# Parameterskattning i linjära dynamiska modeller

Kap 12

## Grundläggande ansats

- Antag (samplade) mätdata ($y$ och $u$)från ett system har insamlats.
- Givet en modell $\mathcal{M}(t, \theta)$ och mätdata , hitta det $\theta$ som ger en så bra (i mistakvadratfelsmening) **prediktion** av systemet som möjligt. ”Prediktionsfelsminimering”.
  Behöver konstruera en prediktor från modellen $\mathcal{M}(t, \theta)$, hur den ser ut beror på hur bruset modelleras.

  Beteckning:
  $\hat{y}(t|\theta)$= prediktion av $y(t)$ givet data upp till $t-1$.

- Skräddarsydda modeller (Grey-box modeller):
  Skatta systemparametrar i fysikalisk modell.
- Konfektionsmodeller (Black-box modeller):
  Parametrar har generellt ingen fysikalisk tolkning.
  Typisk skattas parametrar i en "allmän"
  differensekvation.

## Skräddarsydda modeller (12.1)

- Modellera okända parametrar i fysikalisk modell med parameter $\theta$:

$$\dot{x}(t) = f(x(t), u(t), \theta)$$
$$y(t) = h(x(t), u(t), \theta)$$

- Använd mätningar till att skatta $\theta$. Tillvägagångsätt problemberoende.
- Mera om skräddarsydda modeller senare i kursen.

## Linjära konfektionsmodeller (12.2)

- Olika klasser av konfektionsmodeller med störterm $w(t)$ och störningsfri insignal $\eta(t)$:

$$y(t) = \eta(t) + w(t)$$

- Box-Jenkins

$$\begin{aligned}
y(t) &= G(q,\theta)u(t) + H(q,\theta)e(t) \\
G(q,\theta) &= \frac{B(q)}{F(q)} \\
&= \frac{b_1 q^{-nk} + \ldots + b_{nb} q^{-nk-nb+1}}{1 + f_1 q^{-1} + \ldots + f_{nf} q^{-nf}} \\
H(q,\theta) &= \frac{C(q)}{D(q)} = \frac{1 + c_1 q^{-1} + \ldots + c_{nc} q^{-nc}}{1 + d_1 q^{-1} + \ldots + d_{nd} q^{-nd}}
\end{aligned}$$

där $e(t)$=vitt brus. Notera att insignalen $u(k)$ är tidsfördröjd $nk$ sampel.

- Output-Error (OE). Modellerar ej störsignalens egenskaper.

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

- FIR (specialfall av OE):

$$y(t) = B(q)u(t) + e(t)$$

- ARMAX. Antar $F(q) = D(q) = A(q)$:

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

- **ARX**. Sätt $C(q) = 1$ i ARMAX modellen:

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

Även rena tidsseriemodeller (bara en signal $y$) kan erhållas:

- ARMA ($B = F = 0$):

  $$A(q)y(t) = C(q)e(t)$$

  $C(q) = 1$ ger AR, $A(q) = 1$ ger MA.

- Modellerna används på så sätt att man anger ordningstalen $na, nb, nc, nd, nf$ och $nk$ och beräknar de parametrar som bäst beskriver data.

- Val av modell mm i Kap 14.

## Anpassning av modeller till data (12.3)

- Givet en parametrisk modell och gamla mätdata $\{y(s), u(s)\}_{s=\leq t-1}$ kan man prediktera värdet på utsignalen vid tiden $t$. Prediktionen betecknas $\hat{y}(t|\theta)$.
- Prediktionsfelet definieras som:

$$\epsilon(t, \theta) = y(t) - \hat{y}(t|\theta)$$

- Ide: Välj den modell $\theta$ som minimerar prediktionsfelens varians.

- $\Rightarrow$ Givet mätdata $\{y(t), u(t)\}_{t=1,\ldots,N}$ bilda förlustfunktionen:

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \epsilon^2(t, \theta)$$

och beräkna det $\theta$ som minimierar $V_N(\theta)$:

$$\hat{\theta}_N = \arg\min_{\theta} V_N(\theta)$$

- Om störningarna är Gaussiska så är ovanstående maximum-likelihood (ML) skattningen av $\theta$.

## Prediktion

- Prediktor för generell modell (Box-Jenkins):

$$\hat{y}(t|\theta) = [1 - H^{-1}(q,\theta)]y(t) + H^{-1}(q,\theta)G(q,\theta)u(t)$$

- Prediktor för Output-error modell:

$$\hat{y}(t|\theta) = \frac{B(q,\theta)}{A(q,\theta)}u(t)$$

- Prediktor för ARX-modell:

$$\hat{y}(t|\theta) = [1 - A(q,\theta)]y(t) + B(q,\theta)u(t)$$

## Linjär regression - jmfr F2!

- Minimeringen av förlustfunktionen är speciellt enkel om prediktionen kan skrivas som en linjär funktion av $\theta$.
- För en ARX modell kan prediktionen skrivas som:

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

där

$$\begin{aligned} \theta &= [a_1 \cdots a_{na} \; b_1 \cdots b_{nb}]^T \\ \varphi(t) &= [-y(t-1) \cdots - y(t-na) \\ & \quad u(t-nk) \cdots u(t-nk-nb+1)]^T \end{aligned}$$

- Motsvarande prediktionsfel blir:

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

- Förlustfunktionen $V_N(\theta)$ minimimeras i detta fall av:

$$\hat{\theta}_N = R_N^{-1} f_N$$

där

$$\begin{aligned} f_N &= \frac{1}{N} \sum_{t=1}^{N} \varphi(t) y(t) \\ R_N &= \frac{1}{N} \sum_{t=1}^{N} \varphi(t) \varphi^T(t) \end{aligned}$$

- Notera att (för ARX modellen) består elementen i $f_N$ och $R_N$ alla av skattningar av olika kovariansfunktioner för $y$ och $u$.

- Vid beräkning av skattningen skall invertering av $R_N$ undvikas av numeriska skäl.

## Iterativ sökning

- För många modellstrukturer (B-J, OE, ARMAX...) är funktionen $V_N(\theta)$ en komplicerad funktion av $\theta$. Då måste numerisk sökning användas.

  Ex: Newton-Raphsons metod:

  $$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - \mu_i [V_N''(\hat{\theta}^{(i)})]^{-1} V_N'(\hat{\theta}^{(i)})$$

  Se boken (289-290) för detaljer.

- Det finns färdiga rutiner i Matlab som gör minimeringen.

  `theta='metod'(data,modellordning)` men även ett grafiskt gränssnitt (Blabb 2)

## Modellens egenskaper (12.4)

- Hur bra blir modellskattningar med prediktionsfelsmetoden?
- Modellens **variansfel** beror på att mätningarna och systemet påverkas av brus. Identiska experiment ger aldrig likadana resultat.
- Modellens **biasfel** beror på att modellen ej fullständigt kan beskriva systemet. Biasfelet kavarstår även om mätningarna skulle kunna göras utan inverkan av brus.
- En bra modell har både litet variansfel och biasfel.

## Biasfelet

Antag ett sannt system

$$y(t) = G_o(q)u(t) + w(t)$$

samt en linjär modell med $\theta$-oberoende brusmodell

$$y(t) = G(q,\theta)u(t) + H_*(q)e(t)$$

då gäller

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

## Variansfelet

- Antag att biasfelet = 0. Då gäller:

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

  där

$$\begin{aligned} \overline{R} &= E\Psi(t,\theta_0)\Psi(t,\theta_0)^T \\ \Psi(t,\theta_0) &= \frac{d}{d\theta}\hat{y}(t|\theta) \end{aligned}$$

- OBS. För ARX modell är $\Psi(t,\theta_0) = \varphi(t)$
- Formell beskrivning (sid 299):

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \in AsN(0, \lambda\overline{R}^{-1})$$

- Konfidensintervall för den skattade modellen kan beräknas!
- I frekvensplanet gäller approximativt:

$$\mathrm{var}(G(e^{i\omega}, \hat{\theta}_N)) \approx \frac{n}{N} \frac{\Phi_w(\omega)}{\Phi_u(\omega)}$$

där $n$ är modellens ordningstal. Notera att biasfelet antas vara noll.

## Identifierbarhet

- Olika värden på $\theta$ bör ge olika prediktioner $\hat{y}(t, \theta)$.
- Identifierbarhet: parametrarna kan bestämas från mätdata.
- Icke-identifierbarhet kan bero på:
  - Parametriseringen
  - Insignalen
  - Återkoppling