

Parametriset mallit

- Rakenteelliset mallit
 - parametreillä a priori tulkinta & merkitys
- Black box-mallit
 - parametrit vain laskennan/sovituksen apuvälineitä
- Tarkastellaan pääosin lineaarisia diskreettiaikaisia black-box-malleja
 - 3. harjoitustyössä malli epälineaarinen jatkuva-aikainen
- Tämän diasetin teemoja
 - malliluokat
 - parametrien estimointi - ennustevirhemenetelmä
 - mallin hyvyys vs. parametriestimaattien harhaisuus ja varianssi
 - systeemin / mallin identifioituvuus

Rakenteelliset mallit

- Fysikaalisin tai vastaavin perustein rakennettuja malleja
 - osa parametreista tunnettu: esim. massat, poikkipinta-alat jne...
 - osa estimoitava: kitkakertoimet,...
- Merkintä $\frac{d}{dx} x(t) = f(x(t), u(t), \theta); \hat{y}(t | \theta) = h(x(t), u(t), \theta)$

jossa $\hat{y}(t | \theta)$ on mallin ennustettu ulostulo hetkellä t ja parametrivektorilla θ
- Edellä oletetaan valkoinen mittauskohina $y(t) = h(x(t), u(t), \theta) + e(t)$
 - $E(y(t, \theta)) = \hat{y}(t, \theta)$
 - jos kohinalla on rakennetta, se kannattaa huomioida, esim. ARMA-malli
- Ennustevirhe $\varepsilon(t) = y(t) - \hat{y}(t, \theta)$ (residuaali = ennustevirhe numerona)
- Estimointiongelma: etsi θ s.e. $\min V_N(\theta) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \theta)$

Black box -mallit

- Yleinen lineaarinen diskreettiaikainen malli on muotoa $y(t) = \eta(t) + w(t)$
 - $w(t)$ = häiriötermi
 - $\eta(t)$ = häiriötön ulostulo
- Termit muotoa $\eta(t) = G(q, \theta)u(t)$, $w(t) = H(q, \theta)e(t)$
 - $G()$ ja $H()$ lineaarisia suotimia = diskreettiaikaisia siirtofunktioita
 - käytännössä rationaalisia (muotoa polynomi/polynomi) ja (asymptoottisesti) stabiileja (eräs validointitarkastelu)
 - $G(q, \theta) = B(q, \theta)/F(q, \theta)$, $H(q, \theta) = C(q, \theta)/D(q, \theta)$
 - parametrivektori θ koostuu polynomien B, F, C ja D kertoimista b_i, f_i, h_i, d_i ja kohinan $e(t)$ varianssista
 - ”rakenneparametrit” n_b, n_c, n_d, n_f ja n_k (kuollut aika)

Transfer function parameterizations

The transfer functions $G(q)$ and $H(q)$ in the linear model

$$y[k] = G(q; \theta)u[k] + H(q; \theta)e[k]$$

will be parameterized as

$$G(q; \theta) = q^{-n_k} \frac{b_0 + b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}}$$
$$H(q; \theta) = \frac{1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}}{1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d}}$$

where the parameter vector θ contains the coefficients $\{b_k\}$, $\{f_k\}$, $\{c_k\}$, $\{d_k\}$

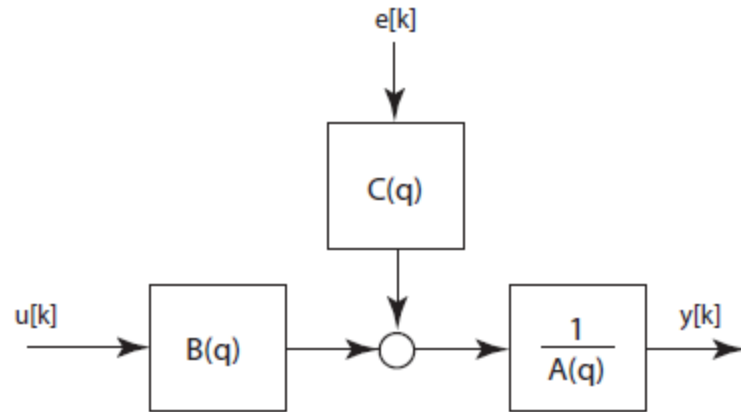
Note: n_k determines dead-time, n_b, n_f, n_c, n_d order of transfer function polynomials.

Erilaisia mallirakenteita

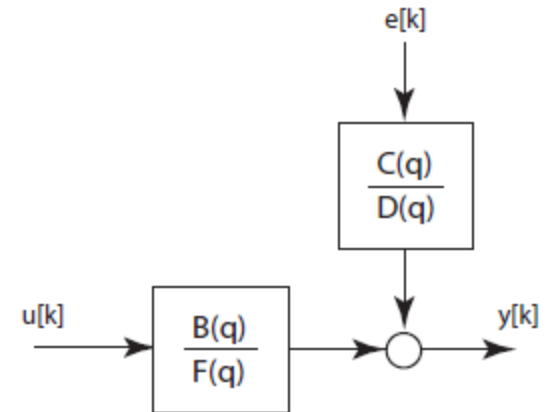
- Box-Jenkins (BJ): $G(q)=B(q)/F(q)$, $H(q)=C(q)/D(q)$
 - ”täydellinen” malli
 - systeemin ja kohinan mallit riippumattomia toisistaan
- Output error (OE): $H(q)=1$ eli $n_c=n_d=0$
 - kohina valkoista $w(t)=e(t)$, eli ei erityistä rakennetta
- ARMAX: $A(q)y(t)=B(q)u(t)+C(q)e(t)$
 - kohina kokee saman dynamiikan kuin u
 - järkevää kun kohina tulee prosessiin sen alkupäässä
- ARX: $A(q)y(t)=B(q)u(t)+e(t)$ (myös *yhtälövirhemalli*)
 - ennustevirhe lineaarinen parametrien suhteen \Rightarrow parametrien estimointi ”helppoa” – tavallinen PNS
 - kohina kokee saman dynamiikan kuin u ; ei haittaa, jos signaali-kohina-suhde on hyvä
- Käyttö: määrää asteluvut, estimoi parametrit
 - käytännössä useita eri mallirakenteita ja astelukuja verrataan

Mallirakenteita

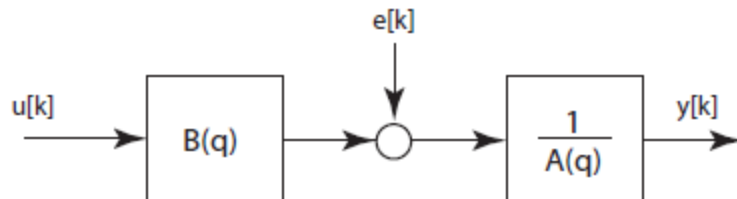
ARMAX (autoregressive moving average
exogeneous input)



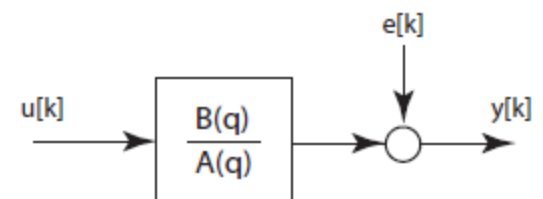
BJ (Box Jenkins)



ARX (autoregressive with exogeneous input)



OE (output error)



Mallin tuottama ennuste $\hat{y}(t)$

- Usein mallin käyttötarkoitus ennustaminen – tarvitaan myös ennustevirheen neliösummassa
- Millainen on kahdella edellisellä dialla esitettyjen mallien ennuste?
 - Malli muotoon $y(t) = \text{JOTAIN} + e(t) \Rightarrow \hat{y}(t) = E(y(t)) = \text{JOTAIN}$, koska $E(e(t)) = 0$
- OE: $\hat{y}(t, \theta) = G(q, \theta)u(t)$
- ARX: $\hat{y}(t, \theta) = (1 - A(q))y(t) + B(q)u(t)$

Consider the linear model

$$y[k] = G(q)u[k] + H(q)e[k]$$

Multiply by $H^{-1}(q)$ (to make noise term white) and re-write as

$$y[k] = (1 - H^{-1}(q))y[k] + H^{-1}(q)G(q)u[k] + e[k]$$

Since $\{e\}$ is a white noise sequence, our best prediction is

$$\hat{y}[k] = (1 - H^{-1}(q))y[k] + H^{-1}(q)G(q)u[k]$$

Mallien sovitus

- Ennustevirhemenetelmät: valitse θ s.e. ennuste on ”hyvä”
 - toimiva hyvyyskriteeri ennustevirheen (otos)variانسsi
 - ennustevirhe $\varepsilon(t) = y(t) - \hat{y}(t, \theta)$ ja
hyvyyskriteeri
$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \theta)$$
- $\hat{\theta}_N = \arg \min V_N(\theta)$
 - voidaan osoittaa, että $V_N(\theta)$ ei ole tuulesta temmattu
 - PNS on ennustevirhemenetelmien erikoistapaus
 - MIMO-mallit: $\varepsilon^2(t)$ matriisiarvoinen, tarvitaan sopiva reaaliarvoinen kuvaus: det, trace, ...
- $V_N(\hat{\theta})$ on kohinan $e(t)$ varianssin estimaatti

Lineaarinen regressio

- Regressio: malli muotoa $y(t) = \theta^T \phi(t) + e(t)$, $\phi(t)$ vektori, jossa viivästettyjä $u(t)$:n ja $y(t)$:n arvoja
 - θ sisältää viiveoperaattoripolynomien kertoimet
- N mittausta \Rightarrow matriisi $X = [\phi(1), \phi(2), \dots, \phi(N)]'$ sekä ulostulo $Y(t) = [y(1), \dots, y(N)]' \Rightarrow \hat{\theta} = (X^T X)^{-1} X^T Y$
- Soveltuu ARX-mallien parametrien estimointiin
 - muut malliluokat epälineaarisia parametrien suhteen
- PNS-oletuksien syytä olla voimassa:
 - viivästetyt y :n ja u :n arvot keskenään kollineaarisia (koesuunnitteluongelma) \Leftrightarrow parametrit tehottomia (suuri varianssi) ja harhaisia, mutta tarkentuvia
 - e :n homoskedastisuus (vakio varianssi) \Leftrightarrow parametrit tehottomia \Rightarrow painotettu PNS-estimointi
 - e :n korreloimattomuus \Leftrightarrow parametrit tehottomia, harhaisia eivätkä tarkentuvia! \Rightarrow kohinalle jonkinlainen rakenne

Iteratiivinen minimointi

- Mallit epälineaarisia parametrien suhteen
 - PNS ei onnistu
- Tarvitaan iteratiivinen menetelmä
 - kyseessä epälineaarinen rajoittamaton optimointitehtävä
- Gradienttimenetelmä: $\theta^{k+1} = \theta^k - \alpha^k V_N'(\theta^k)$
 - α^k viivahaun avulla haettava askelpituus
- Toisen kertaluvun menetelmät: sovelletaan Newton-iterointia optimiratkaisun välttämättömiin ehtoihin
 - välttämätön ehto $V_N'(\theta) = 0$
 - iterointi $\theta^{k+1} = \theta^k - \alpha^k [V_N''(\theta^k)]^{-1} V_N'(\theta^k)$
- Tarvitaan $V(\theta)$:n eli $\varepsilon(t)$:n derivaatat parametrien suhteen
 - riippuvat mallirakenteesta; laskari #9, tehtävä #2

Mikä on hyvä malli?

- Mallin laatu on yhteydessä mallin käyttötarkoitukseen
 - hyvä säätösuunnittelumalli voi olla huono simulointimalli
- Mallin laatu liittyy sen kykyyn kuvata systeemin toiminta
 - systeemin ja mallin ulostulot riittävän samanlaiset
- Hyvä malli yleistää, i.e., mallin ominaisuudet ”eivät riipu estimointidatasta” - mallin tilastolliset ominaisuudet
 - parametrien varianssi
 - varianssia voidaan pienentää kasvattamalla havaintojen lukumäärää
 - mallin ”variance error”
 - Estimaattien konvergenssi väriin arvoihin
 - väärä mallirakenne / puutteellinen koesuunnittelu
 - erotettava mallirakenteen vaikutus ja identifiointikokeen vaikutus
 - mallin ”bias error”

Parametristimaattien harhasta

- Mitä tapahtuu, kun $N \rightarrow \infty$?
- Olkoon ennustevirheen varianssi $E\varepsilon^2(t, \theta) = V(\theta)$
- Jos ennustevirhe $\varepsilon(t, \theta)$ on valkoista kohinaa, niin

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^N \varepsilon^2(t, \theta) \xrightarrow{N \rightarrow \infty} E\varepsilon^2(t, \theta) = V(\theta) \text{ w.p.1}$$

ja lisäksi $\theta \xrightarrow{N \rightarrow \infty} \theta^* = \arg \min V(\theta)$

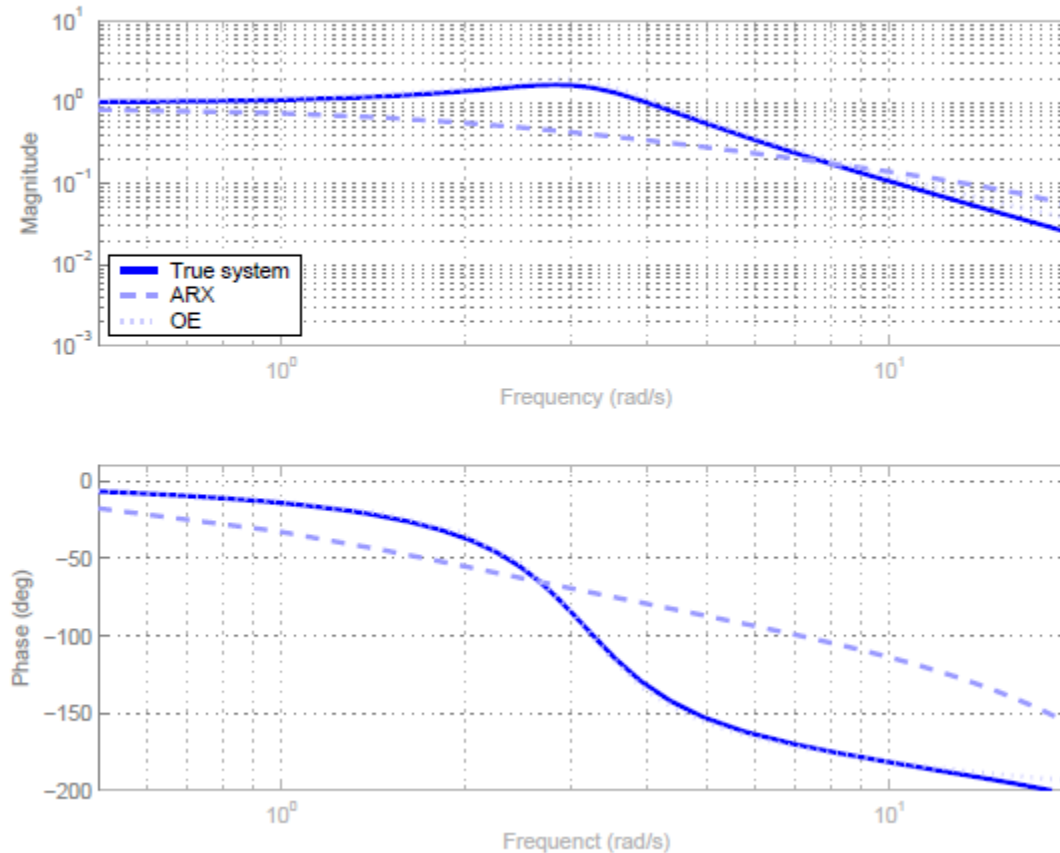
- Jos ennustevirhe ei ole korreloimaton (esim. väärä malli), niin estimaatti on harhainen ja konvergoituu väärään arvoon vs. ”oikean” mallin estimaatin arvo
- Estimaatti minimoi edelleen ennustevirheen varianssin
 - ”paras” malli väärästä malliluokasta

Esimerkki: väärä mallirakenne

- Olkoon data peräisin ARMAX-prosessista
$$y(t) + a_0 y(t-1) = b_0 u(t-1) + e_0(t) + c_0 e_0(t-1)$$
- Sovitetaan ARX-malli $\hat{y}(t | \theta) + a y(t-1) = b u(t-1)$
- Ennustevirheen varianssi on $E(y(t) + a y(t-1) - b u(t-1))^2 = \dots = r_0(1 + a^2 - 2aa_0) + b^2 - 2bb_0 + 2ac_0$
 - ($r_0 = E y^2(t)$, ei riipu a :sta eikä b :stä)
- Ennustevirheen varianssin minimoivat \hat{a} , \hat{b} : $\hat{a} = a_0 - c_0 / r_0$; $\hat{b} = b_0$;
 - $E \hat{a} \neq a_0$ eli estimaatti on harhainen
- Ennustevirheen varianssi näillä \hat{a} , \hat{b} on $1 + c_0^2 - c_0^2 / r_0$
 - todellisilla parametriarvoilla a_0 ja b_0 varianssi on $1 + c_0^2$ eli suurempi
- Siis: vaikka parametriestimaatti on harhainen, se tuottaa pienemmän ennustevirheen varianssin

Model structure matters!

Example: $G(s) = 10/(s^2 + 2s + 10)$, sampling time $h = 0.05$, noise variance 0.1^2



Parametriestimaatin konvergenssi taajuustasossa

Assume that the true system is described by

$$y[k] = G_0(q)u[k] + w[k]$$

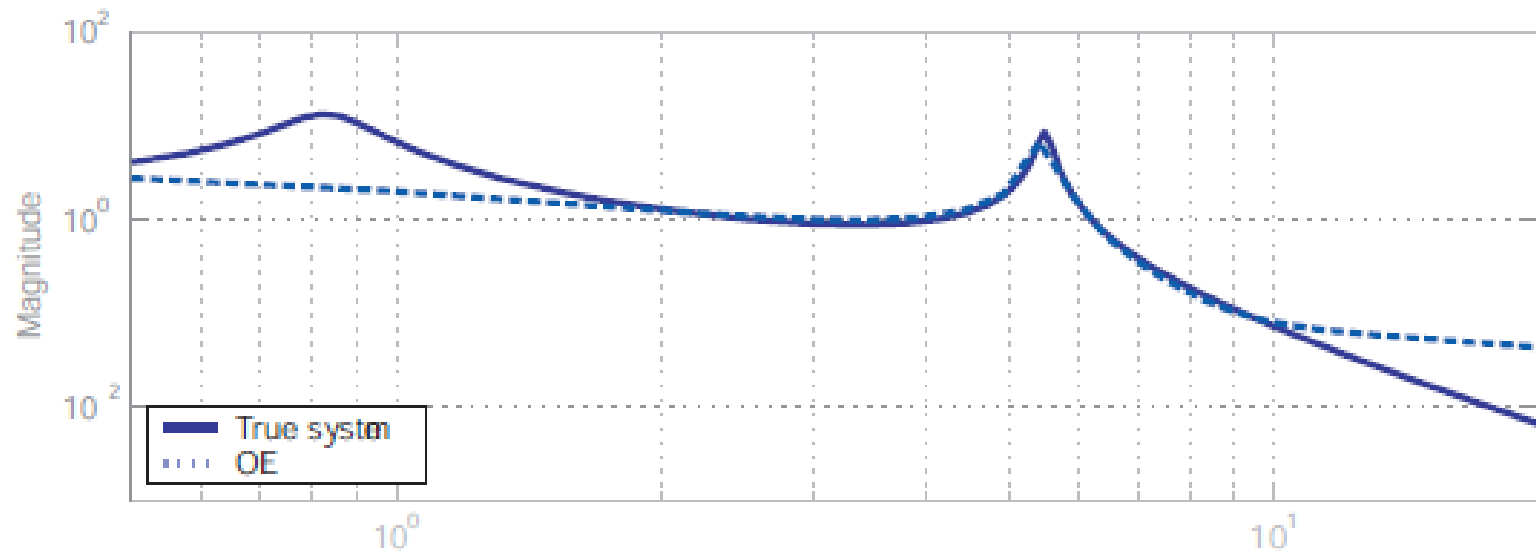
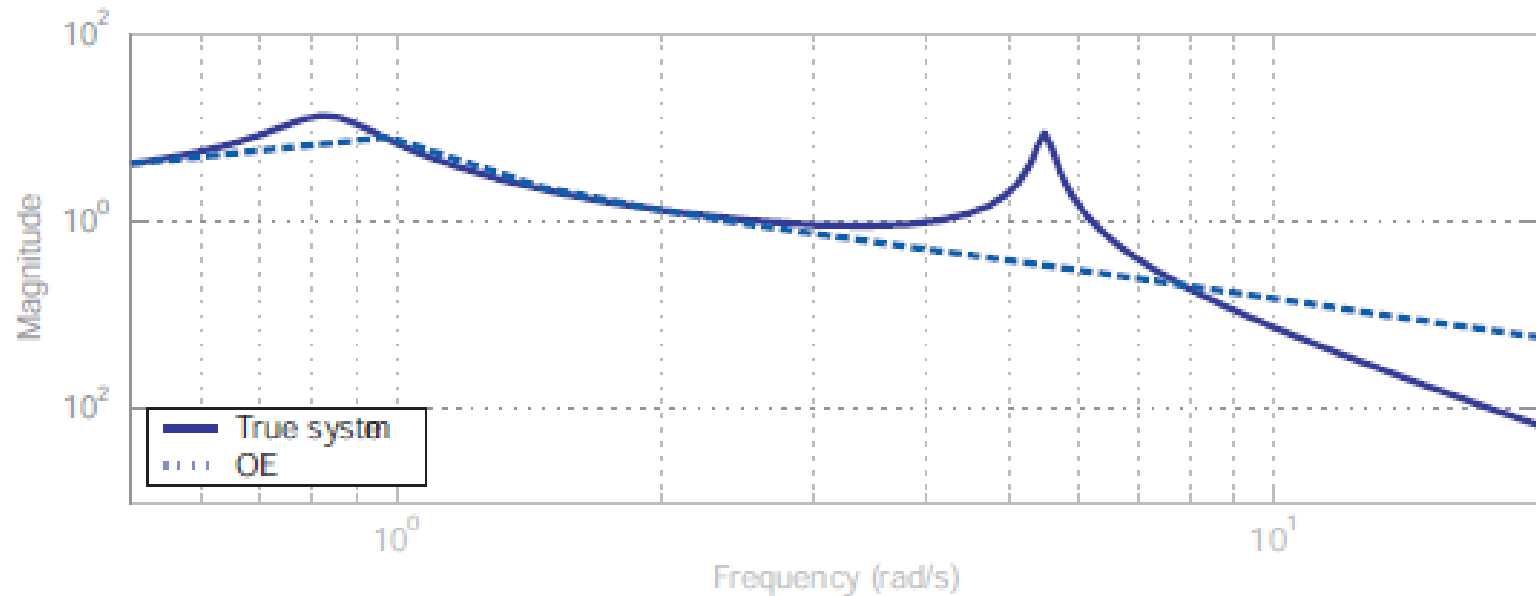
and that we try to estimate a model on the form ($H_*(q)$ independent of θ)

$$y[k] = G(q; \theta)u[k] + H_*(q)e[k]$$

$$\theta^* = \lim_{N \rightarrow \infty} \hat{\theta}_N = \arg \min \int_{-\pi}^{\pi} |G_0(e^{i\omega}) - G(e^{i\omega}, \theta)|^2 \frac{\Phi_u(\omega)}{|H_*(e^{i\omega})|^2} d\omega$$

- $V_N(\theta)$:n minimoiva estimaatti lähestyy arvoa, joka saa mallin taajuusvasteen mahdollisimman lähelle systeemin taajuusvastetta painotettuna
 - ohjauksen spektrillä
 - kohinamallin taajuusvasteen käänteisluvulla
- Koesuunnittelu: valitaan u :n taajuusominaisuudet sopivasti \Rightarrow hyvä sovitus mielenkiintoisilla taajuuksilla

Output error model using low- and high-frequency input signal



Parametriestimaattien varianssi

- Oletetaan, että ennustevirhe $\varepsilon(t, \theta)$ on valkoista kohinaa
- Tällöin parametriestimaatin $\hat{\theta}_N$ kovarianssimatriisi P_N on

$$P_N = E(\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T \approx \frac{1}{N} \lambda \bar{R}^{-1}$$

jossa $R = \psi(t, \theta_0)\psi^T(t, \theta_0)$ ja $\psi(t, \theta) = d/d\theta y^{\wedge}(t, \theta)$

- P_N riippuu
 - kohinan varianssista
 - datapisteiden lukumäärästä
 - ennusteen gradientista (herkkyys!)
- Parametriestimaatit asympotoottisesti normaalijakautuneita
 - tilastollisen merkitsevyyden testaus t-testillä
- Taajuusvasteen modulin varianssi annetulla parametriestimaatilla riippuu
 - parametrien lukumäärästä ja kohinan spektristä
 - datapisteiden lukumäärästä ja ohjauksen spektristä

$$\text{Var} \{G(e^{i\omega}; \theta)\} \approx \frac{n}{N} \frac{\Phi_w(\omega)}{\Phi_u(\omega)}$$

Identifioituvuus

- Identifioituvuus: Voidaanko systeemin/mallin parametrit määrätä yksikäsitteisesti input-output –datasta?
- Olkoon θ_0 hyvyyskriteerin minimoiva parametriestimaatti
- Malli on identifioituva: $y^\wedge(t | \theta_0) = y^\wedge(t | \theta) \iff \theta = \theta_0$ (I)
- Milloin (I) ei päde?
 - 1) kaksi erilaista parametrivektoria tuottaa samanlaisen mallin input-output –käyttäytymiseen
 - rakenteellinen identifioituvuus, ymmärretään systeemin ominaisuutena, vrt. harjoitustyö #3
 - 2) kaksi erilaista parametrivektoria tuottaa erilaisen input-output-käyttäytymiseen ”hyvällä” inputilla, mutta puutteellinen input aiheuttaa samanlaiset ennusteet
 - ”deterministinen identifioituvuus”, ks. laskari #10, tehtävät #1 & #2

Esimerkki: tasavirtamoottori

- Valitaan tiloiksi kulma-asento $y(t)$ ja $-$ nopeus $\omega(t)$ ja ohjaukseksi jännite $u(t)$:

$$\frac{d}{dt} x(t) = \begin{bmatrix} 0 & 1 \\ 0 & -1/\tau \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ \beta/\tau \end{bmatrix} u(t),$$

- Tässä

$$\tau = \frac{JR_1}{fR_1 + k^2}, \quad \beta = \frac{k}{fR_1 + k^2}$$

- Systeemissä on 5 parametriä, mutta mallissa vain 2
 - vaikka mallin parametrit saataisiin estimoitua, niistä saadaan vain 2 yhteyttä systemin parametrien välille
 - tällä parametroinnilla identifiointi ei onnistu – systeemi ei rakeenteellisesti identifioituva

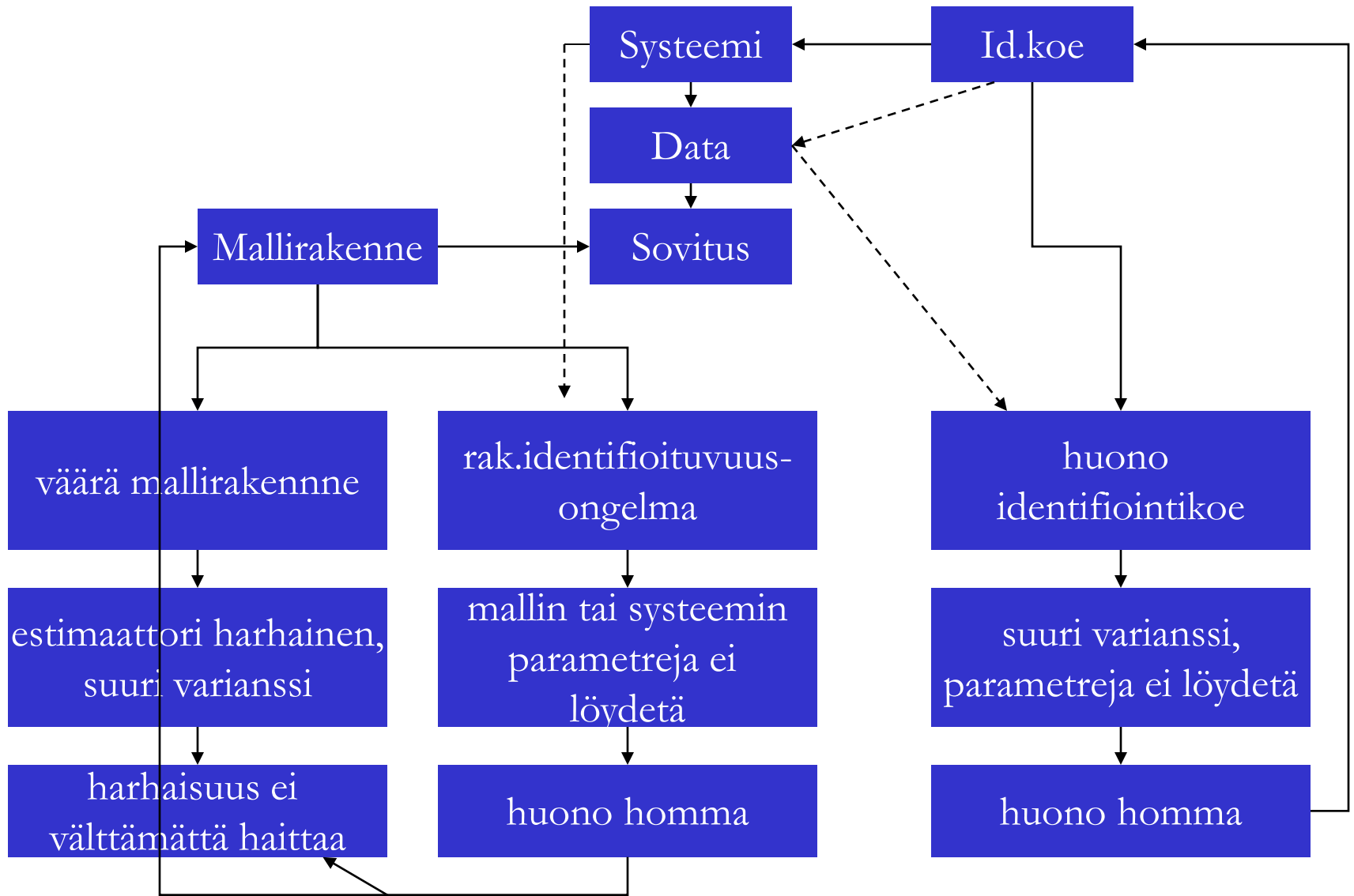
Vaatimukset herätteelle

- Intuitiivisia tuloksia
 - identifiointikokeen pitäisi herättää systeemin mielenkiintoiset toimintatilat
 - sisäänmenon taajuussisältö oleellisessa asemassa
- Jatkuvasti herättävyyden käsite – tarkastellaan diasetillä #10
 - kvantitatiivisia tuloksia herätteen laadun ja parametriestimoinnin onnistumisen välille

Esimerkki

- ARX-malli $\hat{y}(t | \theta) = au(t-1) + bu(t-2)$, $\theta = (a \ b)'$
 - valitaan u vakio-ohjaus u_0 :ksi
- Todellinen ennuste on tällöin $\hat{y}(t | \theta) = (a+b)u_0(t-1)$
 - kaikki parametrit a ja b , joiden summa on sama, antavat saman ennusteen
- Malli (= parametrit a ja b) ei ole identifioituvia (tällä herätteellä)
 - malli on kuitenkin rakenteellisesti identifioituva

Estimoinnin ongelmalähteet



Yhteenveto

- Perusidea: haetaan parametrivektori, joka minimoi ennustevirheen varianssin
 - sitä kuvaa neliöllinen hyvyyskriteeri $V_N(\theta)$
 - sovitus: minimoi ennustevirheen varianssi
 - estimaatin varianssi riippuu kohinan varianssista, datan määrästä ja ennusteen herkkyydestä parametrin suhteen
- Lähestymistavan etuja
 - yleinen käytettävyys, monipuolisuus
 - tuloksena simulointiin soveltuva malli
- Haittoja, muttei välttämättä ylitsepääsemättömiä
 - tarvitaan näkemys systeemin rakenteesta
 - laskentatuki tarpeen