni binarni signal (PNBS)



- Za izhodno stopnjo lahko uporabimo katerokoli stopnjo:
 - Izhodni signal je $+a$, če je vrednost izhodne stopnje 1
 - Izhodni signal je $-a$, če je njena vrednost 0
- Nekaj lastnosti PRBS:
 - PNBS vsebuje $\frac{1}{2}(N + 1)$ vrednosti $+a$ in $\frac{1}{2}(N - 1)$ vrednosti $-a$, njegova srednja vrednost je zato enaka $E\{u(k)\} = \bar{u}(k) = \frac{a}{N}$
 - PNBS je sestavljen iz impulzov različnih dolžin; v vsaki periodi je:
 - $\frac{1}{2} \cdot \frac{N+1}{2}$ impulzov dolžine T (polovica pozitivnih in polovica negativnih)
 - $\frac{1}{4} \cdot \frac{N+1}{2}$ impulzov dolžine $2T$ (polovica pozitivnih in polovica negativnih)
 - $\frac{1}{8} \cdot \frac{N+1}{2}$ impulzov dolžine $3T$ (polovica pozitivnih in polovica negativnih)
 - ...
 - 1 negativni impulz dolžine $(n - 1)T$
 - 1 pozitivni impulz dolžine nT



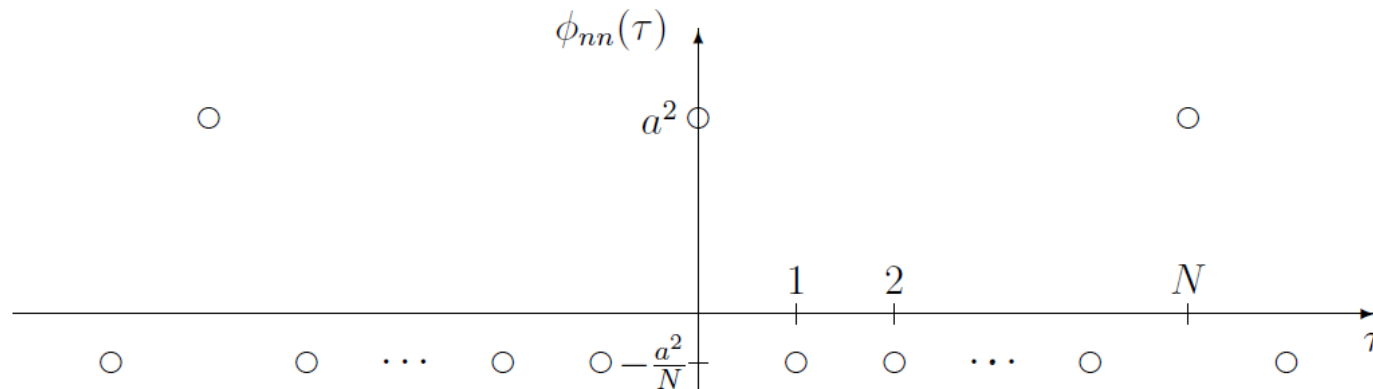
- Avtokorelacijska funkcija signala PNBS:
 - Najprej bomo izračunali avtokorelacijsko funkcijo diskretnega PNBS
 - Pri tem bomo diskretni časovni premik označili s τ_i , in sicer bo premik τ_i ustrezal dejanskemu časovnemu premiku $\tau_i T$
 - Srednja vrednost kvadrata signala (za $\tau_i = 0$) je enaka a^2
 - Ker je signal periodičen, velja isto tudi za premike $\tau_i = N, 2N, 3N, \dots$
 - Za premike $0 < \tau_i < N$ velja, da se pojavi $\frac{1}{2}(N + 1)$ negativnih in $\frac{1}{2}(N - 1)$ pozitivnih produktov $a \times a \rightarrow$ srednja vrednost produkta premaknjenih signalov je enaka $-\frac{a^2}{N}$
 - Glede na vse naštetu in periodičnost s periodo N :

$$\phi_{nn}^d(\tau_i) = \begin{cases} a^2 & \tau_i = 0 \\ -\frac{a^2}{N} & 0 < \tau_i < N \end{cases}$$
$$\phi_{nn}^d(\tau_i \pm iN) = \phi_{nn}^d(\tau_i) \quad i = 0, 1, 2, \dots$$

Psevdonaključni binarni signal (PNBS)



- Avtokorelacijska funkcija signala PNBS:



- Za velike N je avtokorelacijska funkcija PRBS približno enaka
$$\phi_{nn}^d(\tau_i) = a^2 \delta(\tau_i) \quad -N < \tau_i < N$$
- Diskretna spektralna močnostna gostota PNBS:

$$\begin{aligned} \Phi_{nn}^d(m) &= \sum_{\tau_i=0}^{N-1} \phi_{nn}(\tau_i) e^{-j\frac{2\pi}{N}m\tau_i} = a^2 - \frac{a^2}{N} \sum_{\tau_i=1}^{N-1} e^{-j\frac{2\pi}{N}m\tau_i} \\ &= \begin{cases} \frac{a^2}{N} & m = 0 \\ a^2 \left(1 + \frac{1}{N}\right) & 0 < m < N \end{cases} \end{aligned}$$





- Meje integralov v številnih formulah so neskončne. V praksi te meje nadomestimo s končnimi vrednostmi. Rezultirajoča veličina je v tem primeru pravzaprav ocena dejanske veličine in jo označimo s strešico (^). V prisotnosti stohastičnih motenj se pojavita dve vprašanji:
 1. Ali je matematično upanje ocene enako pravi vrednosti? Če je to res, pravimo, da je ocena *nepristranska*. Ocena je *konsistentna*, če se z večanjem intervala opazovanja izboljšuje in limitira k pravi vrednosti, ko gre interval opazovanja proti neskončnosti.
 2. Ali gre varianca napake ocene proti nič, ko gre interval opazovanja proti neskončnosti? Če je to res in če je ocena konsistentna, pravimo, da je ocena konsistentna v srednjekvadratični vrednosti. Takrat gre s podaljševanjem intervala opazovanja tudi varianca napake ocene proti 0 – vsi rezultati so zgoščeni okrog pravega.



- Primer – ocena srednje vrednosti:

- Srednja vrednost stohastične spremenljivke x je:

$$\bar{x} = E\{x(k)\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N x(k)$$

- Ocenjujemo jo po formuli:

$$\hat{\bar{x}} = \frac{1}{N} \sum_{k=1}^N x(k)$$

- Ali je ocena nepristranska, konsistentna in konsistentna v srednjekvadratični vrednosti?
- Matematično upanje ocene srednje vrednosti:

$$E\{\hat{\bar{x}}\} = E\left\{\frac{1}{N} \sum_{k=1}^N x(k)\right\} = \frac{1}{N} \sum_{k=1}^N E\{x(k)\} = \frac{1}{N} N \bar{x} = \bar{x}$$

- Matematično upanje ocene je torej enako pravi vrednosti že za končne intervale opazovanja N
- Ocena je torej **nepristranska** in je seveda tudi **konsistentna**



- Primer – ocena srednje vrednosti:

- Varianca pogreška ocene

$$E\{(\hat{x} - \bar{x})^2\} = E\left\{\left[\frac{1}{N} \sum_{k=1}^N [x(k) - \bar{x}]\right]^2\right\}$$

- Če so posamezne vrednosti naključne spremenljivke x medsebojno statistično neodvisne (beli šum), lahko zapišemo

$$E\{(\hat{x} - \bar{x})^2\} = E\left\{\frac{1}{N^2} \sum_{k=1}^N (x(k) - \bar{x})^2\right\} = \frac{1}{N^2} \sum_{k=1}^N E\{(x(k) - \bar{x})^2\} = \frac{1}{N} \sigma_x^2$$

- Varianca pogreška ocene se z naraščajočim intervalom opazovanja manjša in gre v limiti proti nič
- Ocena je zato **konsistentna v srednjekvadratični vrednosti**





- Primer – ocena variance:

- Varianco stohastične spremenljivke x ocenjujemo po formuli

$$\hat{\sigma}_x^2 = \frac{1}{N} \sum_{k=1}^N [x(k) - \hat{x}]^2$$

- Ali je ocena nepristranska in konsistentna?
- Matematično upanje ocene variance:

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{N} \sum_{k=1}^N (x(k) - \hat{x})^2\right\}$$

- Veljajo povezave:

$$\begin{aligned} \sum_{k=1}^N (x(k) - \hat{x})^2 &= \sum_{k=1}^N [(x(k) - \bar{x}) - (\hat{x} - \bar{x})]^2 = \\ &= \sum_{k=1}^N [x(k) - \bar{x}]^2 + \sum_{k=1}^N (\hat{x} - \bar{x})^2 - 2(\hat{x} - \bar{x}) \sum_{k=1}^N [x(k) - \bar{x}] = \\ &= \sum_{k=1}^N [x(k) - \bar{x}]^2 + N(\hat{x} - \bar{x})^2 - 2(\hat{x} - \bar{x})N(\hat{x} - \bar{x}) = \\ &= \sum_{k=1}^N [x(k) - \bar{x}]^2 - N(\hat{x} - \bar{x})^2 \end{aligned}$$

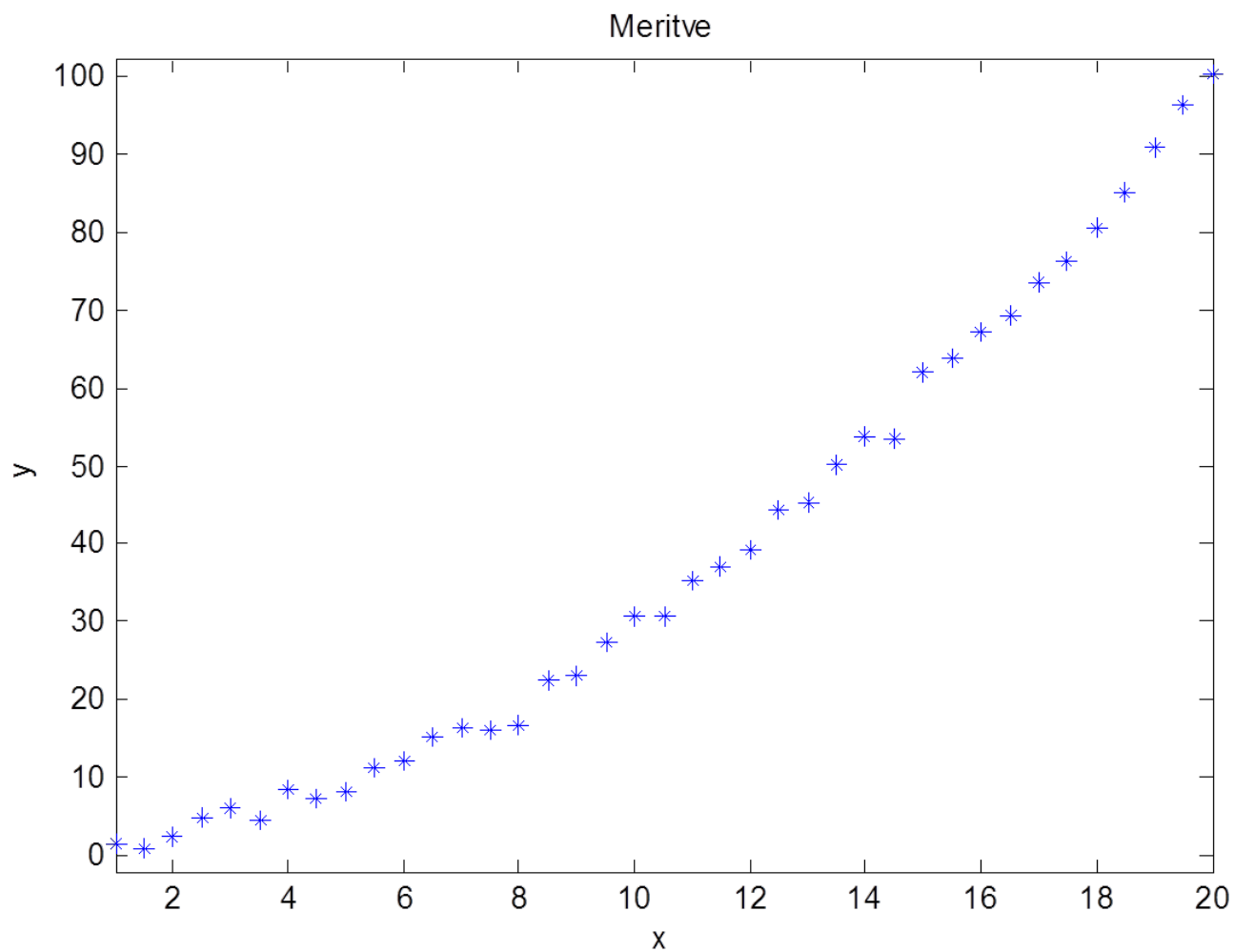
- Zato: $E\{\hat{\sigma}_x^2\} = \frac{1}{N} \left[\sum_{k=1}^N E\{[x(k) - \bar{x}]^2\} - NE\{(\hat{x} - \bar{x})^2\} \right] = \sigma_x^2 - \frac{1}{N} \sigma_x^2$
- Ocena variance je za končne čase N **pristranska**, a **konsistentna**



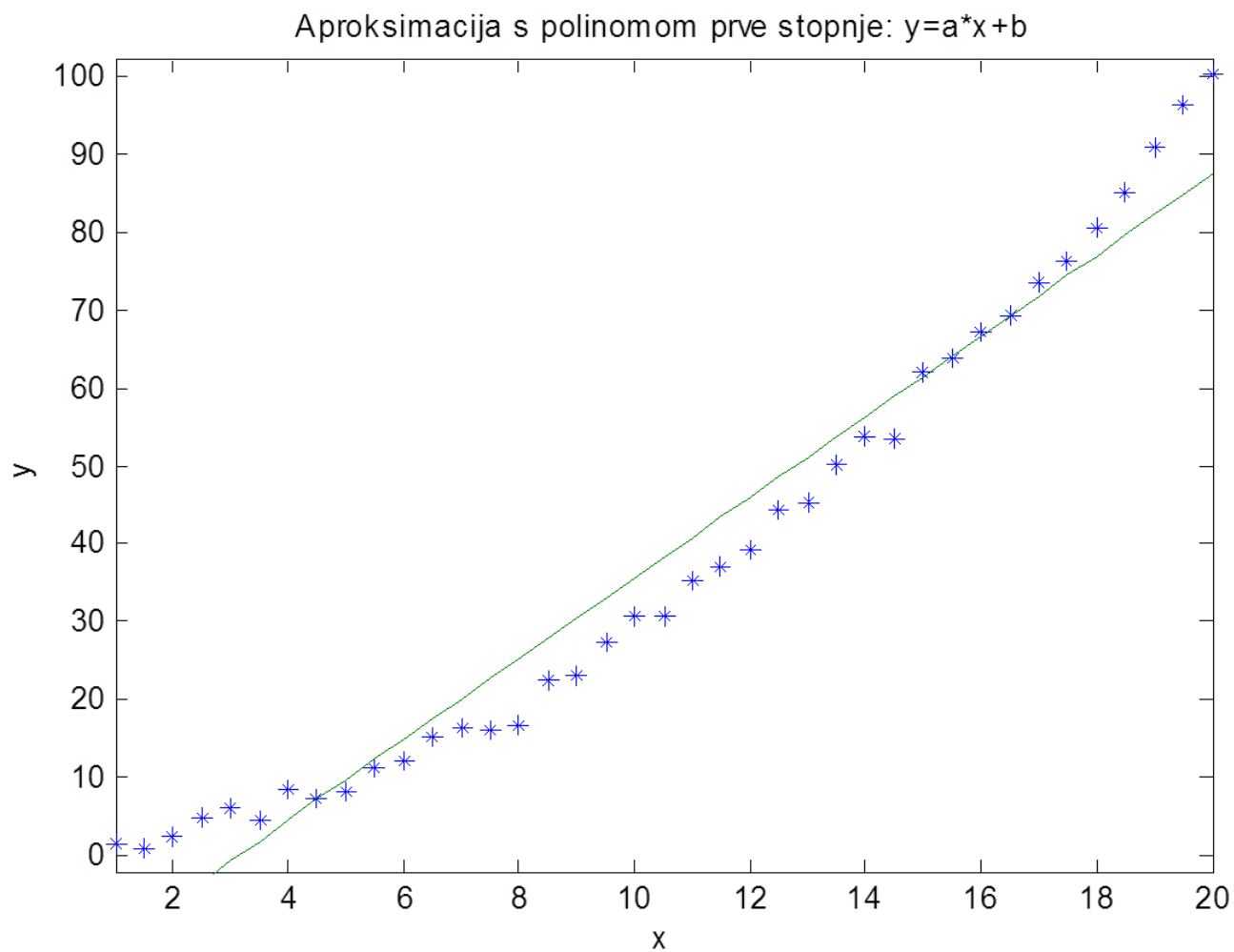


- V splošnem gre za metode, s katerimi določimo neznane parametre sistema – zelo širok spekter metod, ker je tudi široka paleta možnih parametričnih modelov
- Običajno gre za reševanje predoločenega sistema – ima več enačb kot neznank
- Vsem enačbam ne moremo zadostiti, zato se odločimo za kriterijsko funkcijo podobnosti sistema in modela
- Če je kriterijska funkcija vsota kvadratov odstopanj, govorimo o metodi najmanjših kvadratov
- Metodo je zelo stara, saj je bila prvič uporabljena konec 18. stoletja (Gauss)
- Ker je kriterijska funkcija (kvadratna) analitično odvedljiva, se v primeru linearne odvisnosti pogreška od ocenjevanih parametrov problem zelo poenostavi

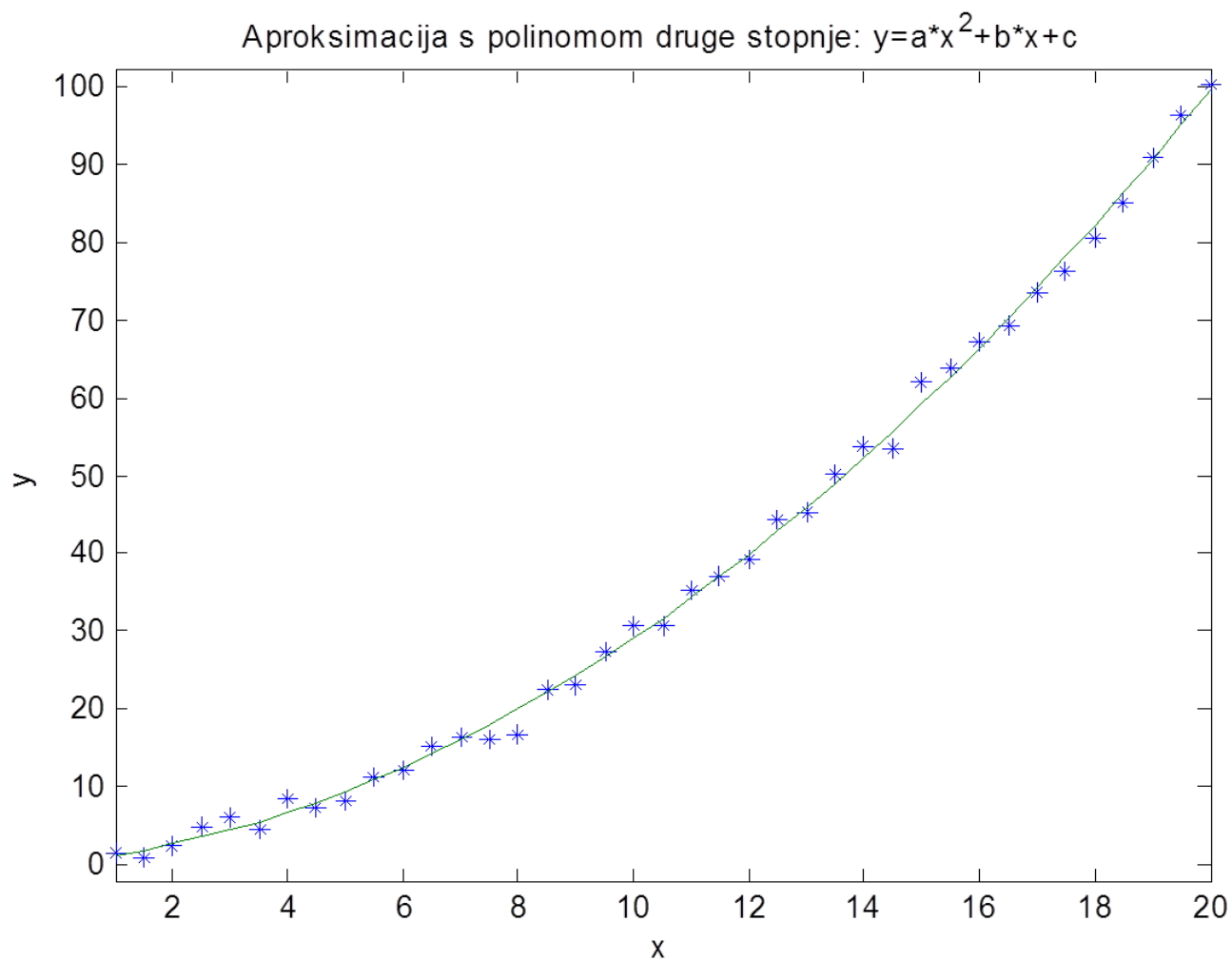
Primer 1



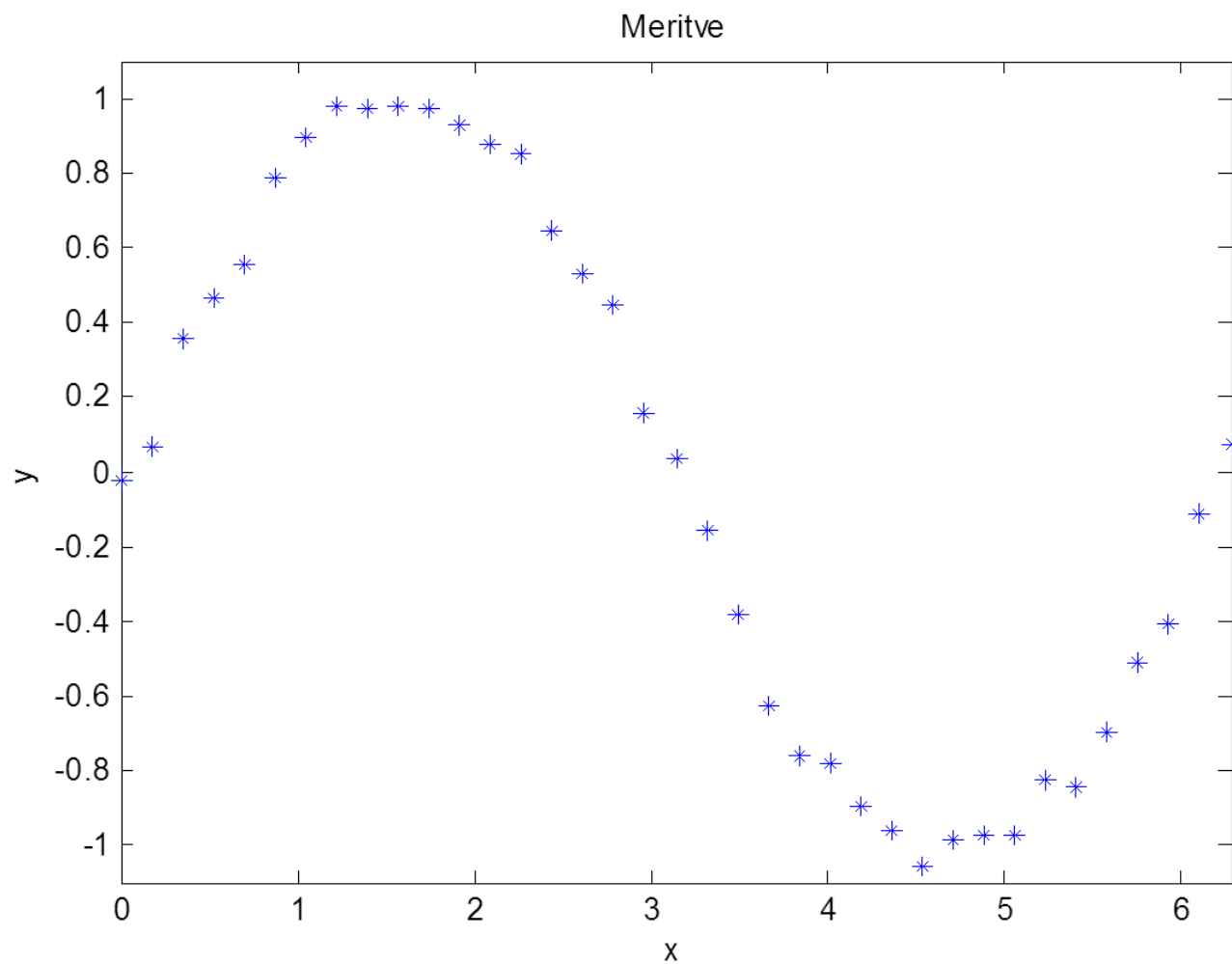
Primer 1



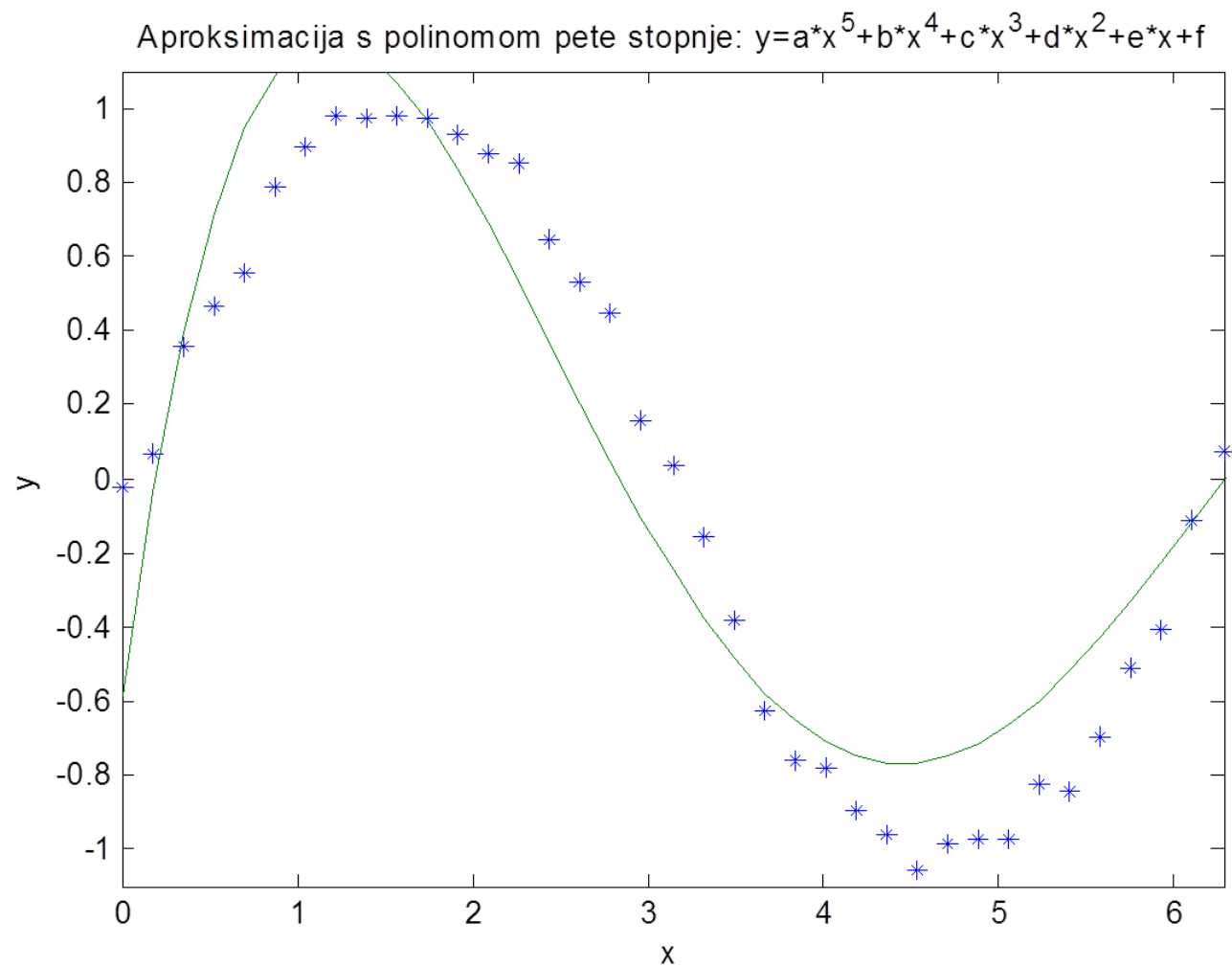
Primer 1



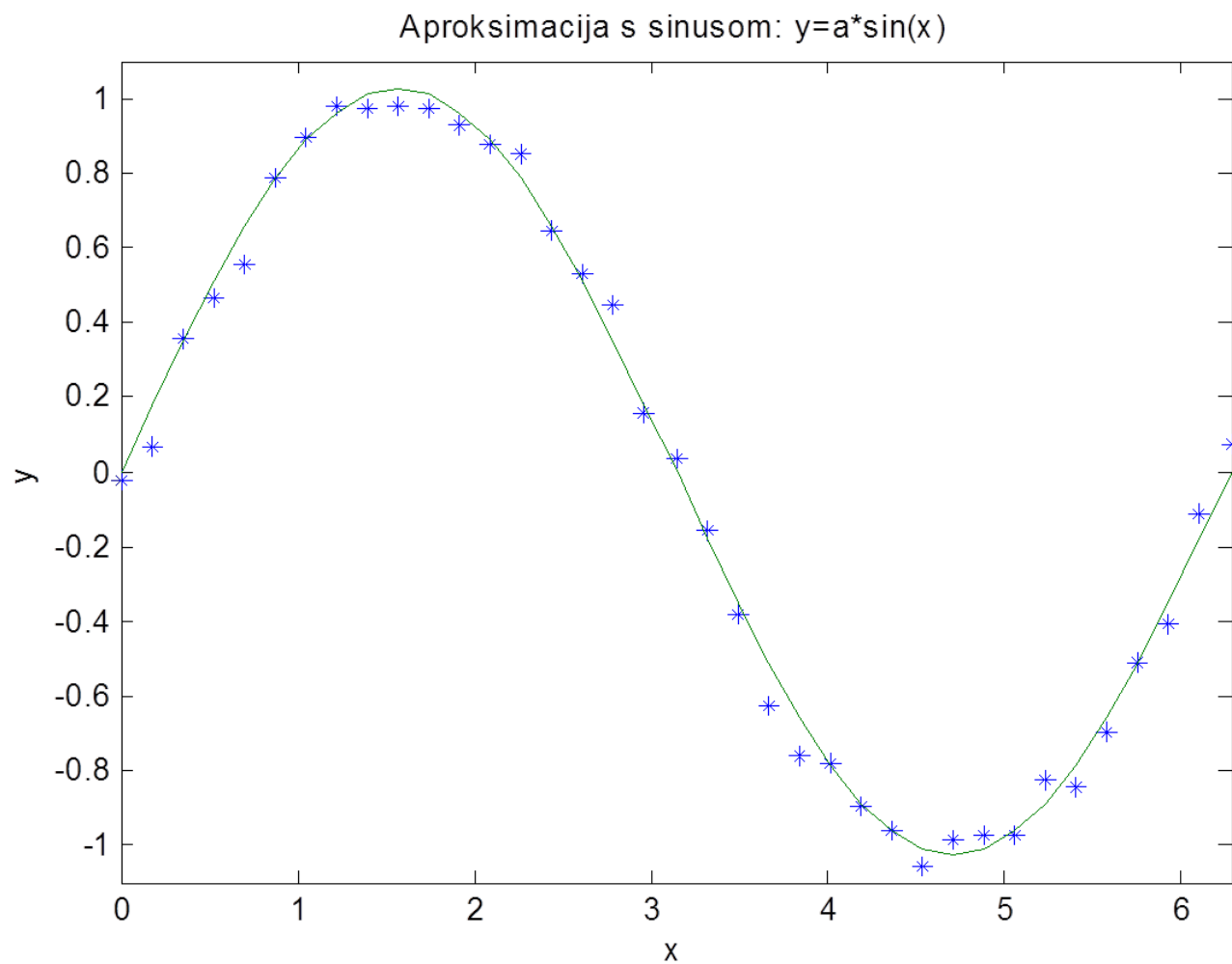
Primer 2



Primer 2



Primer 2



Nekaj zaključkov teh dveh primerov



- Ponavadi se aproksimacija neznane funkcije s povečevanjem ocenjevanih parametrov izboljšuje, ni pa nujno
- Primer 2 je pokazal, da je včasih zelo priporočljivo, če poznamo strukturo sistema, ki ga aproksimiramo ($\sin x$ v našem primeru), ker moramo potem oceniti manj parametrov (le amplitudo), pa še pogrešek med modelom in sistemom je na koncu manjši
- Tudi pri identifikacijah govorimo o strukturni identifikaciji in parametrični identifikaciji:
 - pri strukturni identifikaciji moramo oceniti strukturo sistema
 - nato pa pri parametrični identifikaciji (npr. z metodo najmanjših kvadratov) ocenimo neznane parametre



- Model procesa:

$$y_0(k) = Ku(k)$$

- Moten proces:

$$y(k) = y_0(k) + n(k)$$

- Ocenjeni model procesa:

$$\hat{y}(k) = \hat{K}u(k)$$

- Načeloma je dovolj ena meritev – en par (u, y) , a zaradi motenj uporabimo veliko meritev, ki jih indeksiramo s k (ni nujno, da je k čas; lahko gre za poljubno neodvisno spremenljivko diskretnega značaja)
- Pogrešek ocene meritve:

$$e(k) = y(k) - \hat{y}(k)$$



- Pogreške izračunamo za vseh N opazovanj:

$$y(0) - \hat{K}u(0) = e(0)$$

$$y(1) - \hat{K}u(1) = e(1)$$

$$\vdots$$

$$y(N-1) - \hat{K}u(N-1) = e(N-1)$$

- Kriterijska funkcija je vsota kvadratov pogreškov:

$$V = \sum_{k=0}^{N-1} e^2(k) = \sum_{k=0}^{N-1} [y(k) - \hat{K}u(k)]^2$$

- Kriterijsko funkcijo minimiziramo glede na $\hat{K}$:

$$\frac{\partial V}{\partial \hat{K}} = -2 \sum_{k=0}^{N-1} [y(k) - \hat{K}u(k)]u(k) = 0$$

$$\hat{K} \sum_{k=0}^{N-1} u^2(k) = \sum_{k=0}^{N-1} u(k)y(k)$$



- Rešitev:

$$\hat{K} = \frac{\sum_{k=0}^{N-1} y(k)u(k)}{\sum_{k=0}^{N-1} u^2(k)} = \frac{\hat{\phi}_{uy}(0)}{\hat{\phi}_{uu}(0)}$$

- Drugi odvod pozitiven, zato smo dobili minimum funkcije V :

$$\frac{\partial^2 V}{\partial \hat{K}^2} = 2 \sum_{k=0}^{N-1} u^2(k) \geq 0$$

- Pogoji za obstoj rešitve je, da imenovalec ni enak 0:

$$\sum_{k=0}^{N-1} u^2(k) \neq 0$$

ali

$$\hat{\phi}_{uu}(0) \neq 0$$



- Matematično upanje ocene:

$$\begin{aligned} E\{\hat{K}\} &= E\left\{\frac{\sum_{k=0}^{N-1} [y_0(k) + n(k)]u(k)}{\sum_{k=0}^{N-1} u^2(k)}\right\} = \\ &= \frac{\sum_{k=0}^{N-1} y_0(k)u(k)}{\sum_{k=0}^{N-1} u^2(k)} + E\left\{\frac{\sum_{k=0}^{N-1} n(k)u(k)}{\sum_{k=0}^{N-1} u^2(k)}\right\} \\ &= K + \frac{\sum_{k=0}^{N-1} E\{n(k)u(k)\}}{\sum_{k=0}^{N-1} u^2(k)} \end{aligned}$$

- Ocena je torej nepristranska, če velja:

$$E\{n(k)u(k)\} = 0$$

- Če sta $u(k)$ in $n(k)$ nekorelirana (križna kovarianca je enaka 0):
$$E\{n(k)u(k)\} = E\{n(k)\}E\{u(k)\}$$
- Če ima vsaj en od signalov $u(k)$ ali $n(k)$ srednjo vrednost enako 0, je torej produkt enak 0 in **ocena nepristranska**

Skalarni problem – konsistentnost v srednjkvadratični vrednosti



- Matematično upanje ocene variance:

$$\sigma_K^2 = E\{[\hat{K} - K]^2\} = \frac{E\left\{\left[\sum_{k=0}^{N-1} n(k)u(k)\right]^2\right\}}{\left[\sum_{k=0}^{N-1} u^2(k)\right]^2}$$

- Analiza števca gornjega izraza:

$$\begin{aligned} & E\left\{\sum_{k=0}^{N-1} n(k)u(k) \times \sum_{k'=0}^{N-1} n(k')u(k')\right\} \\ &= \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} E\{n(k)n(k')u(k)u(k')\} \\ &= \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} \phi_{nn}(k - k')\phi_{uu}(k - k') \end{aligned}$$

– k je diskretni čas, in ne le indeks meritve

Skalarni problem – konsistentnost v srednjkvadratični vrednosti



- Poseben primer – motnja je beli šum:

$$\phi_{nn}(\tau) = \Phi_{n0}\delta(\tau) = \sigma_n^2\delta(\tau)$$

- Takrat za izraz v števcu variance ocenjenega parametra velja:

$$\begin{aligned} & \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} \phi_{nn}(k-k')\phi_{uu}(k-k') = \\ & = \sigma_n^2 \sum_{k=1}^{N-1} \phi_{uu}(0) = \sigma_n^2 N \phi_{uu}(0) \approx \sigma_n^2 \sum_{k=0}^{N-1} u^2(k) \end{aligned}$$

↖ enako v primeru neskončnega N

- Izraz za varianco ocene se torej poenostavi:

$$\sigma_K^2 \approx \frac{\sigma_n^2}{\sum_{k=0}^{N-1} u^2(k)} \approx \frac{1}{N} \times \frac{\sigma_n^2}{\sigma_u^2}$$

- V enačbi enakosti, če je srednja vrednost u enaka 0

Skalarni problem – konsistentnost v srednjekvadratični vrednosti



- Poseben primer – u je beli šum:

$$\phi_{uu}(\tau) = \Phi_{u0}\delta(\tau) = \sigma_u^2\delta(\tau)$$

- Izraz v števcu variance ocenjenega parametra:

$$\begin{aligned} \sum_{k=0}^{N-1} \sum_{k'=0}^{N-1} \phi_{nn}(k-k')\phi_{uu}(k-k') &= \\ &= \sigma_u^2 N \phi_{nn}(0) \approx \sigma_u^2 N \sigma_n^2 \end{aligned}$$

- V enačbi enakosti, če je srednja vrednost n enaka 0

- Varianca ocenjenega parametra:

$$\sigma_K^2 \approx \frac{N\sigma_u^2\sigma_n^2}{[\sum_{k=0}^{N-1} u^2(k)]^2} \approx \frac{1}{N} \times \frac{\sigma_n^2}{\sigma_u^2}$$

- V enačbi enakosti, če je srednja vrednost u enaka 0

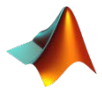
Skalarni problem – konsistentnost v srednjekvadratični vrednosti



- Če je u ali n beli šum in imata oba signala srednjo vrednost 0:

$$\sigma_K^2 \approx \frac{1}{N} \times \frac{\sigma_n^2}{\sigma_u^2}$$

- Varianca napake ocene parametra je premo sorazmerna varianci motnje, obratno sorazmerna varianci vzbujanja in z večanjem števila meritev N pada proti 0, zato je **ocena konsistentna v srednjekvadratični vrednosti**
- To pravilo velja kot dokaj splošno pravilo pri identifikaciji:
 - Varianca ocenjenega parametra oz. pogrešek ocene bo tem večji čim večja je jakost motilnega signala
 - Varianca ocenjenega parametra oz. pogrešek ocene bo padal z večanjem energije vzbujalnega signala





- Ocenjujemo več neznanih parametrov hkrati
- Proces (statično odvisnost izhoda od vhoda) brez motenj lahko zapišemo z enačbo:

$$y_0 = a_1 f_1(u) + a_2 f_2(u) + \dots + a_n f_n(u)$$

- Zelo pomembno dejstvo: y_0 je **linearno odvisen od parametrov** a_i , četudi so $f_i(u)$ poljubne, torej tudi nelinearne funkcije
- Zgornjih n parametrov lahko določimo, če imamo n meritev
- Če pa je meritev več, imamo predoločen sistem – takrat uporabimo metodo najmanjših kvadratov, ki minimizira vsoto kvadratov odstopanj izhoda modela od meritev
- Poseben primer vektorskega problema je ocenjevanje parametrov regresijskih polinomov



- Matematični model regresijske krivulje reda q :

$$y_0 = K_0 + K_1 u + K_2 u^2 + \dots + K_q u^q$$

- Spet imamo N meritev:

$$y_0(k) = K_0 + K_1 u(k) + K_2 u^2(k) + \dots + K_q u^q(k),$$
$$k = 0, 1, \dots, N - 1$$

- Spet imamo tudi šum:

$$y(k) = y_0(k) + n(k)$$

- Ocenjeni model procesa:

$$\hat{y}(k) = \hat{K}_0 + \hat{K}_1 u(k) + \hat{K}_2 u^2(k) + \dots + \hat{K}_q u^q(k)$$

- Pogrešek ocene:

$$e(k) = \hat{y}(k) - y(k)$$



- Zapis vseh N pogreškov ocene:

$$y(0) - \hat{K}_0 - \hat{K}_1 u(0) - \hat{K}_2 u^2(0) - \dots - \hat{K}_q u^q(0) = e(0)$$

$$y(1) - \hat{K}_0 - \hat{K}_1 u(1) - \hat{K}_2 u^2(1) - \dots - \hat{K}_q u^q(1) = e(1)$$

$$\vdots$$

$$y(N-1) - \hat{K}_0 - \hat{K}_1 u(N-1) - \dots - \hat{K}_q u^q(N-1) = e(N-1)$$

- Kompaktni zapis pogreškov ocene:

$$\mathbf{y} - \mathbf{U}\hat{\mathbf{K}} = \mathbf{e}$$

- Zapis enačb procesa:

$$y(0) - K_0 - K_1 u(0) - K_2 u^2(0) - \dots - K_q u^q(0) = n(0)$$

$$y(1) - K_0 - K_1 u(1) - K_2 u^2(1) - \dots - K_q u^q(1) = n(1)$$

$$\vdots$$

$$y(N-1) - K_0 - K_1 u(N-1) - \dots - K_q u^q(N-1) = n(N-1)$$

- Kompaktni zapis enačb procesa:

$$\mathbf{y} - \mathbf{U}\mathbf{K} = \mathbf{n}$$

Vektorski problem



- Uporabljene matrike in vektorji:

$$\mathbf{U} = \begin{bmatrix} 1 & u(0) & u^2(0) & \cdots & u^q(0) \\ 1 & u(1) & u^2(1) & \cdots & u^q(1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & u(N-1) & u^2(N-1) & \cdots & u^q(N-1) \end{bmatrix}$$
$$\mathbf{y} = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} \quad \hat{\mathbf{y}} = \begin{bmatrix} \hat{y}(0) \\ \hat{y}(1) \\ \vdots \\ \hat{y}(N-1) \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e(0) \\ e(1) \\ \vdots \\ e(N-1) \end{bmatrix}$$
$$\mathbf{n} = \begin{bmatrix} n(0) \\ n(1) \\ \vdots \\ n(N-1) \end{bmatrix} \quad \mathbf{K} = \begin{bmatrix} K_0 \\ K_1 \\ \vdots \\ K_q \end{bmatrix} \quad \hat{\mathbf{K}} = \begin{bmatrix} \hat{K}_0 \\ \hat{K}_1 \\ \vdots \\ \hat{K}_q \end{bmatrix}$$



- Kriterijska funkcija – vsota kvadratov pogreškov v vektorsko matrični obliki:

$$\begin{aligned} V &= \mathbf{e}^T \mathbf{e} = [\mathbf{y} - \mathbf{U}\hat{\mathbf{K}}]^T [\mathbf{y} - \mathbf{U}\hat{\mathbf{K}}] = \\ &= [\mathbf{y}^T - \hat{\mathbf{K}}^T \mathbf{U}^T] [\mathbf{y} - \mathbf{U}\hat{\mathbf{K}}] = \\ &= \mathbf{y}^T \mathbf{y} - \hat{\mathbf{K}}^T \mathbf{U}^T \mathbf{y} - (\mathbf{U}^T \mathbf{y})^T \hat{\mathbf{K}} + \hat{\mathbf{K}}^T \mathbf{U}^T \mathbf{U} \hat{\mathbf{K}} \end{aligned}$$

- Kriterijsko funkcijo odvajamo po vektorju parametrov:

$$\frac{\partial V}{\partial \hat{\mathbf{K}}} = -2\mathbf{U}^T \mathbf{y} + 2\mathbf{U}^T \mathbf{U} \hat{\mathbf{K}} = -2[\mathbf{U}^T \mathbf{y} - \mathbf{U}^T \mathbf{U} \hat{\mathbf{K}}] = \mathbf{0}$$

- Dobimo rešitev:

$$\hat{\mathbf{K}} = [\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T \mathbf{y}$$

- Rešitev je minimum, ker velja:

$$\frac{\partial^2 V}{\partial \hat{\mathbf{K}}^2} = 2\mathbf{U}^T \mathbf{U} \quad - \text{pozitivno definitna matrika}$$

Vektorski problem – nepristranskost (šum na y)



- Odvisnost ocene od motenj $\mathbf{n}$:

$$\hat{\mathbf{K}} = [\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T (\mathbf{U} \mathbf{K} + \mathbf{n}) = \mathbf{K} + [\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T \mathbf{n}$$

- Matematično upanje ocene:

$$E\{\hat{\mathbf{K}}\} = \mathbf{K} + E\{[\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T \mathbf{n}\}$$

- Če vhod in motnja nista korelirana, lahko matematično upanje produkta razbijemo na produkt matematičnih upanj (prvi faktor je determinističen in operator matematičnega upanja ni potreben):

$$E\{\hat{\mathbf{K}}\} = \mathbf{K} + [\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T E\{\mathbf{n}\}$$

- Ocena je torej **nepristranska**, če vhodni signal ni koreliran z motnjo in če je srednja vrednost motnje enaka 0



- Kovariančna matrika napake ocene:

$$\begin{aligned}\text{cov}[\hat{\mathbf{K}} - \mathbf{K}] &= E\{[\hat{\mathbf{K}} - \mathbf{K}] \times [\hat{\mathbf{K}} - \mathbf{K}]^T\} \\ &= E\{[\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T \mathbf{nn}^T \mathbf{U} [\mathbf{U}^T \mathbf{U}]^{-1}\}\end{aligned}$$

- Če vhod in motnja nista korelirana, nadalje velja:

$$\text{cov}[\hat{\mathbf{K}} - \mathbf{K}] = [\mathbf{U}^T \mathbf{U}]^{-1} \mathbf{U}^T E\{\mathbf{nn}^T\} \mathbf{U} [\mathbf{U}^T \mathbf{U}]^{-1}$$

- Matematično upanje časovno premaknjenih elementov šuma je enako avtokorelaciji šuma:

$$E\{\mathbf{nn}^T\} = \begin{bmatrix} \phi_{nn}(0) & \phi_{nn}(1) & \cdots & \phi_{nn}(N-1) \\ \phi_{nn}(1) & \phi_{nn}(0) & \cdots & \phi_{nn}(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{nn}(N-1) & \phi_{nn}(N-2) & \cdots & \phi_{nn}(0) \end{bmatrix}$$

Vektorski problem – konsistentnost v srednjekvadratični vrednosti



- Če je motnja beli šum, je njena avtokorelacija delta-impulz:

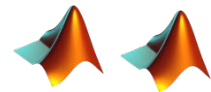
$$\phi_{nn}(\tau) = \Phi_{n0} \delta(\tau)$$

$$E\{\mathbf{nn}^T\} = \Phi_{n0} \mathbf{I}$$

- Potem velja:

$$\text{cov}[\hat{\mathbf{K}} - \mathbf{K}] = \Phi_{n0} [\mathbf{U}^T \mathbf{U}]^{-1} = \frac{1}{N} \Phi_{n0} \left[\frac{1}{N} \mathbf{U}^T \mathbf{U} \right]^{-1}$$

- Izraz v oglatem oklepaju je približno enak kovariančni matriki vhodnega signala (in njegovih potenc) in je zato končen
- Varianca napake ocene gre torej proti 0, če gre število opazovanj proti neskončno – ocena je konsistentna v srednjekvadratični vrednosti





- Problem ocenjevanja parametrov predoločenih sistemov lahko v splošnem opišemo na naslednji način:

$$\mathbf{f}(\mathbf{d}, \Theta) = 0$$

- Vektor $\mathbf{d}$ vsebuje podatke oz. meritve, ki jih poznamo
- Vektor Θ vsebuje neznane oz. iskane parametre
- Nelinearna vektorska funkcija $\mathbf{f}$ podaja povezavo med meritvami in par.

- Pogosto lahko podatke razdelimo na dva dela in zapišemo:

$$\mathbf{y} = \mathbf{g}(\mathbf{x}, \Theta)$$

- Del podatkov, zbran v vektorju $\mathbf{y}$, imenujemo izhod modela – v splošnem nima zveze z izhodom sistema, kot ga običajno definiramo
- Del podatkov, zbran v vektorju $\mathbf{x}$, imenujemo vhod modela

- Problem ocenjevanja parametrov je **linearen v parametrih**, če lahko enačbo sistema zapišemo na naslednji način

$$\mathbf{y} = \mathbf{H}(\mathbf{x})\Theta$$



- V prejšnji enačbi je y poznan vektor, Θ neznan vektor, $H(x)$ pa poznana matrika ustreznih dimenzij
- Sistem, ki ga opisuje enačba $y = H(x)\Theta$, je v splošnem nelinearen (zaradi nelinearne preslikave H nad vektorjem x), vendar je linearen glede na ocenjevani vektor Θ
- Primer: Sistem $y = kx^2$, v katerem na osnovi meritev x in y ocenjujemo k , je nelinearen glede na x , a linearen glede na k
- Primer: Sistem $y = A \sin(\omega t + \varphi)$, v katerem na osnovi meritev t in y ter poznane frekvence ω , ocenjujemo A in φ , je linearen glede na A in nelinearen glede na φ . Sistem torej ni linearen v parametrih. Nato problem zapišemo nekoliko drugače:

$$y = A \sin \omega t \cos \varphi + A \cos \omega t \sin \varphi = \begin{bmatrix} \sin \omega t & \cos \omega t \end{bmatrix} \begin{bmatrix} A \cos \varphi \\ A \sin \varphi \end{bmatrix}$$

- Problem postane linearen glede na nove parametre $A \cos \varphi$ in $A \sin \varphi$



- Če je problem ocenjevanja linearen v parametrih, lahko torej napišemo:

$$\text{znana veličina} = \text{znana veličina} \times \text{neznana veličina}$$

- Kot rečeno, ni nujno, da je veličina na levi strani enačaja izhod sistema, veličina na desni pa vhod v sistem
- Primer: Sistem opisuje enačba:

$$y(k) = Ku^2(k) + \sqrt{v(k)}$$

- $u(k)$ in $v(k)$ sta merljiva vhoda, $y(k)$ pa merljiv izhod sistema
- Enačba za ocenjevanje parameta K se glasi:

$$y(k) - \sqrt{v(k)} = u^2(k) \times K$$

- Izhod modela ocenjevanja K je torej $y(k) - \sqrt{v(k)}$
- Pri ocenjevanju parametrov je smiselno uporabiti karseda veliko informacije, ki je na voljo, zato imamo v praksi običajno veliko več meritev, kot je ocenjevanih parametrov → predoločen sistem



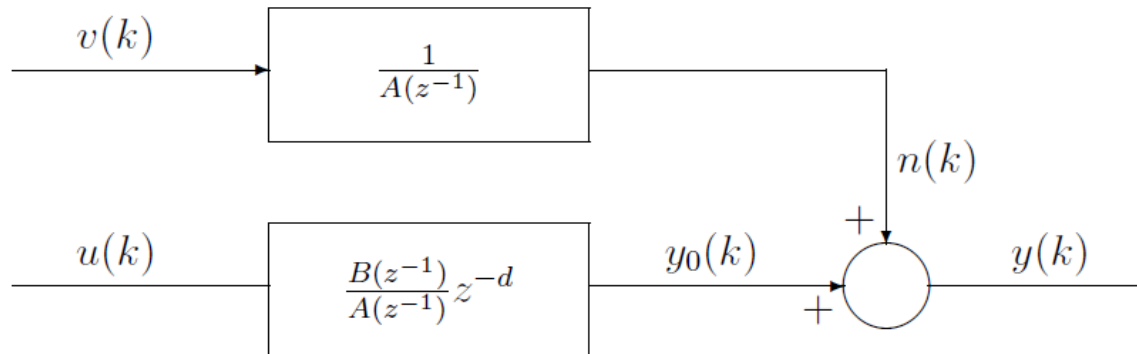
- Metoda najmanjših kvadratov izredno uporabna za reševanje predoločenega sistema enačb, še posebej če neznanke nastopajo linearno oz. je problem ocenjevanja linearen v parametrih
- Metoda se pogosto uporablja za ocenjevanje neznanih parametrov statičnega sistema
- Če so motnje in neodvisne spremenljivke (vhodi v sistem) nekorelirani ter je srednja vrednost motenj enaka 0, potem je metoda nepristranska (torej tudi konsistentna) in konsistentna v srednjekvadratični vrednosti
- Videli bomo, da je zelo pravšnja tudi za ocenjevanje parametrov dinamičnih sistemov (ob upoštevanju določenih predpostavk)



- Obravnavali bomo diskretni dinamični proces s prenosno funkcijo

$$G_p(z) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} = \frac{b_1 z^{-1} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}} z^{-d}$$

- d – časovna zakasnitev – moramo jo vnaprej poznati
 - n – red sistema – moramo ga vnaprej poznati
 - a_i in b_i ($i = 1, \dots, n$) so parametri prenosne funkcije, ki jih določimo z metodo najmanjših kvadratov
- Izhod je moten s šumom n , ki ga generiramo iz šuma v s filtriranjem preko filtra s prenosno funkcijo $G_n(z) = \frac{1}{A(z^{-1})}$





- Za sedaj nismo postavili nobene zahteve glede lastnosti šuma v
- Smiselnost predlaganega šumnega filtra bomo utemeljili kasneje
- Enačba motenega procesa v prostoru z se tako glasi

$$y(z) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(z) + \frac{1}{A(z^{-1})} v(z)$$

- Ta zapis preuredimo:

$$A(z^{-1})y(z) = B(z^{-1})z^{-d}u(z) + v(z)$$

$$\begin{aligned} (1 + a_1 z^{-1} + \dots + a_n z^{-n})y(z) &= \\ &= (b_1 z^{-1} + \dots + b_n z^{-n})z^{-d}u(z) + v(z) \end{aligned}$$

- Sedaj enačbo pretvorimo v časovni prostor (inverzna z -trans.):

$$\begin{aligned} y(k) + a_1 y(k-1) + a_2 y(k-2) + \dots + a_n y(k-n) \\ = b_1 u(k-d-1) + \dots + b_n u(k-d-n) + v(k) \end{aligned}$$



- Naša naloga je oceniti parametre a_i in b_i iz meritev vhodnih ($u(k)$) in izhodnih ($y(k)$) veličin procesa
- Ta naloga se razlikuje od ocenjevanja parametrov nelinearnega statičnega procesa le v tem, da meritve v sistemu enačb niso dobljene v enem samem časovnem trenutku, temveč v nekaj zaporednih časovnih trenutkih

- Vektorski zapis enačbe:

$$y(k) = \boldsymbol{\psi}^T(k) \boldsymbol{\theta} + v(k)$$

- $\boldsymbol{\psi}^T(k) = [-y(k-1), \dots, -y(k-n), u(k-d-1), \dots, u(k-d-n)]$
- $\boldsymbol{\theta}^T = [a_1, \dots, a_n, b_1, \dots, b_n]$ – neznan vektor parametrov

- Ustrezni model tako zapišemo v obliki

$$\hat{y}(k) = \boldsymbol{\psi}^T(k) \hat{\boldsymbol{\theta}} = -\hat{a}_1 y(k-1) - \dots - \hat{a}_n y(k-n) + \\ + \hat{b}_1 u(k-d-1) + \dots + \hat{b}_n u(k-d-n)$$

- $\hat{\boldsymbol{\theta}}^T = [\hat{a}_1, \dots, \hat{a}_n, \hat{b}_1, \dots, \hat{b}_n]$ – ocena parametrov



- Izhod modela $\hat{y}$ v literaturi pogosto imenujejo enokoračno napoved izhoda modela v trenutku k na osnovi meritev v trenutkih $k - 1, k - 2, \dots, k - d - n$
- Zapis $\hat{y}$ v prostoru z (z-transformacija prejšnje diferenčne en.):
$$\hat{y}(z) = [1 - \hat{A}(z^{-1})]y(z) + \hat{B}(z^{-1})z^{-d}u(z)$$
 - $\hat{A}(z^{-1}) = 1 + \hat{a}_1z^{-1} + \dots + \hat{a}_nz^{-n}$
 - $\hat{B}(z^{-1}) = \hat{b}_1z^{-1} + \dots + \hat{b}_nz^{-n}$

- Pogrešek modela:

$$e(k) = y(k) - \hat{y}(k)$$

- Povezava med pogreškom modela in oceno parametrov:

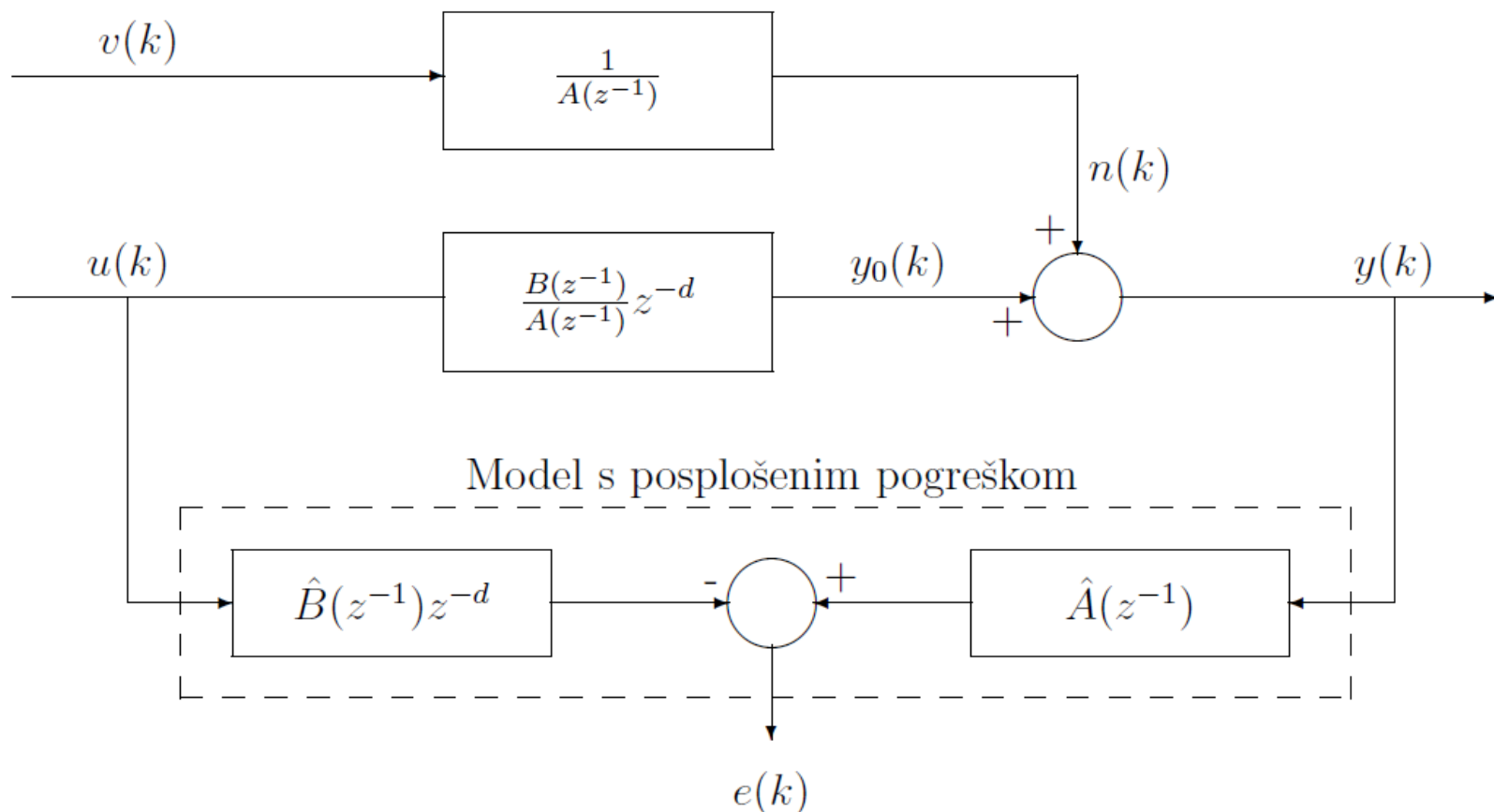
$$e(k) = y(k) - \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\theta}}$$

- Z upoštevanjem $\hat{y}(z)$ lahko enačbo pretvorimo v prostor z :

$$e(z) = \hat{A}(z^{-1})y(z) - \hat{B}(z^{-1})z^{-d}u(z)$$



- Obravnavani model pogreška je pravzaprav model s posplošenim pogreškom:





- Sistem enačb za določitev ocene parametrov $\hat{\theta}$ dobimo tako, da napišemo enačbo $e(k) = y(k) - \psi^T(k)\hat{\theta}$ za N opazovanj
- Število opazovanj mora biti enako ali večje (v praksi mnogo večje) od števila parametrov, ki jih ocenjujemo $N \geq 2n$
- Ob tekočem času k zapišemo zgornjo enačbo za trenutni čas in za preteklih $N - 1$ časovnih trenutkov in dobimo sistem enačb:
$$\begin{aligned}e(k - N + 1) &= y(k - N + 1) - \psi^T(k - N + 1)\hat{\theta} \\e(k - N + 2) &= y(k - N + 2) - \psi^T(k - N + 2)\hat{\theta} \\&\vdots \\e(k - 1) &= y(k - 1) - \psi^T(k - 1)\hat{\theta} \\e(k) &= y(k) - \psi^T(k)\hat{\theta}\end{aligned}$$
- Bolj kompakten zapis enačb dobimo, če uporabimo matrično-vektorski zapis



- Vektorji:

$$\mathbf{y} = \begin{bmatrix} y(k - N + 1) \\ y(k - N + 2) \\ \vdots \\ y(k - 1) \\ y(k) \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e(k - N + 1) \\ e(k - N + 2) \\ \vdots \\ e(k - 1) \\ e(k) \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} v(k - N + 1) \\ v(k - N + 2) \\ \vdots \\ v(k - 1) \\ v(k) \end{bmatrix}$$

$$\boldsymbol{\psi}^T(k) = [-y(k - 1) \quad \dots \quad -y(k - n) \quad u(k - 1 - d) \quad \dots \quad u(k - n - d)]$$

- Matrika:

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}^T(k - N + 1) \\ \boldsymbol{\psi}^T(k - N + 2) \\ \vdots \\ \boldsymbol{\psi}^T(k - 1) \\ \boldsymbol{\psi}^T(k) \end{bmatrix}$$

$$= \begin{bmatrix} -y(k - N) & \dots & -y(k - N - n + 1) & u(k - N - d) & \dots & u(k - N - n - d + 1) \\ \vdots & & \vdots & \vdots & & \vdots \\ -y(k - 1) & \dots & -y(k - n) & u(k - 1 - d) & \dots & u(k - n - d) \end{bmatrix}$$



- Sistem enačb lahko zapišemo v vektorsko-matrični obliki:

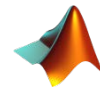
$$\mathbf{y} - \mathbf{\Psi}^T \hat{\boldsymbol{\theta}} = \mathbf{e}$$

- Na podoben način bi lahko v istih časovnih trenutkih zapisali tudi enačbo $y(k) - \boldsymbol{\psi}^T(k) \boldsymbol{\theta} = v(k)$:

$$\mathbf{y} - \mathbf{\Psi}^T \boldsymbol{\theta} = \mathbf{v}$$

- Pri obravnavi ocene parametrov za vektorski primer smo prišli do rezultata:

$$\hat{\boldsymbol{\theta}} = [\mathbf{\Psi}^T \mathbf{\Psi}]^{-1} \mathbf{\Psi}^T \mathbf{y}$$



- Matrika $\mathbf{\Psi}^T \mathbf{\Psi}$ je vedno simetrična kvadratna matrika dimenzije $2n \times 2n$ (ne glede na število opazovanj)
- Ker jo je potrebno invertirati, mora biti nesingularna (determinanta različna od 0 oz. polnega ranga):
 - dosežemo s primernim vzbujanjem
 - vhodni signal vsebuje dovolj frekvenc



- Če v enačbo ocene vnesemo povezavo med $\mathbf{y}$ in $\mathbf{v}$, dobimo:

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T (\boldsymbol{\Psi} \boldsymbol{\theta} + \mathbf{v}) = \boldsymbol{\theta} + [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T \mathbf{v}$$

- Matematično upanje ocene je torej:

$$E\{\hat{\boldsymbol{\theta}}\} = \boldsymbol{\theta} + E\{[\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T \mathbf{v}\}$$

- Če je $E\{\boldsymbol{\Psi}^T \mathbf{v}\}$ različen od 0, je metoda pristranska:

- V tem vektorju so sami produkti tipov $v(i)y(i-c)$ in $v(i)u(i-c)$, kjer je i poljubno število, c pa poljubno pozitivno število
- Ocena $\hat{\boldsymbol{\theta}}$ je torej nepristranska, če je šum $v(k)$ nekoreliran tako s preteklimi vrednostmi vhoda kot s preteklimi vrednostmi izhoda, hkrati pa mora biti srednja vrednost ali šuma ali vhoda in izhoda 0
- Nekoreliranost šuma z vhodom ni problematična
- Nekoreliranost šuma z izhodom ni samoumevna, ker šum vpliva na izhod, vendar pa lahko zagotovimo nekoreliranost šuma $v(k)$ s preteklimi vrednostmi izhoda, če $v(k)$ ni koreliran z $v(k-1)$ – beli šum



- Ocena parametrov $\hat{\theta}$ nepristranska, če je izhod procesa moten s šumom $n(k)$, ki je dobljen iz belega šuma preko šumnega filtra s prenosno funkcijo $1/A(z^{-1})$ in če je ali srednja vrednost vzbujevalnega signala ali srednja vrednost šuma enaka nič
- V primeru nepristranskosti sta $\hat{\theta}$ in θ enaka, kar pomeni, da sta enaka tudi e in v in posledično velja:
$$e(k) = v(k)$$
- Pogrešek modela $e(k)$ torej konvergira k šumu $v(k)$:
 - Pogrešek modela $e(k)$ nastane v nadomestni shemi iz belega šuma $v(k)$, ki je filtriran preko šumnega filtra $1/A(z^{-1})$ ter preko desne strani posplošenega modela pogreška $\hat{A}(z^{-1})$
 - Zaradi nepristranskosti $\hat{A}(z^{-1}) = A(z^{-1})$ in pr. f. filtra je 1: $e(k) = v(k)$
 - Pogrešek modela $e(k)$ ni vodljiv iz vhoda $u(k)$, ker se v primeru nepristranskosti obe veji iz vhoda $u(k)$ uničita



- Izračunajmo še kovariančno matriko pogreška ocene $\hat{\boldsymbol{\theta}}$:
$$\text{cov}[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}] = E \left\{ [\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}] \times [\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}]^T \right\} = E \{ [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T \mathbf{v} \mathbf{v}^T \boldsymbol{\Psi} [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \}$$
- Če je $v(k)$ beli šum, potem ni koreliran z $\boldsymbol{\Psi}$, je pa enak $e(k)$:
$$E \{ \mathbf{v} \mathbf{v}^T \} = E \{ \mathbf{e} \mathbf{e}^T \} = \sigma_e^2 \mathbf{I}$$
- Kovariančna matrika je torej:
$$\text{cov}[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}] = \sigma_e^2 E \{ [\boldsymbol{\Psi}^T \boldsymbol{\Psi}]^{-1} \}$$
- Elementi po glavni diagonali kovariančne matrike predstavljajo varianco (kvadrat standardne deviacije) ocene posameznega parametra
- Matrika $\boldsymbol{\Psi}^T \boldsymbol{\Psi}$ je matrika korelacijskih funkcij, pomnožena z N
- Ker to matriko invertiramo, gre z naraščanjem N proti $0 \rightarrow$ ocena je konsistentna v srednjekvadratični vrednosti

Rekurzivna metoda najmanjših kvadratov



- Doslej obravnavana metoda najmanjših kvadratov je enokoračna metoda, to se pravi, da najprej shranjujemo merjene signale, nato pa po koncu meritev izračunamo oceno parametrov
- Če želimo ocenjevati parametre sprotno (on line) v realnem času, nas zanima ocena parametrov med samim opazovanjem vhodnega in izhodnega signala (med samo meritvijo)
- Vsako novo opazovanje (v času $k + 1$) doprinese nekaj nove informacije, ki se matematično pokaže v eni dodatni enačbi, oziroma v podaljšanju vektorja $\mathbf{y}$ za vrednost $y(k + 1)$ ter v podaljšanju matrike Ψ za vrstico $\Psi^T(k + 1)$
- Če bi uporabili nerekurzivno metodo najmanjših kvadratov, bi morali v vsakem koraku invertirati produkt matrik $\Psi^T \Psi$ ter ga pomnožiti z vektorjem $\Psi^T \mathbf{y}$ (zelo zamudno)
- Enačbe raje transformiramo v rekurzivno obliko

Rekurzivna metoda najmanjših kvadratov



- Rekurzivna metoda najmanjših kvadratov je matematično ekvivalentna navadni metodi najmanjših kvadratov, a se pri njej ocena parametrov v vsakem koraku izračuna iz ocene v prejšnjem koraku
- Matrike in vektorji se s časom spreminjajo, zato jih označimo kot časovne funkcije (z dodatkom argumenta k)
- Nerekurzivna enačba za ocenjevanje parametrov postane:

$$\hat{\boldsymbol{\theta}}(k) = [\boldsymbol{\Psi}^T(k)\boldsymbol{\Psi}(k)]^{-1}\boldsymbol{\Psi}^T(k)\mathbf{y}(k)$$

- Uvedemo novo matriko

$$\mathbf{P}(k) = [\boldsymbol{\Psi}^T(k)\boldsymbol{\Psi}(k)]^{-1}$$

- Podaljševanje vektorjev in matrik ob času $k + 1$:

$$\mathbf{y}(k + 1) = \begin{bmatrix} \mathbf{y}(k) \\ y(k + 1) \end{bmatrix} \quad \boldsymbol{\Psi}(k + 1) = \begin{bmatrix} \boldsymbol{\Psi}(k) \\ \boldsymbol{\psi}^T(k + 1) \end{bmatrix}$$

$$\boldsymbol{\psi}^T(k + 1) = [-y(k) \quad \dots \quad -y(k - n + 1) \quad u(k - d) \quad \dots \quad u(k - n - d + 1)]$$

Rekurzivna metoda najmanjših kvadratov



- Najprej zapišemo rekurzijo za $\mathbf{P}^{-1}$:

$$\mathbf{P}^{-1}(k+1) = \mathbf{P}^{-1}(k) + \boldsymbol{\psi}(k+1)\boldsymbol{\psi}^T(k+1)$$

- Rekurzivno enačbo za $\hat{\boldsymbol{\theta}}$ dobimo po krajši izpeljavi:

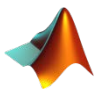
$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + \mathbf{P}(k+1)\boldsymbol{\psi}(k+1)[y(k+1) - \boldsymbol{\psi}^T(k+1)\hat{\boldsymbol{\theta}}(k)]$$

$$\begin{array}{c} \text{nova} \\ \text{ocena} \end{array} = \begin{array}{c} \text{stara} \\ \text{ocena} \end{array} + \begin{array}{c} \text{ojačenje} \\ \text{korekcije} \end{array} \left[\begin{array}{c} \text{nova} \\ \text{meritev} \end{array} - \begin{array}{c} \text{napoved meritve} \\ \text{na osnovi stare ocene} \end{array} \right]$$

- Za rešitev te diferenčne enačbe potrebujemo začetni pogoj $\hat{\boldsymbol{\theta}}(0)$
- Za določitev $\hat{\boldsymbol{\theta}}(k)$ potrebujemo tudi diferenčno enačbo za $\mathbf{P}$, ki jo dobimo iz rekurzivne enačbe za $\mathbf{P}^{-1}$ z uporabo izreka o inverziji matrik:

$$\mathbf{P}(k+1) = \mathbf{P}(k) - \frac{\mathbf{P}(k)\boldsymbol{\psi}(k+1)\boldsymbol{\psi}^T(k+1)\mathbf{P}(k)}{1 + \boldsymbol{\psi}^T(k+1)\mathbf{P}(k)\boldsymbol{\psi}(k+1)}$$

- Za rešitev te diferenčne enačbe potrebujemo začetni pogoj $\mathbf{P}(0)$
- Običajno izberemo: $\mathbf{P}(0) = \alpha \mathbf{I} \quad \alpha \gg 1$



Metoda uteženih najmanjših kvadratov



- Pri navadni metodi najmanjših kvadratov smo utežili pogreške vseh enačb oz. meritev enako
- Če je stopnja zaupanja v posamezne meritve zaradi kakršnega koli razloga različna, lahko različnim enačbam predpišemo različne uteži

- Kriterijska funkcija V ima namesto oblike $V = \mathbf{e}^T \mathbf{e}$ obliko:

$$V = \mathbf{e}^T \mathbf{W} \mathbf{e}$$

- $\mathbf{W}$ je (diagonalna) matrika uteži posameznih meritev:

$$\mathbf{W} = \begin{bmatrix} w(k - N + 1) & 0 & 0 & \dots & 0 \\ 0 & w(k - N + 2) & 0 & \dots & 0 \\ 0 & 0 & w(k - N + 3) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & w(k) \end{bmatrix}$$

- Pri metodi uteženih najmanjših kvadratov je lahko matrika $\mathbf{W}$ v splošnem tudi poljubna pozitivno definitna matrika

Metoda uteženih najmanjših kvadratov



- Na zelo podoben način kot pri izpeljavi ocene parametrov po metodi najmanjših kvadratov dobimo:

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Psi}^T \mathbf{W} \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T \mathbf{W} \mathbf{y}$$

- Tudi pogoji za pristranskost in konsistenco so enaki
- Kovariančna matrika pogreška ocene parametrov je:

$$\text{cov}[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}] = E\{[\boldsymbol{\Psi}^T \mathbf{W} \boldsymbol{\Psi}]^{-1} \boldsymbol{\Psi}^T \mathbf{W} \mathbf{v} \mathbf{v}^T \mathbf{W} \boldsymbol{\Psi} [\boldsymbol{\Psi}^T \mathbf{W} \boldsymbol{\Psi}]^{-1}\}$$

- Najmanjšo varianco dobimo, ko je utežna matrika $\mathbf{W}$ enaka inverzni kovariančni matriki šuma $E\{\mathbf{v} \mathbf{v}^T\}$:

$$\mathbf{W}^{-1} = E\{\mathbf{v} \mathbf{v}^T\}$$

- Ocena z najmanjšo varianco se imenuje Markovska ocena
- V primeru, ko je $v(k)$ beli šum, je njegova kovariančna matrika enaka enotini matriki in ocenitvena enačba postane ocenitvena enačba enostavne metode najmanjših kvadratov



- Pozabljanje pomeni, da imajo enačbe, ki temeljijo na starih podatkih, manjšo težo kakor tiste, ki temeljijo na novejših
- Pomembno predvsem pri ocenjevanju parametrov časovno spremenljivih procesov
- Eksponencialno pozabljanje je poseben primer uteževanja:
 - Utež posamezne enačbe v vsakem naslednjem trenutku zmanjšamo za faktor λ
 - Ko so podatki aktualni (v trenutku k), je utež enaka $w(k)$
 - V naslednjem trenutku ($k + 1$) pa je enaka $w(k + 1) = \lambda(k + 1)w(k)$
 - Če naj informacija izgublja na teži, mora biti λ manjši od 1
 - Za $\lambda = 1$ dobimo navadno metodo uteženih najmanjših kvadratov brez pozabljanja
 - Interval parametra λ :

$$0 < \lambda \leq 1 \text{ (v praksi } 0,95 \leq \lambda \leq 0,99\text{)}$$

Rekurzivna metoda uteženih najmanjših kvadratov s pozabljanjem



- Tudi pozabljanje je lahko časovno spremenljivo, zato $\lambda(k)$
- Za zapis utežne matrike lahko napišemo rekurzivno enačbo:

$$\mathbf{W}(k+1) = \begin{bmatrix} \lambda(k+1)\mathbf{W}(k) & \mathbf{0} \\ \mathbf{0}^T & w(k+1) \end{bmatrix}$$

- Definicija matrike $\mathbf{P}(k)$ je tokrat nekoliko drugačna:

$$\mathbf{P}(k) = [\boldsymbol{\Psi}^T(k)\mathbf{W}(k)\boldsymbol{\Psi}(k)]^{-1}$$

- Spet lahko zapišemo oceno parametrov kot:

$$\hat{\boldsymbol{\theta}}(k) = \mathbf{P}(k)\boldsymbol{\Psi}^T(k)\mathbf{W}(k)\mathbf{y}(k)$$

- Rekurzivna enačba za $\mathbf{P}^{-1}$ se v primeru rekurzivne metode uteženih najmanjših kvadratov s pozabljanjem glasi:

$$\mathbf{P}^{-1}(k+1) = \lambda(k+1)\mathbf{P}^{-1}(k) + w(k+1)\boldsymbol{\Psi}(k+1)\boldsymbol{\Psi}^T(k+1)$$

- Z upoštevanjem izreka o inverziji matrik izpeljemo:

$$\mathbf{P}(k+1) = \left(\mathbf{P}(k) - \frac{w(k+1)\mathbf{P}(k)\boldsymbol{\Psi}(k+1)\boldsymbol{\Psi}^T(k+1)\mathbf{P}(k)}{\lambda(k+1) + w(k+1)\boldsymbol{\Psi}^T(k+1)\mathbf{P}(k)\boldsymbol{\Psi}(k+1)} \right) \frac{1}{\lambda(k+1)}$$

Rekurzivna metoda uteženih najmanjših kvadratov s pozabljanjem



- Potrebujemo še rekurzivno enačbo za oceno parametrov:
$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + w(k+1)\mathbf{P}(k+1)\boldsymbol{\psi}(k+1)[y(k+1) - \boldsymbol{\psi}^T(k+1)\hat{\boldsymbol{\theta}}(k)]$$
- Rekurzija za oceno parametrov $\hat{\boldsymbol{\theta}}$ se razlikuje od ustrezne ocene pri navadni metodi najmanjših kvadratov le po faktorju $w(k+1)$, to je uteži aktualnih podatkov, ki je dodana ojačenju korekcije
- Rekurzija za $\mathbf{P}$ skriva v sebi veliko nevarnost:
 - V primeru, ko je $\lambda < 1$ in ni vzbujanja ($\boldsymbol{\psi}(k+1) = \mathbf{0}$), matrika $\mathbf{P}^{-1}(k)$ konvergira proti nič, kar pomeni da gre njen inverz proti ∞
 - $\mathbf{P}(k)$ je ojačenje korekcije, zato pride takoj, ko se pojavi najmanjša napaka med novo meritvijo in njeno napovedjo na osnovi parametrov zadnjega koraka, do velikih sprememb ocene parametrov (pravimo, da algoritem najprej „zaspi“, nato pa „eksplodira“)
 - Velik $\mathbf{P}(k)$ pomeni tudi veliko varianco napake ocene parametrov
 - Opisani fenomen nezaželen $\rightarrow$ spremenljivo eksponentialno pozabljanje, kjer se faktor $\lambda(k)$ prilagaja količini informacije v $\boldsymbol{\psi}$



- Če želimo ocenjevati parametre časovno spremenljivih procesov, moramo dopustiti stalno spremembo ocene vektorja ocenjenih parametrov $\rightarrow$ matrika $\mathbf{P}(k)$, ki pri vseh metodah pomeni nekakšno „ojačenje“ ne sme iti proti nič (konec popravkov), zato:
 - Pozabljanje – časovno pozabljanje onemogoči, da bi elementi matrike $\mathbf{P}(k)$ postali zelo majhni in s tem dopusti spremembe $\hat{\boldsymbol{\theta}}(k)$
 - Resetiranje – pri tej metodi občasno postavimo matriko $\mathbf{P}(k)$ na neko vrednost (če se parametri spreminjajo časovno diskretno, bi bilo idealno, da resetiramo matriko $\mathbf{P}(k)$ prav v trenutkih sprememb parametrov; če so ti trenutki neznani, si lahko pomagamo tako, da resetiramo matriko periodično, oziroma takrat, ko se njena „velikost“ (sled matrike) zmanjša pod neko določeno vrednost)
- Obe metodi za ocenjevanje parametrov časovno spremenljivih procesov imata svoje prednosti in slabosti (resetiranje uporabno pri občasnih in skokovitih spremembah, pozabljanje pri lezenju)



- Meritve $u(k)$ in $y(k)$, ki tvorijo vektor $\Psi(k)$, ne predstavljajo absolutnih vrednosti merjenega vhoda in izhoda procesa, temveč le odstopanja od ustaljenih vrednosti, ki ju imenujemo tudi delovna točka, torej

$$u(k) = U(k) - U_{00}$$

$$y(k) = Y(k) - Y_{00}$$

- $U(k)$ in $Y(k)$ sta absolutni vrednosti vhoda in izhoda, z U_{00} in Y_{00} pa pripadajoči ustaljeni vrednosti
- $U(k) = U_{00}$ povzroči v stacionarnem stanju signal $Y(k) = Y_{00}$
- V daljšem časovnem obdobju morata biti srednji vrednosti tako vhodnega, kakor tudi izhodnega signala enaki nič, torej $\sum_{k=0}^N u(k) = 0$ in $\sum_{k=0}^N y(k) = 0$, zato:

$$U_{00} = \frac{1}{N+1} \sum_{k=0}^N U(k) \quad \text{in} \quad Y_{00} = \frac{1}{N+1} \sum_{k=0}^N Y(k)$$



- Na izhodni signal seveda vplivajo tudi motnje $n(k)$
- Če je srednja vrednost motnje različna od nič, se to pozna le pri ustaljeni vrednosti izhodnega signala, zato lahko vedno privzamemo, da je srednja vrednost (matematično upanje) motnje enaka nič ($E\{n(k)\} = 0$)
- Delovno točko bi torej lahko ocenili kot $Y_{00} = \frac{1}{N+1} \sum_{k=0}^N Y(k)$, vendar lahko to metodo uporabimo le pri nesprotnem ocenjevanju parametrov
- Pri sprotnem ocenjevanju pa uporabimo eno izmed naslednjih dveh metod:
 - Delovno točko ocenjujemo hkrati s parametri
 - Vpliv delovne točke lahko izničimo z diferenciranjem



- Diferenčno enačbo zapišemo za $u(k) = U(k) - U_{00}$ in $y(k) = Y(k) - Y_{00}$:
$$Y(k) = -a_1 Y(k-1) - \dots - a_n Y(k-n) + b_1 U(k-d-1) + \dots + b_n U(k-d-n) + Y_{00} + a_1 Y_{00} + \dots + a_n Y_{00} - b_1 U_{00} - \dots - b_n U_{00} + v(k)$$
- Vse člene, ki združujejo ustaljene vrednosti, združimo:
$$K = Y_{00} + a_1 Y_{00} + \dots + a_n Y_{00} - b_1 U_{00} - \dots - b_n U_{00}$$
- Sedaj imamo dodaten neznan parameter in dodaten regresor:
$$\boldsymbol{\theta}^T = [a_1, \dots, a_n, b_1, \dots, b_n, K]$$
$$\boldsymbol{\psi}^T(k) = [-Y(k-1), \dots, -Y(k-n), U(k-d-1), \dots, U(k-d-n), 1]$$
 - Regresorji so sedaj meritve in ne njihova odstopanja od delovnih vrednosti
- Enačba sistema:
$$Y(k) = \boldsymbol{\psi}^T(k) \boldsymbol{\theta} + v(k)$$
 - Klasična oblika, zato tudi klasično reševanje

Izničenje delovne točke z diferenciranjem



- Zaradi konstantnosti ustaljenih vrednosti sta diferenci absolutnih signalov enaki diferencam odstopanj signalov od delovne točke:

$$\Delta U(k) = U(k) - U(k-1) = u(k) - u(k-1) = \Delta u(k)$$

$$\Delta Y(k) = Y(k) - Y(k-1) = y(k) - y(k-1) = \Delta y(k)$$

- Če zapišemo diferenčno enačbo procesa še za trenutek $k-1$ in obe enačbi odštejemo drugo od druge, dobimo:

$$\begin{aligned} \Delta y(k) = & -a_1 \Delta y(k-1) - a_2 \Delta y(k-2) - \dots - a_n \Delta y(k-n) \\ & + b_1 \Delta u(k-d-1) + \dots + b_n \Delta u(k-d-n) + v(k) - v(k-1) \end{aligned}$$

- Ker pa velja $\Delta Y(k) = \Delta y(k)$ in $\Delta U(k) = \Delta u(k)$:

$$\begin{aligned} \Delta Y(k) = & -a_1 \Delta Y(k-1) - a_2 \Delta Y(k-2) - \dots - a_n \Delta Y(k-n) \\ & + b_1 \Delta U(k-d-1) + \dots + b_n \Delta U(k-d-n) + v(k) - v(k-1) \end{aligned}$$

- Sedaj imamo sistem:

$$\Delta Y(k) = \boldsymbol{\Psi}^T(k) \boldsymbol{\theta} + v(k) - v(k-1)$$

$$\boldsymbol{\Psi}^T(k) = [-\Delta Y(k-1), \dots, -\Delta Y(k-n), \Delta U(k-d-1), \dots, \Delta U(k-d-n)]$$



- V zadnjem izrazu imamo motilni šum v obliki $v(k) - v(k - 1)$ namesto $v(k)$, zato je nadaljnji postopek enak klasičnemu
- Metoda je nepristranska, če je $v(k) - v(k - 1)$ beli šum (motilni signal $v(k)$ dobljen iz belega šuma s pomočjo filtra $\frac{1}{1-z^{-1}}$) oz. $n(k)$ dobljen iz belega šuma preko filtra $G_n(z) = \frac{1}{(1-z^{-1})A(z^{-1})}$
- Obe metodi ima seveda dobre in slabe lastnosti:
 - Dobra lastnost prve metode je, da je manj občutljiva na šum, vendar pa zaradi dodatnega parametra konvergira počasneje
 - Druga metoda ima enako število ocenjevanih parametrov kot originalna metoda, vendar pa diferenciranje merjenih vrednosti ojači šum, kar zopet slabo vpliva na konvergenco
 - Pri prvi metodi dobimo parameter $\hat{K}$, ki povezuje ustaljeni vrednosti vhoda in izhoda neposredno, medtem, ko ga je potrebno pri drugi metodi oceniti ločeno



- Numerične težave lahko nastopijo, ko uporabimo algoritem ocenjevanja parametrov na slabo pogojenem problemu
- Parametrška identifikacija procesa zahteva, da poznamo red procesa in njegovo zakasnitev:
 - Ko ne poznamo točnega reda procesa, lahko izberemo višji red → razširiti moramo vektor regresorjev, ki pa ne vplivajo na izhod, zato so „prave“ vrednosti dodatnih parametrov enake 0
 - Tudi določitev premajhne zakasnitve procesa vodi do slabo pogojenega procesa (pri izbranem mrtvem času d , ki je manjši od dejanskega, so parametri $b_1, \dots, b_d$ odvečni)
 - Posledica redundantnosti nekaterih parametrov je slabo pogojen problem ocenjevanja parametrov – ocenjujemo pare polov in ničel, ki so si podobni, hkrati pa ne prispevajo skoraj nič k vhodno-izhodnemu obnašanju procesa (odvečni poli in ničle imajo majhen efekt na lastnosti modela vse dokler ostanejo poli stabilni)



- Slabo pogojen proces, oziroma linearna odvisnost vektorjev v matriki $\Psi^T \Psi$ lahko nastopi tudi, če vzorčimo proces prepogosto ali če je ustaljena vrednost, ki nastopa v signalnih vrednostih prevelika
- Slabo pogojenost matrike $\Psi^T \Psi$ se zrcali na naslednji način:
 - majhni pogreški v izmerjenih vrednostih imajo velik vpliv na ocene parametrov
 - dobljeni parametri, ki precej odstopajo od dejanskih rešitev, dajo le majhne pogreške enačb
- Pojav kvantificiramo s faktorjem pogojenosti κ , ki ga bomo izpeljali na primeru linearne enačbe:

$$\mathbf{A}\boldsymbol{\theta} = \mathbf{b} \quad \mathbf{A} = \Psi^T \Psi, \mathbf{b} = \Psi^T \mathbf{y}$$

- Če pride do napake v vektorju $\mathbf{b}$ za $\Delta \mathbf{b}$, se ustrezno pokvari tudi $\boldsymbol{\theta}$

$$\mathbf{A}(\boldsymbol{\theta} + \Delta \boldsymbol{\theta}) = \mathbf{b} + \Delta \mathbf{b} \Rightarrow \Delta \boldsymbol{\theta} = \mathbf{A}^{-1} \Delta \mathbf{b}$$



- V nadaljevanju bomo analizirali normi vektorjev $\mathbf{b} = \mathbf{A}\boldsymbol{\theta}$ in $\Delta\boldsymbol{\theta} = \mathbf{A}^{-1}\Delta\mathbf{b}$:

$$\|\mathbf{b}\| = \|\mathbf{A}\boldsymbol{\theta}\| \leq \|\mathbf{A}\| \cdot \|\boldsymbol{\theta}\|$$

$$\|\Delta\boldsymbol{\theta}\| = \|\mathbf{A}^{-1}\Delta\mathbf{b}\| \leq \|\mathbf{A}^{-1}\| \cdot \|\Delta\mathbf{b}\|$$

- S preureditvijo neenačb dobimo:

$$\left. \begin{array}{l} \frac{1}{\|\boldsymbol{\theta}\|} \leq \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|} \\ \|\Delta\boldsymbol{\theta}\| \leq \|\mathbf{A}^{-1}\| \cdot \|\Delta\mathbf{b}\| \end{array} \right\} \Rightarrow \frac{\|\Delta\boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|} \leq \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

- Relativna napaka vektorja $\mathbf{b}$ na relativno napako vektorja ocenjenih parametrov $\boldsymbol{\theta}$ vpliva preko faktorja pogojenosti:

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| \geq 1$$

- Čim manjši je faktor pogojenosti, tem manjša je občutljivost ocenjenih parametrov na napake v meritvah $\mathbf{b}$



- Kaj sploh je $\|\mathbf{A}\|$?
 - Matrična norma je vedno vsiljena oz. inducirana z izborom vektorske norme
 - Če za normo vektorja izberemo evklidsko normo (Pitagorovo dolžino), je inducirana matrična norma t.i. spektralna norma:
$$\|\mathbf{A}\| = \max \sigma_i(\mathbf{A})$$
- Kaj pa je $\sigma_i(\mathbf{A})$?
 - $\sigma_i(\mathbf{A})$ so singularne vrednosti matrike, ki so vedno realne in pozitivne
 - Če je matrika kvadratna (dim. $n \times n$), ima n singularnih vrednosti:
$$\sigma_i(\mathbf{A}) = \lambda_i(\mathbf{A}^T \mathbf{A}) = \lambda_i(\mathbf{A} \mathbf{A}^T) \quad i = 1, 2, \dots, n$$
- Norma $\|\mathbf{A}\|$ je torej enaka največji singularni vrednosti $\sigma_{\max}(\mathbf{A})$, norma $\|\mathbf{A}^{-1}\|$ pa inverzu najmanjše singularne vrednosti $\frac{1}{\sigma_{\min}(\mathbf{A})}$
- Faktor pogojenosti κ je torej $\kappa(\mathbf{A}) = \frac{\sigma_{\max}}{\sigma_{\min}} \geq 1$



Metoda razširjenih najmanjših kvadratov



- Pogoji za nepristransko oceno parametrov pri originalni metodi najmanjših kvadratov je posebna oblika motilnega signala $n(k)$, ki mora biti dobljen iz belega šuma $v(k)$ preko filtra $G_n(z) = \frac{1}{A(z^{-1})}$
- To je seveda huda zahteva, ki je v praksi le redko izpolnjena
- Prva možnost za omiljenje pogoja za nepristransko oceno je v nastavku bolj splošnega šumnega filtra

$$G_n(z) = \frac{D(z^{-1})}{A(z^{-1})} = \frac{1 + d_1 z^{-1} + \dots + d_n z^{-n}}{1 + a_1 z^{-1} + \dots + a_n z^{-n}}$$

- Enačba motenega procesa:

$$y(z) = \frac{B(z^{-1})}{A(z^{-1})} z^{-d} u(z) + \frac{D(z^{-1})}{A(z^{-1})} v(z)$$

$$A(z^{-1})y(z) = B(z^{-1})z^{-d}u(z) + D(z^{-1})v(z)$$

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \dots - a_n y(k-n) + b_1 u(k-d-1) + \dots + b_n u(k-d-n) + v(k) + d_1 v(k-1) + \dots + d_n v(k-n)$$

Metoda razširjenih najmanjših kvadratov



- Definirajmo razširjena vektorja:

$$\Psi_0^T(k) = [-y(k-1), \dots, -y(k-n), u(k-d-1), \dots, u(k-d-n), v(k-1), \dots, v(k-n)]$$
$$\theta^T = [a_1, \dots, a_n, b_1, \dots, b_n, d_1, \dots, d_n]$$

- Enačba procesa:

$$y(k) = \Psi_0^T(k)\theta + v(k)$$

- Če bi poznali šum $v(k)$, oziroma njegove zakasnjene vrednosti $v(k-1), \dots, v(k-n)$, bi lahko uporabili navadno metodo najmanjših kvadratov za oceno razširjenega vektorja parametrov

$$\hat{\theta}^T = [\hat{a}_1, \dots, \hat{a}_n, \hat{b}_1, \dots, \hat{b}_n, \hat{d}_1, \dots, \hat{d}_n]$$

- Šuma $v(k)$ seveda ne poznamo, poznamo pa pogrešek modela $e(k)$
- V primeru nepristranske ocene velja $v(k) = e(k)$, zato v vektorju $\Psi_0^T(k)$ zamenjamo zakasnjene vrednosti šuma z zakasnjenimi vrednostmi pogreška in definiramo:

$$\Psi^T(k) = [-y(k-1), \dots, -y(k-n), u(k-d-1), \dots, u(k-d-n), e(k-1), \dots, e(k-n)]$$



- Sedaj lahko uporabimo enačbe poljubne rekurzivne različice metode najmanjših kvadratov, pri čemer so dimenzije matrike $\mathbf{P}(k)$ ter vektorjev $\hat{\boldsymbol{\theta}}(k)$ in $\boldsymbol{\psi}(k)$ ustrezno večje ($3n \times 3n$ oz. $3n \times 1$)
- Pri nerekurzivni metodi je definicija $e(k)$ jasna:
$$e(k) = y(k) - \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\theta}}$$
- Pri rekurzivni metodi pa se pojavi vprašanje, kateri $\hat{\boldsymbol{\theta}}$ vzamemo:
 - Če vzamemo oceno parametrov iz prejšnjega trenutka, dobimo ***apriorno oceno pogreška***:
$$e(k) = y(k) - \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\theta}}(k-1)$$
 - Če vzamemo oceno parametrov iz tekočega trenutka, dobimo ***aposteriorno oceno pogreška***:
$$\varepsilon(k) = y(k) - \boldsymbol{\psi}^T(k)\hat{\boldsymbol{\theta}}(k)$$
- Dve vrsti pogreška $\rightarrow$ dve različici metode razširjenih najmanjših kvadratov – razlika je le v zadnjih n elementih v $\boldsymbol{\psi}$ (e ali ε)



- V praksi se lastnosti obeh variant le malo razlikujejo, zato uporabljamo pretežno varianto z apriornim pogreškom $e(k)$, ki ga tako in tako potrebujemo za oceno parametrov v trenutku k ($\hat{\boldsymbol{\theta}}(k)$)
- Algoritem za primer apriornega pogreška je torej zanka po k :

$$\boldsymbol{\Psi}^T(k) = [-y(k-1), \dots, -y(k-n), u(k-d-1), \dots, u(k-d-n), e(k-1), \dots, e(k-n)]$$

$$\mathbf{P}(k) = \left(\mathbf{P}(k-1) - \frac{w(k)\mathbf{P}(k-1)\boldsymbol{\Psi}(k)\boldsymbol{\Psi}^T(k)\mathbf{P}(k-1)}{\lambda(k) + w(k)\boldsymbol{\Psi}^T(k)\mathbf{P}(k-1)\boldsymbol{\Psi}(k)} \right) \frac{1}{\lambda(k)}$$

$$e(k) = y(k) - \boldsymbol{\Psi}^T(k)\hat{\boldsymbol{\theta}}(k-1)$$

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + w(k)\mathbf{P}(k)\boldsymbol{\Psi}(k) \underbrace{\left[y(k) - \boldsymbol{\Psi}^T(k)\hat{\boldsymbol{\theta}}(k-1) \right]}_{e(k)}$$

- Apriorni pogrešek lahko zapišemo v daljši (izpisani obliki):

$$e(k) = y(k) + \hat{a}_1 y(k-1) + \dots + \hat{a}_n y(k-n) - \hat{b}_1 u(k-d-1) - \dots - \hat{b}_n u(k-d-n) - \hat{d}_1 e(k-1) - \dots - \hat{d}_n e(k-n)$$

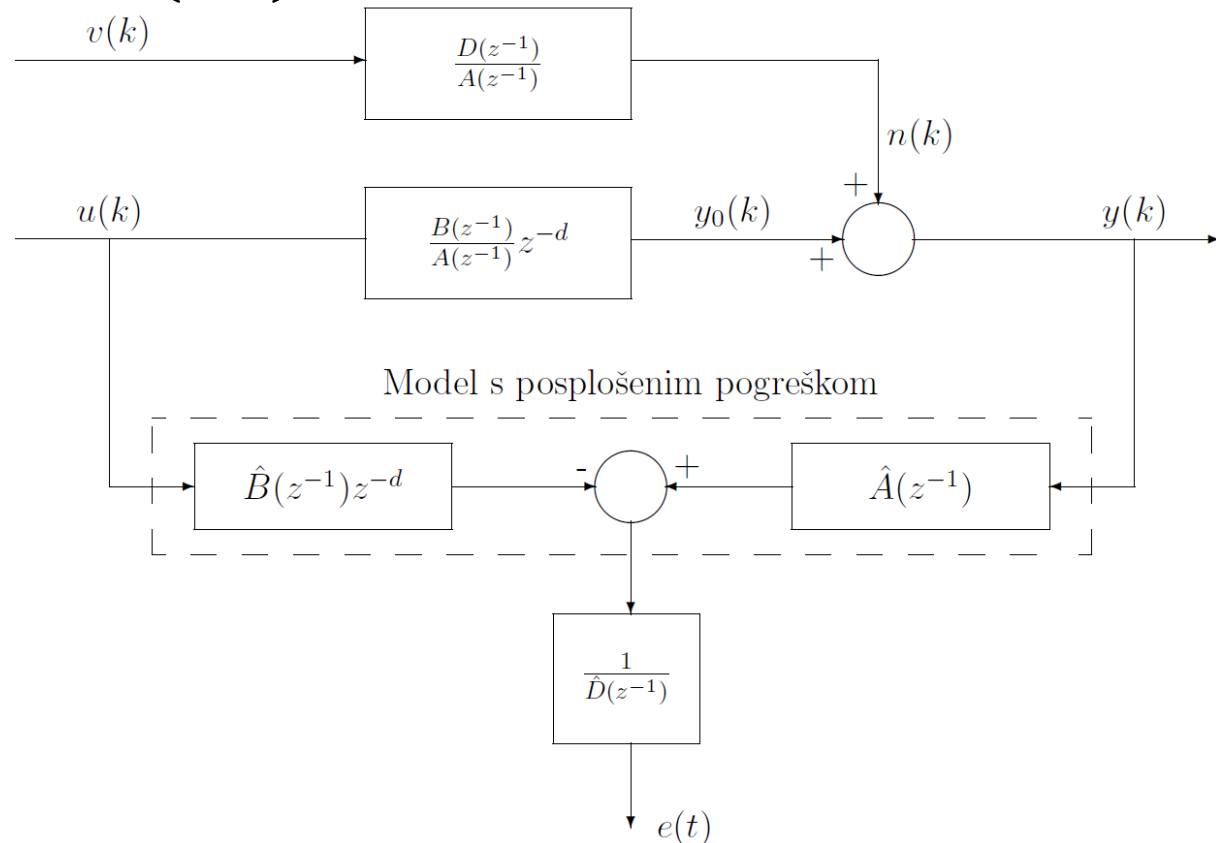
$$e(z) = \hat{A}(z^{-1})y(z) - \hat{B}(z^{-1})z^{-d}u(z) + [1 - \hat{D}(z^{-1})]e(z)$$

Metoda razširjenih najmanjših kvadratov



- Iz prejšnje enačbe lahko izpeljemo $e(z)$:

$$e(z) = \frac{1}{\hat{D}(z^{-1})} (\hat{A}(z^{-1})y(z) - \hat{B}(z^{-1})z^{-d}u(z))$$



Metoda razširjenih najmanjših kvadratov



- Ta slika ilustrira bistvo metode razširjenih najmanjših kvadratov:
 - Če filtriramo beli šum s filtrom s prenosno funkcijo $G_n(z)$, potem dobimo varianco izhodnega šuma $n(k)$ tako, da pomnožimo varianco belega šuma $v(k)$ z vsoto kvadratov impulznega odziva filtra $G_n(z)$:
$$\sigma_n^2 = \sigma_v^2 (g_{n0}^2 + g_{n1}^2 + g_{n2}^2 + \dots)$$
 - Filtri z moničnima (prvi koeficient je 1) polinomoma v števcu in imenovalcu lahko torej varianco šuma le poslabšajo
 - Tvorjenje signala pogreška $e(k)$ na prejšnji sliki je pravzaprav filtriranje belega šuma $v(k)$ preko filtra z moničnim števcem in moničnim imenovalcem
 - Metoda najmanjših kvadratov išče najmanjšo varianco pogreška $e(k)$, torej skuša narediti iz signala $e(k)$ beli šum $v(k)$
 - Ker pa vsebuje šumni filter $G_n(z)$ števec $D(z^{-1})$, moramo signal pogreška, ki bi ga dobili pri navadni metodi najmanjših kvadratov filtrirati s filtrom $\frac{1}{\hat{D}(z^{-1})}$



- Prednost pred navadno metodo najmanjših kvadratov:
 - Šumni filter $\frac{D(z^{-1})}{A(z^{-1})}$ je bolj splošen od šumnega filtra $\frac{1}{A(z^{-1})}$, ki smo ga predstavili pri navadni metodi najmanjših kvadratov, zato je razred šumnih signalov, pri katerih daje metoda razširjenih najmanjših kvadratov nepristransko oceno, širši od razreda šumnih signalov pri navadni metodi najmanjših kvadratov
- Slabosti:
 - Povečano število ocenjevanih parametrov poslabša hitrost konvergence metode razširjenih najmanjših kvadratov
 - Večja težava:
 - Signal pogreška $e(k)$ oziroma $\varepsilon(k)$, ki smo ga izračunali v koraku k , uporabimo za ocenjevanje parametrov v naslednjih korakih
 - S tem smo vpeljali v algoritem ocenjevanja povratno zanko
 - Sistemi s povratno zanko pa lahko postanejo nestabilni



- Metoda razširjenih najmanjših kvadratov je lokalno stabilna v okolici $\tilde{\theta} = \mathbf{0}$ (torej za majhne vrednosti začetnih pogojev), če je izpolnjen pogoj:

$$\operatorname{Re} \left[\frac{1}{D(e^{j\omega})} \right] > 0 \quad \text{za vse realne } \omega$$

- Metoda razširjenih najmanjših kvadratov je globalno stabilna, če je izpolnjen pogoj:

$$\operatorname{Re} \left[\frac{1}{D(e^{j\omega})} - \frac{\lambda}{2} \right] > 0 \quad \text{za vse realne } \omega$$

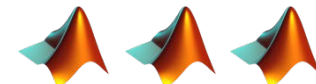
- Stabilnost metode pomeni, da signali ne divergirajo in torej ostanejo omejeni
- Stabilnost še ne pomeni, da je metoda nepristranska in konsistentna



- Ocena parametrov konvergira k pravi vrednosti, če sta poleg pogoja stabilnosti izpolnjena še naslednja pogoja:
 - Matrika $\mathbf{P}(k)$ gre proti 0
 - Šum na vhodu šumnega filtra $v(k)$ je beli šum
- Prvi pogoj je izpolnjen, če je vektor $\boldsymbol{\psi}(k)$ oz. $\boldsymbol{\psi}_\varepsilon(k)$ stalno različen od nič (poleg tega pa morajo biti njegove komponente medsebojno linearno neodvisne), kar imenujemo pogoj stalnega vzbujanja
- Če pogoji niso izpolnjeni, dobimo pristranske rezultate
- Navadna metoda najmanjših kvadratov, pri kateri je $D(e^{j\omega})=1$, je lokalno in globalno stabilna

$$\operatorname{Re} \left[\frac{1}{D(e^{j\omega})} - \frac{\lambda}{2} \right] > 0,$$

$$0 < \lambda \leq 1$$



Metoda največje podobnosti



- Metoda največje podobnosti (s tujko *Maximum likelihood*) je splošen postopek ocenjevanja parametrov statističnega modela
- Že zelo zgodaj so jo uporabili za ocenjevanje parametrov dinamičnih sistemov (Åström in Bohlin)
- Ideja metode je v tem, da kot rešitev identifikacijskega problema poišče vrednosti parametrov $\hat{\theta}$, ki dajo opazovanemu rezultatu največjo verjetnost (kateri $\hat{\theta}$ iz celotnega nabora možnih je najbolj verjeten glede na izmerjene meritve in dani model?)
- V praksi to naredimo, da definiramo t.i. podobnostno funkcijo
- Vrednosti parametrov, ki maksimizirajo podobnostno funkcijo, predstavljajo rešitev identifikacijskega problema
- Metoda se zelo poenostavi, če predpostavimo normalno porazdeljene in statistično neodvisne signale (predpostavimo enako obliko motenj kot pri metodi razširjenih najmanjših kvadratov)



- Obstajata originalna nerekurzivna verzija in rekurzivna verzija
- Poudariti je potrebno, da je tudi nerekurzivna metoda iterativna, ker je signal pogreška nelinearna funkcija parametrov
- Izpeljava metode je precej zapletena. Najprej definiramo filtrirane vektorje signalov:

$$y_f(k) = y(k) - \hat{d}_1 y_f(k-1) - \dots - \hat{d}_n y_f(k-n)$$

$$u_f(k) = u(k) - \hat{d}_1 u_f(k-1) - \dots - \hat{d}_n u_f(k-n)$$

$$e_f(k) = e(k) - \hat{d}_1 e_f(k-1) - \dots - \hat{d}_n e_f(k-n)$$

$$\boldsymbol{\varphi}^T(k) = [-y_f(k-1), \dots, -y_f(k-n), u_f(k-d-1), \dots, u_f(k-d-n), e_f(k-1), \dots, e_f(k-n)]$$

– Za generiranje uporabimo ocene parametrov števca šumnega filtra

- Rekurzivni formuli za $\hat{\boldsymbol{\theta}}$ in $\mathbf{P}$ (5x $\boldsymbol{\varphi}$ namesto $\boldsymbol{\psi}$, 1 $\boldsymbol{\psi}$ ostane):

$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + w(k+1)\mathbf{P}(k+1)\boldsymbol{\varphi}(k+1)[y(k+1) - \boldsymbol{\psi}^T(k+1)\hat{\boldsymbol{\theta}}(k)]$$

$$\mathbf{P}(k+1) = \left(\mathbf{P}(k) - \frac{w(k+1)\mathbf{P}(k)\boldsymbol{\varphi}(k+1)\boldsymbol{\varphi}^T(k+1)\mathbf{P}(k)}{\lambda(k+1) + w(k+1)\boldsymbol{\varphi}^T(k+1)\mathbf{P}(k)\boldsymbol{\varphi}(k+1)} \right) \frac{1}{\lambda(k+1)}$$



- Spomnimo se: metoda najmanjših kvadratov daje nepristransko oceno parametrov le, če so ocene križnokorelacijskih funkcij $\phi_{yv}(i)$ ($i = 1, 2 \dots n$) enake nič (izpolnjeno le pri posebni obliki šumnega filtra $1/A(z^{-1})$)
- Metoda pomožnih spremenljivk (s tujko *instrumental variables*) odpravi pristranskost tako, da zamenja spremenljivke y , ki so seveda „pokvarjene“ s šumom s tako imenovanimi pomožnimi spremenljivkami x , ki so nekorelirane z motnjo v
- Tako daje metoda nepristransko oceno pri poljubni obliki šumnega filtra, vendar pa je, kot bomo videli, metoda stabilna le lokalno



- Izpeljava:
 - Enačbo za oceno parametrov po metodi najmanjših kvadratov bi lahko neformalno izpeljali tudi tako, da bi enačbo $\mathbf{y} - \mathbf{\Psi}\mathbf{\theta} = \mathbf{v}$ premultiplicirali s $\mathbf{\Psi}^T$ in izrazili vektor $\mathbf{\theta}$:
$$\mathbf{\theta} = [\mathbf{\Psi}^T \mathbf{\Psi}]^{-1} \mathbf{\Psi}^T \mathbf{y} - [\mathbf{\Psi}^T \mathbf{\Psi}]^{-1} \mathbf{\Psi}^T \mathbf{v}$$
 - Oceno $\hat{\mathbf{\theta}}$ dobimo, če upoštevamo prvi člen zgornje enačbe
 - Drugi člen predstavlja pristranskost $\mathbf{b}$
 - Če premultipliciramo enačbo z $\mathbf{W}^T$ namesto z $\mathbf{\Psi}^T$, dobimo oceno:
$$\hat{\mathbf{\theta}} = [\mathbf{W}^T \mathbf{\Psi}]^{-1} \mathbf{W}^T \mathbf{y}$$
 - Pristranskost pa je na podoben način kot zgoraj:
$$\mathbf{b} = [\mathbf{W}^T \mathbf{\Psi}]^{-1} \mathbf{W}^T \mathbf{v}$$
- Če so elementi matrike pomožnih spremenljivk $\mathbf{W}$ nekorelirani s šumom $\mathbf{v}$, je pristranskost enaka 0
- Glavni problem: Kako izbrati elemente matrike $\mathbf{W}$?



- Spremenljivke oz. elementi matrike $\mathbf{W}$ morajo biti:
 - nekorelirane z motilnim signalom
 - močno korelirane s koristnima signaloma $u(k)$ in $y_0(k) \rightarrow$ ta zahteva je pomembna zato, ker mora biti matrika $\mathbf{W}^T \mathbf{\Psi}$ pozitivno definitna
- Matrika $\mathbf{\Psi}$ vsebuje vhodne signale u , ki so nekorelirani s šumnim signalom, in izhodne signale y , ki so seveda korelirani s šumom
- Idealna matrika pomožnih spremenljivk $\mathbf{W}$ bi morala vsebovati namesto motenih izhodnih signalov y nemotene signale y_0 , vendar le-ti niso dostopni $\rightarrow$ nadomestimo jih s pomožnimi spremenljivkami, ki jih dobimo s pomočjo ocenjenega vektorja parametrov procesa (po navadni metodi najmanjših kvadratov)
- Izvedemo torej simulacijo odziva sistema na podan vhodni signal in sistem, ki ga definira vektor parametrov $\hat{\boldsymbol{\theta}} \rightarrow y_s(k)$



- Algoritem:
 1. Iz meritev procesa $u(k)$ in $y(k)$ sestavimo matriko Ψ
 2. Po navadni metodi najmanjših kvadratov določimo oceno $\hat{\theta}_{LS}$:
$$\hat{\theta}_{LS} = [\Psi^T \Psi]^{-1} \Psi^T y$$
 3. Izračunamo simulacijo sistema, podanega z $\hat{\theta}_{LS}$, na vzbujanje $u(k)$:
$$y_s(k)$$
 4. Sestavimo matriko W (meritev y nadomestimo s simulirano y_s)
 5. Izračunamo oceno parametrov po metodi pomožnih spremenljivk:
$$\hat{\theta}_{IV} = [W^T \Psi]^{-1} W^T y$$
- Pri nerekurzivni metodi pomožnih spremenljivk lahko postopamo iterativno (vračamo se na korak 3, kjer namesto $\hat{\theta}_{LS}$ pri simulaciji uporabimo oceno $\hat{\theta}_{IV}$) – ponavljamo dokler se ocena parametrov ne spreminja več (v splošnem zadošča malo iteracij)
- Obstaja tudi rekurzivna različica metode



- Metoda stohastične aproksimacije uporablja gradientni algoritem za minimizacijo kriterijske funkcije

$$V(k + 1) = \frac{1}{2} e^2(k + 1)$$

- Kriterijska funkcija upošteva le trenutno vrednost pogreška modela:

$$e(k + 1) = y(k + 1) - \boldsymbol{\Psi}^T(k + 1) \hat{\boldsymbol{\theta}}(k)$$

- Po gradientnem algoritmu dobimo novo vrednost ocene parametrov v smeri največjega **padca** kriterijske funkcije:

$$\hat{\boldsymbol{\theta}}(k + 1) = \hat{\boldsymbol{\theta}}(k) - \rho(k + 1) \frac{d}{d\hat{\boldsymbol{\theta}}(k)} V(k + 1)$$

- Izračun odvoda:

$$\frac{d}{d\hat{\boldsymbol{\theta}}(k)} V(k + 1) = e(k + 1) \frac{de(k + 1)}{d\hat{\boldsymbol{\theta}}(k)} = -e(k + 1) \boldsymbol{\Psi}(k + 1)$$



- Popravki parametrov se torej izvajajo v skladu s formulo:

$$\begin{aligned}\hat{\boldsymbol{\theta}}(k+1) &= \hat{\boldsymbol{\theta}}(k) + \rho(k+1)e(k+1)\boldsymbol{\psi}(k+1) \\ &= \hat{\boldsymbol{\theta}}(k) + \rho(k+1)\boldsymbol{\psi}(k+1)[y(k+1) - \boldsymbol{\psi}^T(k+1)\hat{\boldsymbol{\theta}}(k)]\end{aligned}$$

- Algoritem zelo podoben rekurzivni metodi najmanjših kvadratov, le da uporablja namesto kovariančne matrike $\mathbf{P}$ skalar ρ
- Metoda konvergira, če so izpolnjene naslednje zahteve za ρ :

$$\lim_{k \rightarrow \infty} \rho(k) = 0 \quad \sum_{k=1}^{\infty} \rho = \infty \quad \sum_{k=1}^{\infty} \rho^2(k) < \infty$$

- Možne izbire:

- $\rho(k) = \frac{1}{k}$
- $\rho(k)$ je enak sledi matrike $\mathbf{P}(k)$ in se izračunava rekurzivno:
$$\rho^{-1}(k+1) = \rho^{-1}(k) + \boldsymbol{\psi}^T(k+1)\boldsymbol{\psi}(k+1), \quad \rho^{-1}(0) = 0$$

- Metoda stohastične aproksimacije numerično enostavnejša od rekurzivne metode najmanjših kvadratov, a konvergira počasi



- Včasih želimo ocenjevati parametre procesov direktno v zveznem časovnem prostoru:
 - Z diskretizacijo namreč parametri prenosnih funkcij izgubijo svoj fizikalni pomen, kar otežuje interpretacijo in ovrednotenje procesnih parametrov (npr. časovnih konstant, lastnih frekvenc, koeficientov dušenja) – posebno pomembno, ko so zvezni procesi delno znani in ocenjujemo le nekaj neznanih parametrov
 - Druga prednost ocenjevanja parametrov direktno v časovnem prostoru je v tem, da odpade težavna izbira časa vzorčenja, saj ocenjeni parametri niso odvisni od časa vzorčenja (numerični problemi!); ni nujno, da je čas vzorčenja konstanten
- Ideja ocenjevanja parametrov zveznih procesov je enaka tisti za ocenjevanje parametrov diskretnih procesov: z opazovanjem vhoda in izhoda procesa tvorimo sistem enačb (običajno zelo predoločen), katerega rešitev je vektor procesnih parametrov



- Prenosna funkcija procesa:

$$G_p(s) = \frac{Y(s)}{U(s)} = \frac{b_0 + b_1s + \dots + b_ns^n}{1 + a_1s + \dots + a_ns^n} \quad a_n \neq 0$$

- Zapis modela sistema v obliki diferencialne enačbe:

$$\sum_{i=0}^n a_i \frac{d^i}{dt^i} y(t) = \sum_{i=0}^n b_i \frac{d^i}{dt^i} u(t) \quad a_0 = 1$$

- Izrazimo $y(t)$:

$$y(t) = -a_1 \frac{d}{dt} y(t) \dots - a_n \frac{d^n}{dt^n} y(t) + b_0 u(t) + b_1 \frac{d}{dt} u(t) + \dots + b_n \frac{d^n}{dt^n} u(t)$$

- Desno stran enačbe bi lahko zapisali kot produkt vektorja regresorjev (merjeni signali in njihovi odvodi) in vektorja parametrov
- Odvodi seveda niso merljivi – možno bi bilo tvoriti približke, vendar bi s tem zelo ojačili šum (poslabšanje rezultatov!)



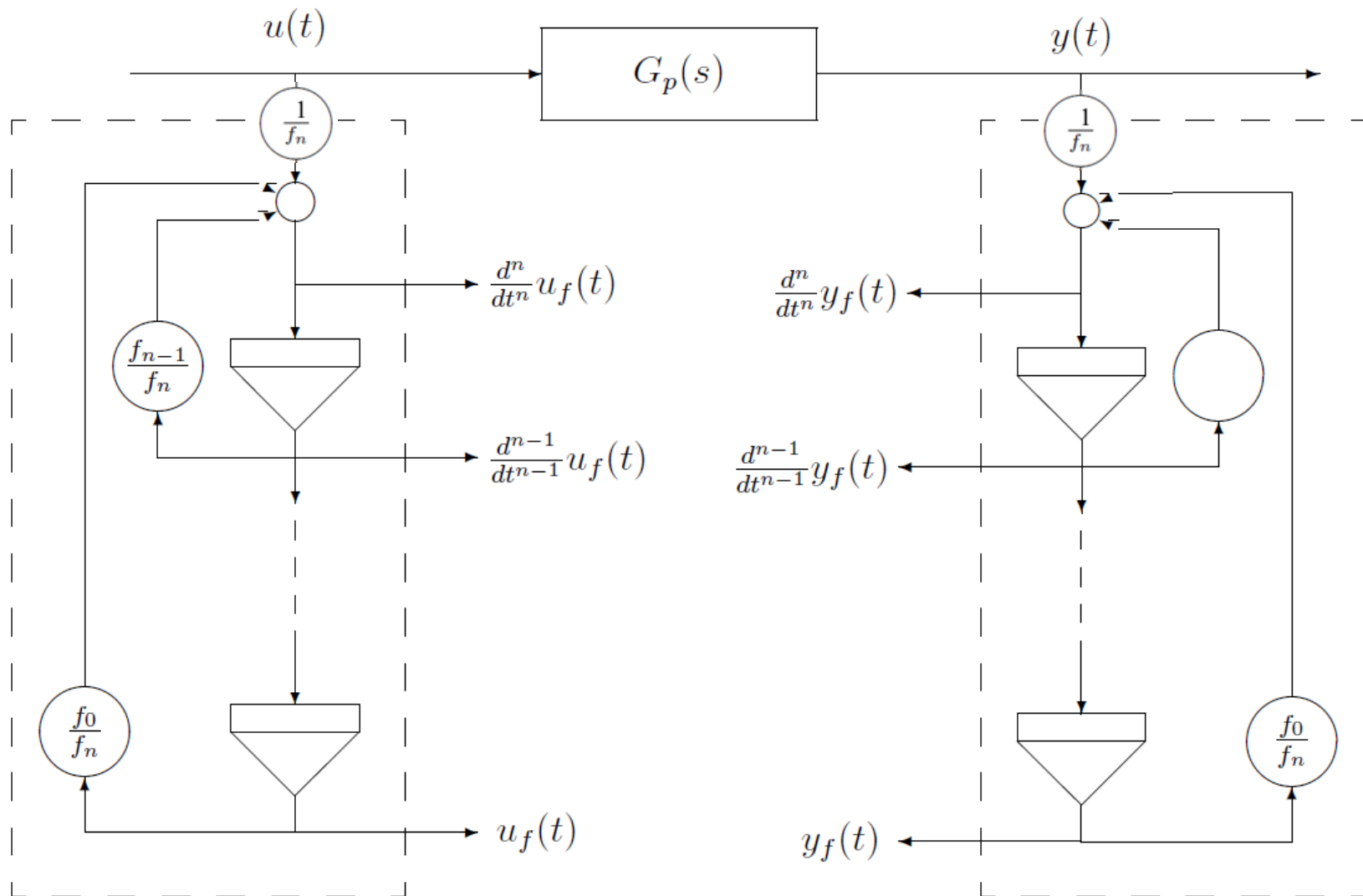
- Alternativa prej omenjeni metodi je uvedba *filtrov spremenljivk stanja*, s katerima filtriramo vhodni in izhodni signal
- Če vhod in izhod sistema filtriramo preko poljubnega filtra, veže filtrirani vhodni in filtrirani izhodni signal enaka prenosna funkcija kakor originalni vhodni in izhodni signal procesa
- Razlika med redom imenovalca in števca filtra mora biti vsaj tako velika, kakor je red prenosne funkcije ocenjevanega procesa
- Najenostavnejša rešitev je, da je red prenosne funkcije filtra $G_f(s)$ enak redu prenosne funkcije procesa, red števca pa je 0:

$$G_f(s) = \frac{1}{f_0 + f_1 s + \dots + f_n s^n} \quad f_n \neq 0$$

- Filtrirane vhodne in izhodne signale lahko izrazimo v obliki:

$$U_f(s) = G_f(s)U(s) \quad Y_f(s) = G_f(s)Y(s)$$

Ocenjevanje parametrov zveznih procesov





- Odvisnost med filtriranim vhodnim in filtriranim izhodnim signalom lahko izrazimo v naslednji obliki:

$$\frac{Y_f(s)}{U_f(s)} = \frac{G_f(s)Y(s)}{G_f(s)U(s)} = \frac{Y(s)}{U(s)} = G_p(s)$$

- Zapis v obliki diferencialne enačbe:

$$y_f(t) = -a_1 \frac{d}{dt} y_f(t) - \dots - a_n \frac{d^n}{dt^n} y_f(t) + b_0 u_f(t) + b_1 \frac{d}{dt} u_f(t) + \dots + b_n \frac{d^n}{dt^n} u_f(t)$$

- Diferencialna enačba po strukturi in parametrih enaka dif. enačbi originalnega sistema, a vsebuje filtrirane signale namesto originalnih

- Vpeljemo vektor neznanih parametrov in vektor merjenih signalov oz. regresorjev, ki jih dobimo v vezju filtra:

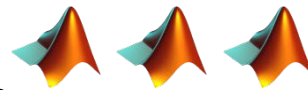
$$\begin{aligned} \boldsymbol{\theta}^T &= [a_1 \quad \dots \quad a_n \quad b_0 \quad b_1 \quad \dots \quad b_n] \\ \boldsymbol{\Psi}^T(t) &= \left[-\frac{d}{dt} y_f(t) \quad \dots \quad -\frac{d^n}{dt^n} y_f(t) \quad u_f(t) \quad \frac{d}{dt} u_f(t) \quad \dots \quad \frac{d^n}{dt^n} u_f(t) \right] \\ y_f(t) &= \boldsymbol{\Psi}^T(t) \boldsymbol{\theta} \end{aligned}$$



- Nadaljnji postopek je standarden postopek najmanjših kvadratov
- Pri nerekurzivni metodi tvorimo s pomočjo N opazovanj ob časih $t_1, t_2, \dots, t_N$ matriko Ψ (v i -ti vrsti ima $\Psi^T(t_i)$) in vektor $\mathbf{y}$ (v i -ti vrsti ima $y_f(t_i)$), iz katerih izračunamo rešitev:

$$\hat{\boldsymbol{\theta}} = [\Psi^T \Psi]^{-1} \Psi^T \mathbf{y}$$

- Seveda lahko algoritem zapišemo tudi v rekurzivni obliki
- Paziti je potrebno, da je filter $G_f(s)$ zvezni filter in bi ga morali realizirati na analogen način
- Za izvedbo filtra $G_f(s)$ lahko uporabimo tudi digitalno simulacijo zveznih sistemov, a moramo paziti na razliko med realizacijo vhodnega in izhodnega signala:
 - Vhodni signal je med dvema vzorcema običajno konstanten
 - Izhodni signal se med dvema vzorcema spreminja – če vzorčimo dovolj hitro, lahko predpostavimo linearno interpolacijo med vzorci





- Izbira vhodnega signala
 - Nekatere metode zahtevajo poseben tip vhodnega signala, kar velja v največji meri za neparametrične identifikacijske metode
 - Če želimo oceniti parametre procesa reda n (tipično $2n$ parametrov), mora vzbujanje vsebovati vsaj $2n$ različnih frekvenc (matematično zahteva izhaja iz zahteve po nesingularnosti matrike Ψ)
 - Skupno vsem metodam – vhodni signali dovolj frekvenčno bogati!
 - Amplituda – dobro razmerje signal/šum, vendar so velike amplitude problematične zaradi nelinearnosti (zapustimo linearno področje) in tehnoloških zahtev (velikih amplitud ne smemo uporabiti)
- Izbira časa vzorčenja
 - Prehitro vzorčenje – numerične težave!
 - Prepočasnno vzorčenje – izguba informacije – visoke frekvence
 - Običajno izberemo približno 10% dviznega časa (čas od 10% do 90% spremembe pri odzivu na stopničasto spremembo vhoda)



- Predhodna obdelava signalov:
 - Analogno filtriranje pred analogno-digitalnim pretvornikom (preprečimo zgibanje frekvenc – Shannon!)
 - Odstranitev enosmernih komponent in signalov lezenja (drift)
 - Zmanjšanje vpliva šuma z digitalnim filtriranjem (pomembno, da uporabimo za filtriranje vhodnega in izhodnega signala enak filter)
- Izbira modela:
 - Metode za ocenjevanje parametrov dajejo le najboljše možne parametre (glede na izbrani kriterij) pri izbrani strukturi modela, zato določitev strukture (pri linearnih sistemih le red) zelo pomembna:
 - Če izberemo prenizek red modela (pravimo, da model „podparametriziramo“), je tak model lahko zelo nenatančen
 - Če izberemo prevelik red modela (njegova „preparametrizacija“), imamo numerične težave (pogojenost), poleg tega na lego odvečnih parametrov vplivajo predvsem šum in motnje, kar vodi do napačnih rezultatov



- Izbira modela:
 - Bistvena pri izboru modela je dilema med *fleksibilnostjo* in *kompleksnostjo* modela:
 - Model mora biti dovolj fleksibilen, da lahko z njim dovolj dobro opišemo dejanski proces
 - Preveč kompleksen model vsebuje pole in ničle, ki se (vsaj aproksimativno) krajšajo in kot taki ne vplivajo na njegovo vhodno-izhodno obnašanje
 - Pri odločitvah si lahko pomagamo z:
 - apriornimi znanji o procesu oziroma z apriornim obdelovanjem podatkov (izvajamo operacije, ki dajejo vpogled v strukturo modela, ne da bi izračunali model v celoti):
 - določitev ranga informacijske matrike $\Psi^T \Psi$ (enak redu modela)
 - ocena neparametričnega modela (frekvenčnega odziva) – ocenimo red
 - aposteriornimi preizkusi dobljenega modela (tovrstni preskusi se ujemajo s preskusi za veljavnost modela –obravnavali bomo kasneje)



- Preskus veljavnosti modela ali **validacija**:
 - Običajno je rezultat identifikacijskega postopka več modelov – izberemo tistega, ki najbolje prestane preskus veljavnosti
 - Preskus veljavnosti modela pove, kako dobro opisuje model dejanski proces, ki ga identificiramo – izredno važen namen uporabe modela, saj je včasih važna natančnost, oziroma verodostojnost parametrov, včasih pa le ujemanje frekvenčnih odzivov modela in procesa
- Pri preskusu veljavnosti uporabljamo več postopkov:
 - Preskus pogreška modela – srednja vrednost mora biti 0, distribucija simetrična, nekoreliranost s preteklimi vhodi, avtokovariančna funkcija delta-impulz
 - Konsistenca vhodno-izhodnega obnašanja (na drugih podatkih!)
 - Konsistenca frekvenčnega obnašanja (po možnosti na drugih podatkih)
 - Smiselnost parametrov (še posebej, če imamo fizikalne parametre)
 - Redukcija modela (če reduciran model dovolj dober, ga izberemo)
 - Kovariančna matrika pogreškov parametrov (če velika, ocena nezanesljiva; žal ne velja, da majhna ocena pomeni, da je ocena dobra)
 - Uporabnost modela (če model služi svojemu namenu, je to prava ocena vrednosti – zadnji in odločujoči preskus njegove veljavnosti)



- Matlab\Uvod\Primer01.m
- Matlab\Uvod\Primer02.m
- Matlab\Uvod\Primer03.m
- Matlab\Uvod\Primer04.m
- Matlab\Uvod\Primer05.m
- Matlab\Uvod\Primer06.m
- Matlab\Uvod\y_je_konv_u_g.m
- Matlab\Strejc\Primer01_P_proces_.m
- Matlab\Strejc\Primer02_I_proces_.m
- Matlab\Prilagajanje\primer1.m
- Matlab\Prilagajanje\primer2.m
- Matlab\Prilagajanje\primer3.m
- Matlab\Prilagajanje\primer4_1.m
- Matlab\Prilagajanje\primer4_2.m
- Matlab\Sumi\Primer01_Distribucija.m
- Matlab\Sumi\Primer02_Avtokor_Belega.m
- Matlab\Sumi\Primer03_Avtokor_Barvnega_Gauss.m
- Matlab\Sumi\Primer04_Avtokor_Barvnega_enakom.m
- Matlab\Sumi\Primer05_Avtokor_Barvnega_binar.m
- Matlab\Sumi\Primer06_Poissonov_sum.m
- Matlab\Sumi\Primer07_Zvezni_NBS.m
- Matlab\Sumi\Primer08_Diskretni_NBS.m
- Matlab\Sumi\Primer09_Diskret_PNBS.m
- Matlab\Pistranskost\Primer01_Konsistenca_ocene_sr_vred.m
- Matlab\Pistranskost\Primer02_Pistranskost_ocene_var.m
- Matlab\MNK\Primer01_Skalarni_problem.m
- Matlab\MNK\Primer02_Vektorski_problem.m
- Matlab\MNK\Primer03_Bazne_funkcije.m
- Matlab\MNK\Primer04_Dinamicni_sis_nerekurzivno.m
- Matlab\MNK\Primer05_Dinamicni_sis_rekurzivno.m
- Matlab\MNK\Primer06_Lastne_vred_vekt_Singular_vred.m
- Matlab\MNK\Primer07_Faktor_pogojenosti.m
- Matlab\MNK_Mod\Primer01_RNK_beli_sum.m
- Matlab\MNK_Mod\Primer02_RNK_poz_real.m
- Matlab\MNK_Mod\Primer03_RNK_1_A.m
- Matlab\MNK_Mod\Primer04_NK_zvezno_majhen_sum.m
- Matlab\MNK_Mod\Primer05_NK_zvezno_sred_sum.m
- Matlab\MNK_Mod\Primer06_NK_zvezno_velik_sum.m