

Identifikáció

Szabályozástechnika előadás

Harmati István, Kiss Bálint

Budapesti Műszaki és Gazdaságtudományi Egyetem
Irányítástechnika és Informatika Tanszék

Budapest, 2020. május 4.



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszabványai
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modell esetén



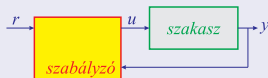
Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszabványai
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



A visszacsatolásról

Visszacsatolás



Alapvető eszköz

A kívánt irányba módosítja a szakasz dinamikáját zárt körben.

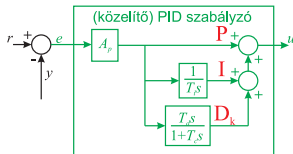
Korlátozások

- 1 modell megbízhatósága
- 2 beavatkozó jel korlátos

Lehetőségeink (ebben a félévben)

- 1 statikus tulajdonságok beállítása
- 2 tranziensek beállítása (domináns pólus, fázistartalék, sáv szélesség)
- 3 labilis szakasz stabilizálása
- 4 belső jelek megfigyelése
- 5 kis és nagy holtidő kezelése
- 6 véges impulzusválasz elérése (diszkrét időben)
- 7 zárt kör pólusainak áthelyezése
- 8 zavar elnyomása vagy becslése

Visszacsatolások - leltár



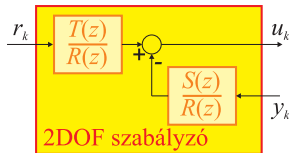
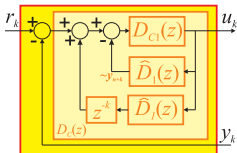
soros PID

- 1 fázistartalék beállítása
- 2 sáv szélesség növelése
- 3 típuszám növelése
- 4 paraméterek beállítása egyenletekből (grafikus megoldás vagy fsolve)

soros PID mintavételesen

- 1 mint a PID
- 2 egyszerű algoritmus a számítógépen futtatható
- 3 mintavételezés rontja a fázistöbbletet
- 4 T a méretezésből adódó ω_c alapján

Visszacsatolások - leltár



Smith prediktor

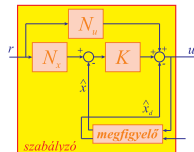
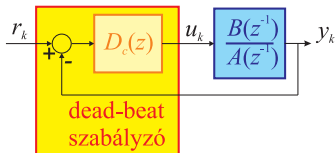
- 1 futtatja a szakasz modelljét
- 2 „nagy” holtidő esetére
- 3 $D_c(z)$ méretezése a holtidő nélküli szakaszhoz

2DOF szabályzó

- 1 modellalapú tervezés
- 2 diophantoszi egyenlet megoldása kell hozzá



Visszacsatolások - leltár



Véges beállítású szabályozás

- 1 speciálisan méretezett, soros mintavételes szabályzó
- 2 zárt körben FIR
- 3 figyelni kell T megválasztására

Állapotteres szabályzó

- 1 pólusáthelyezés
- 2 megfigyelő
- 3 opciókkal



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés**
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszerepei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Bevezetés



Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- Indirekt: Modell megalkotása és szabályozótervezés a a modellre

Megoldandó feladat

Identifikáció: A modell megalkotása

Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- Indirekt: Modell megalkotása és szabályozótervezés a a modellre

Megoldandó feladat

Identifikáció: A modell megalkotása

Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- Indirekt: Modell megalkotása és szabályozótervezés a a modellre

Megoldandó feladat

Identifikáció: A modell megalkotása

Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- Indirekt: Modell megalkotása és szabályozótervezés a a modellre

Megoldandó feladat

Identifikáció: A modell megalkotása

Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- **Indirekt: Modell megalkotása és szabályozótervezés a a modellre**

Megoldandó feladat

Identifikáció: A modell megalkotása

Motiváció

Probléma

A szabályozó tervezéséhez szükség van a szakasz modelljére, de az sok esetben nem áll rendelkezésre!

Megoldás módszere

Adaptív szabályozás

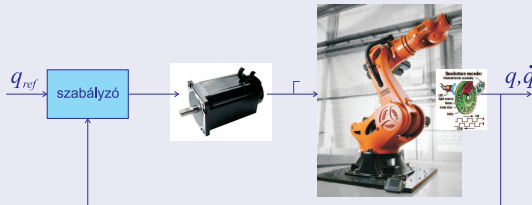
- Direkt: A beavatkozó jel hangolása közvetlenül a szabályozási hiba alapján történik
- Indirekt: Modell megalkotása és szabályozótervezés a a modellre

Megoldandó feladat

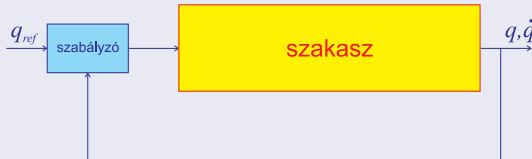
Identifikáció: A modell megalkotása

Megoldandó feladat

Eljutni innen



Ide



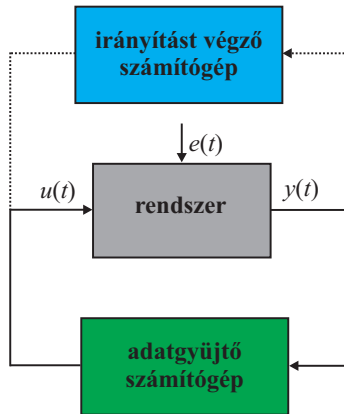
Az identifikáció sémája

Ismert (mért)

- Bemenet (u)
- Kimenet (y) [e zajjal terhelt]

Ismeretlen

- Zaj (e)
- A rendszer paraméterei
(átviteli függvény, állapotteres
leírás)



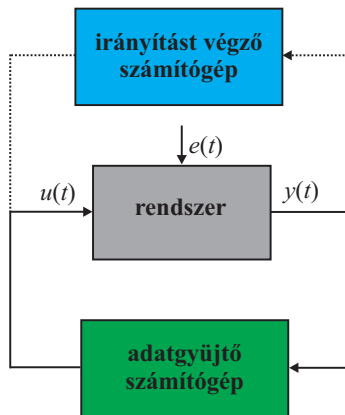
Az identifikáció sémája

Ismert (mért)

- Bemenet (u)
- Kimenet (y) [e zajjal terhelt]

Ismeretlen

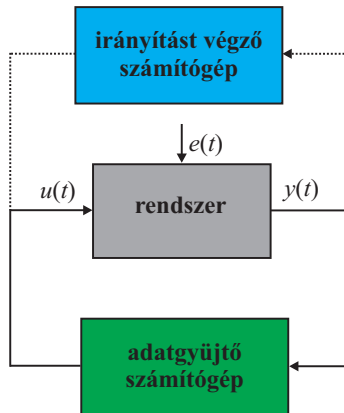
- Zaj (e)
- A rendszer paraméterei
(átviteli függvény, állapotteres leírás)



Rendszertechnikai felépítés

Adatgyűjtés

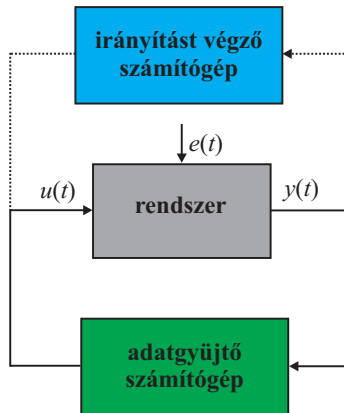
- Az adatgyűjtő számítógép végzi
- A mintavételezett, mért u , y értékeket tárolja
- Az adatgyűjtő számítógép funkcionálisan elkülönül a szabályozástól
- Az adatgyűjtés és szabályozás gyakran ugyanazon a számítógépen fut



Rendszertechnikai felépítés

Adatgyűjtés

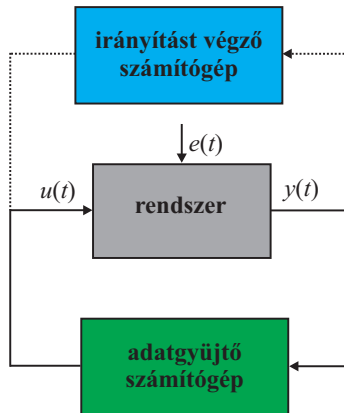
- Az adatgyűjtő számítógép végzi
- A mintavételezett, mért u , y értékeket tárolja
- Az adatgyűjtő számítógép funkcionálisan elkülönül a szabályozástól
- Az adatgyűjtés és szabályozás gyakran ugyanazon a számítógépen fut



Rendszertechnikai felépítés

Adatgyűjtés

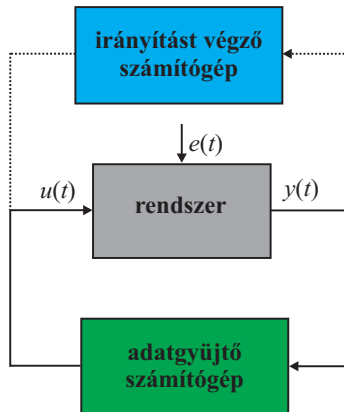
- Az adatgyűjtő számítógép végzi
- A mintavételezett, mért u , y értékeket tárolja
- Az adatgyűjtő számítógép funkcionálisan elkülönül a szabályozástól
- Az adatgyűjtés és szabályozás gyakran ugyanazon a számítógépen fut



Rendszertechnikai felépítés

Adatgyűjtés

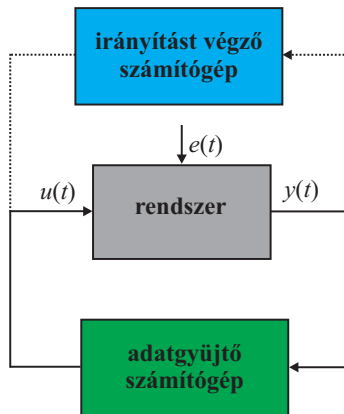
- Az adatgyűjtő számítógép végzi
- A mintavételezett, mért u , y értékeket tárolja
- Az adatgyűjtő számítógép funkcionálisan elkülönül a szabályozástól
- Az adatgyűjtés és szabályozás gyakran ugyanazon a számítógépen fut



Rendszertechnikai felépítés

Adatgyűjtés

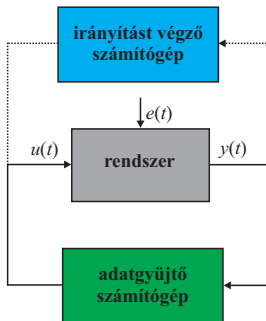
- Az adatgyűjtő számítógép végzi
- A mintavételezett, mért u , y értékeket tárolja
- Az adatgyűjtő számítógép funkcionálisan elkülönül a szabályozástól
- Az adatgyűjtés és szabályozás gyakran ugyanazon a számítógépen fut



Rendszertechnikai felépítés

Identifikáció megvalósítása nyílt körben

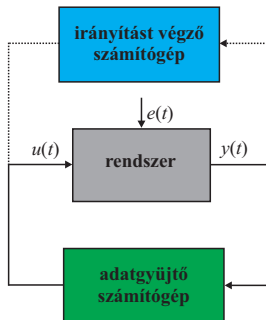
- Az identifikációt végző egységek (adatgyűjtő számítógép) nem részei a szabályozási huroknak
- Az adatgyűjtő számítógép előírhatja az u bemenetet



Rendszertechnikai felépítés

Identifikáció megvalósítása nyílt körben

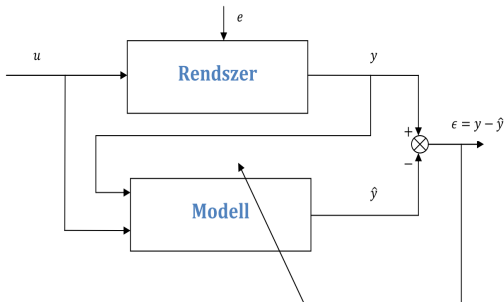
- Az identifikációt végző egységek (adatgyűjtő számítógép) nem részei a szabályozási huroknak
- Az adatgyűjtő számítógép előírhatja az u bemenetet



Rendszerteknikai felépítés

Identifikáció megvalósítása nyílt körben

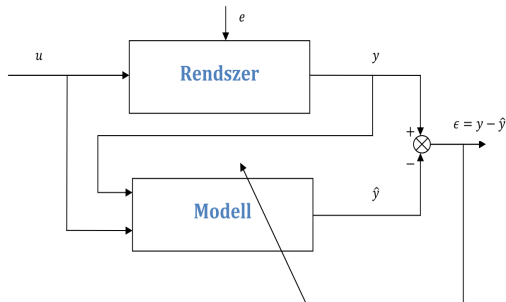
- A mért y kimenet nincs visszacsatolva a folyamat u bemenetére
- Előny: y és u nem korrelál (függetlenek), amely az identifikációs módszerek számára előnyös
- Hátrány: Identifikáció közben nincs szabályozás



Rendszertechnikai felépítés

Identifikáció megvalósítása nyílt körben

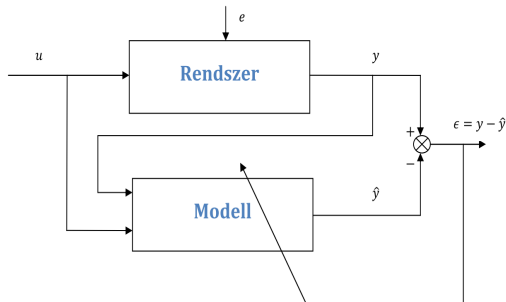
- A mért y kimenet nincs visszacsatolva a folyamat u bemenetére
- Előny: y és u nem korrelál (függetlenek), amely az identifikációs módszerek számára előnyös
- Hátrány: Identifikáció közben nincs szabályozás



Rendszerteknikai felépítés

Identifikáció megvalósítása nyílt körben

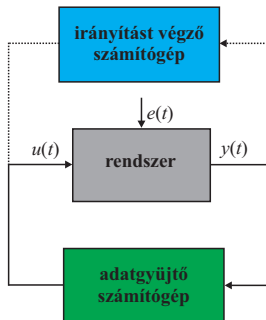
- A mért y kimenet nincs visszacsatolva a folyamat u bemenetére
- Előny: y és u nem korrelál (függetlenek), amely az identifikációs módszerek számára előnyös
- Hátrány: Identifikáció közben nincs szabályozás



Rendszertechnikai felépítés

Identifikáció megvalósítása zárt körben

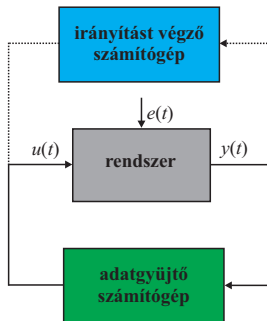
- Az identifikációt végző egységek (adatgyűjtő számítógép) részei a szabályozási huroknak
- Az adatgyűjtő számítógép csak mér és tárol



Rendszertechnikai felépítés

Identifikáció megvalósítása zárt körben

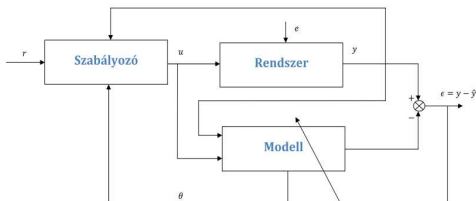
- Az identifikációt végző egységek (adatgyűjtő számítógép) részei a szabályozási huroknak
- Az adatgyűjtő számítógép csak mér és tárol



Rendszertechnikai felépítés

Identifikáció megvalósítása zárt körben

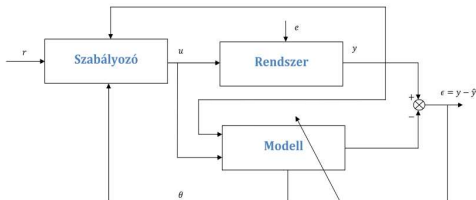
- A mért y kimenet vissza van csatolva a folyamat u bemenetére
- Hátrány: y és u korrelál, amely az identifikációs módszerek számára nem előnyös
- Előny: Identifikáció közben van szabályozás



Rendszertechnikai felépítés

Identifikáció megvalósítása zárt körben

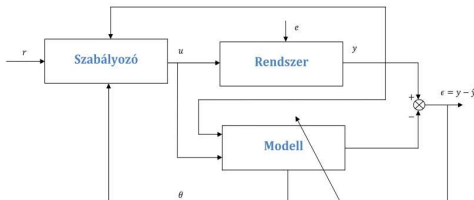
- A mért y kimenet vissza van csatolva a folyamat u bemenetére
- Hátrány: y és u korrelál, amely az identifikációs módszerek számára nem előnyös
- Előny: Identifikáció közben van szabályozás



Rendszertechnikai felépítés

Identifikáció megvalósítása zárt körben

- A mért y kimenet vissza van csatolva a folyamat u bemenetére
- Hátrány: y és u korrelál, amely az identifikációs módszerek számára nem előnyös
- Előny: Identifikáció közben van szabályozás



Identifikáció, mint optimalizálási probléma - Feltevések

Feltevések

- Diszkrétidejű modell
- Ekvidisztánsan mintavételezett jelek
- Lineáris modell (munkaponti linearizálás)

Megjegyzés

Nemlineáris rendszerek identifikációjára a **mesterséges intelligencia módszerek** adnak népszerű alternatívát

Érdeklődőknek: VIIIIM129 Lágyszámítási módszerek

(MSC Inf. Autonóm irányító rendszerek és robotok főszakirány tárgya)



Identifikáció, mint optimalizálási probléma - Feltevések

Feltevések

- Diszkrétidejű modell
- Ekvidisztánsan mintavételezett jelek
- Lineáris modell (munkaponti linearizálás)

Megjegyzés

Nemlineáris rendszerek identifikációjára a **mesterséges intelligencia módszerek** adnak népszerű alternatívát

Érdeklődőknek: VIIIIM129 Lágyszámítási módszerek

(MSC Inf. Autonóm irányító rendszerek és robotok főszakirány tárgya)



Identifikáció, mint optimalizálási probléma - modell kimenet

- Adott (ismertnek vehető)
 - a diszkrétidejű bemeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - $y(t)$ a diszkrétidejű kimeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - A nem mérhető $e(t)$ zajsorozat statisztikai jellemzői
- Gyűjtsük a rendszer ismeretlen (azaz meghatározandó) paramétereit a θ paramétervektorba.
- A lineáris modell kimenete (◀ modellosztályok ▶ prediktor ▶ LS módszer ARX modellre ▶)

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t).$$

Megjegyzések

- Különböző θ értékek különböző modell kimeneteket produkálnak
- A modellkimenet összehasonlítható a valódi kimenettel egy **költségfüggvény** segítségével

Identifikáció, mint optimalizálási probléma - modell kimenet

- Adott (ismertnek vehető)
 - a diszkrétidejű bemeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - $y(t)$ a diszkrétidejű kimeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - A nem mérhető $e(t)$ zajsorozat statisztikai jellemzői
- Gyűjtsük a rendszer ismeretlen (azaz meghatározandó) paramétereit a θ paramétervektorba.
- A lineáris modell kimenete (◀ modellosztályok ▶ prediktor ▶ LS módszer ARX modellre ▶)

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t).$$

Megjegyzések

- Különböző θ értékek különböző modell kimeneteket produkálnak
- A modellkimenet összehasonlítható a valódi kimenettel egy **költségfüggvény** segítségével

Identifikáció, mint optimalizálási probléma - modell kimenet

- Adott (ismertnek vehető)
 - a diszkrétidejű bemeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - $y(t)$ a diszkrétidejű kimeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - A nem mérhető $e(t)$ zajsorozat statisztikai jellemzői
- Gyűjtsük a rendszer ismeretlen (azaz meghatározandó) paramétereit a θ paramétervektorba.
- A lineáris modell kimenete (◀ modellosztályok ◀ prediktor ◀ LS módszer ARX modellre)

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t).$$

Megjegyzések

- Különböző θ értékek különböző modell kimeneteket produkálnak
- A modellkimenet összehasonlítható a valódi kimenettel egy **költségfüggvény** segítségével

Identifikáció, mint optimalizálási probléma - modell kimenet

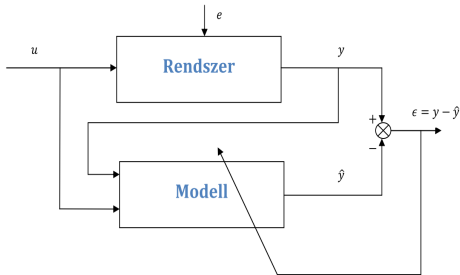
- Adott (ismertnek vehető)
 - a diszkrétidejű bemeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - $y(t)$ a diszkrétidejű kimeneti jelsorozat ($t = \dots, -2, -1, 0, 1, 2, \dots$)
 - A nem mérhető $e(t)$ zajszorozat statisztikai jellemzői
- Gyűjtsük a rendszer ismeretlen (azaz meghatározandó) paramétereit a θ paramétervektorba.
- A lineáris modell kimenete (◀ modellosztályok ▶ prediktor ▶ LS módszer ARX modellre)

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t).$$

Megjegyzések

- Különböző θ értékek különböző modell kimeneteket produkálnak
- A modellkimenet összehasonlítható a valódi kimenettel egy **költségfüggvény** segítségével

Identifikáció, mint optimalizálási probléma - modell kimenet



Megjegyzések

- Különböző θ értékek különböző modell kimeneteket produkálnak
- A modellkimenet összehasonlítható a valódi kimenettel egy **költséggfüggvény** segítségével

Identifikáció, mint optimalizálási probléma - A feladat megfogalmazása

- A modell kimenete újra: $\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$
- A költségfüggvény jelölése: $V(u(t), y(t), \theta)$
- Az optimális paramétervektor, aminek megtalálása az identifikáció célja:

$$\hat{\theta} = \arg \min V(u(t), y(t), \theta)$$

A megoldáshoz tisztázni kell ...

- G és H lehetséges alakjait
- az $e(t)$ zaj leírása után az $e(t)$ zaj hatását
- a V költségfüggvény felírásával a zaj kiküszöbölését $\hat{y}(t)$ számításából
- optimalizálási eljárásokat $\hat{\theta}$ meghatározásához



Identifikáció, mint optimalizálási probléma - A feladat megfogalmazása

- A modell kimenete újra: $\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$
- A költségfüggvény jelölése: $V(u(t), y(t), \theta)$
- Az optimális paramétervektor, aminek megtalálása az identifikáció célja:

$$\hat{\theta} = \arg \min V(u(t), y(t), \theta)$$

A megoldáshoz tisztázni kell ...

- G és H lehetséges alakjait ▶
- az $e(t)$ zaj leírása ▶ után az $e(t)$ zaj hatását ▶
- a V költségfüggvény felírásával a zaj kiküszöbölését $\hat{y}(t)$ számításából ▶
- optimalizálási eljárásokat $\hat{\theta}$ meghatározásához ▶



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében**
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszervei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Stabil, lineáris operátorok a sorozatok terében

◀ Megoldás lépései

▶▶ Fejezet kihagyása



Eltolás operátor

- **A mintavételezett jelek reprezentációja:** Számsorozatok
- **A rendszer reprezentációja:** Ezen számsorozatokon értelmezett operátorok.



Eltolás operátor

- **A mintavételezett jelek reprezentációja:** Számsorozatok
- **A rendszer reprezentációja:** Ezen számsorozatokon értelmezett operátorok.

Akkor alkalmazzunk (egyoldalas) $\mathcal{Z}$ transzformációt!?

NEM! Mert,...

- az $e(t)$ zajt stacionárius folyamat $\rightarrow$ folyamat jelei nemnullák $t < 0$ időpontokra
 $\leftrightarrow$ Egyoldalas $\mathcal{Z}$ transzformáció alkalmazásának feltétele.



Eltolás operátor

- **A mintavételezett jelek reprezentációja:** Számsorozatok
- **A rendszer reprezentációja:** Ezen számsorozatokon értelmezett operátorok.

Akkor alkalmazzunk (egyoldalas) $\mathcal{Z}$ transzformációt!?

NEM! Mert,...

- az $e(t)$ zajt stacionárius folyamat $\rightarrow$ folyamat jelei nemnullák $t < 0$ időpontokra
 $\leftrightarrow$ Egyoldalas $\mathcal{Z}$ transzformáció alkalmazásának feltétele.

Akkor mit tegyünk?

A q, q^{-1} eltolási operátorok bevezetése:



Eltolás operátor

- **A mintavételezett jelek reprezentációja:** Számsorozatok
- **A rendszer reprezentációja:** Ezen számsorozatokon értelmezett operátorok.

Akkor alkalmazzunk (egyoldalas) $\mathcal{Z}$ transzformációt!?

NEM! Mert,...

- az $e(t)$ zajt stacionárius folyamat $\rightarrow$ folyamat jelei nemnullák $t < 0$ időpontokra
 $\leftrightarrow$ Egyoldalas $\mathcal{Z}$ transzformáció alkalmazásának feltétele.

Akkor mit tegyünk?

A q, q^{-1} eltolási operátorok bevezetése:



Jel és rendszerreprezentáció kérdése

- Jelölje ℓ az (általában komplex)

$$x = \{x_n\} = \{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\} = \{\dots, x(-2), x(-1), x(0), x(1), x(2), \dots\}$$

sorozatok végtelen dimenziós vektorterét

A vektortérben definiált összeadás és a skalárral való szorzás:

$$x, y \in \ell, c \in \mathbb{C} \quad z = \{z_n\} = x + y = \{x_n + y_n\} \quad c \cdot x = \{c \cdot x_n\}$$

- Legyen ℓ_2 az ℓ azon részhalmaza, melyre teljesül, hogy

$$\|x\|^2 = \sum_{n=-\infty}^{\infty} x_n x_n^* < \infty, \forall x \in \ell_2 \quad (* \text{ a konjugált})$$

Megjegyzés

Az ℓ_2 szintén vektortér és ezen belül Hilbert-tér, mivel,

- Értelmezhető a skalár szorzat: $\langle x, y \rangle = \sum_{n=-\infty}^{\infty} x_n y_n^*$ az
- $\|x\|^2 = \sum_{n=-\infty}^{\infty} x_n x_n^* < \infty, \forall x \in \ell_2$ normával
- Minden Cauchy sorozatnak van határértéke ℓ_2 -ben

Jel és rendszerreprezentáció kérdése

- Jelölje ℓ az (általában komplex)

$$x = \{x_n\} = \{\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots\} = \{\dots, x(-2), x(-1), x(0), x(1), x(2), \dots\}$$

sorozatok végtelen dimenziós vektorterét

A vektortérben definiált összeadás és a skalárral való szorzás:

$$x, y \in \ell, c \in \mathbb{C} \quad z = \{z_n\} = x + y = \{x_n + y_n\} \quad c \cdot x = \{c \cdot x_n\}$$

- Legyen ℓ_2 az ℓ azon részhalmaza, melyre teljesül, hogy

$$\|x\|^2 = \sum_{n=-\infty}^{\infty} x_n x_n^* < \infty, \forall x \in \ell_2 \quad (* \text{ a konjugált})$$

Megjegyzés

Az ℓ_2 szintén vektortér és ezen belül Hilbert-tér, mivel,

- Értelmezhető a skalár szorzat: $\langle x, y \rangle = \sum_{n=-\infty}^{\infty} x_n y_n^*$ az
- $\|x\|^2 = \sum_{n=-\infty}^{\infty} x_n x_n^* < \infty, \forall x \in \ell_2$ normával
- Minden Cauchy sorozatnak van határértéke ℓ_2 -ben

Eltolás operátor

A q eltolási (vagy shift) operátor

Vezessük be az i identitás és q, q^{-1} ún. shift-, vagy eltolás operátorokat az ℓ_2 vektortérben:

$$\begin{aligned} q &: \ell_2 \rightarrow \ell_2 & q \cdot x(t) &= x(t+1) \\ q^{-1} &: \ell_2 \rightarrow \ell_2 & q^{-1} \cdot x(t) &= x(t-1) \\ i &: \ell_2 \rightarrow \ell_2 & i \cdot x(t) &= x(t). \end{aligned}$$

Definíció - $A : \ell_2 \rightarrow \ell_2$ **korlátos, lineáris** operátor

$\exists c > 0$, hogy $\forall x \in \ell_2$ esetén $\|A \cdot x\| \leq c\|x\|$.

A korlátos lineáris operátorok halmazának jelölése: $K(\ell_2 \rightarrow \ell_2)$

Az A korlátos lineáris operátor normája:

$$\begin{aligned} \|A\| &= \inf \{c > 0 : \|Ax\| \leq c\|x\|, \forall x \in \ell_2\} < \infty \\ \|A\| &= \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| < \infty \end{aligned}$$

q, q^{-1} és i korlátos lineáris operátorok ($c = 1$)



Eltolás operátor

A q eltolási (vagy shift) operátor

Vezessük be az i identitás és q, q^{-1} ún. shift-, vagy eltolás operátorokat az ℓ_2 vektortérben:

$$\begin{aligned} q &: \ell_2 \rightarrow \ell_2 & q \cdot x(t) &= x(t+1) \\ q^{-1} &: \ell_2 \rightarrow \ell_2 & q^{-1} \cdot x(t) &= x(t-1) \\ i &: \ell_2 \rightarrow \ell_2 & i \cdot x(t) &= x(t). \end{aligned}$$

Definíció - $A : \ell_2 \rightarrow \ell_2$ **korlátos, lineáris** operátor

$\exists c > 0$, hogy $\forall x \in \ell_2$ esetén $\|A \cdot x\| \leq c\|x\|$.

A korlátos lineáris operátorok halmazának jelölése: $K(\ell_2 \rightarrow \ell_2)$

Az A korlátos lineáris operátor normája:

$$\begin{aligned} \|A\| &= \inf \{c > 0 : \|Ax\| \leq c\|x\|, \forall x \in \ell_2\} < \infty \\ \|A\| &= \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| < \infty \end{aligned}$$

q, q^{-1} és i korlátos lineáris operátorok ($c = 1$)



Eltolás operátor

A q eltolási (vagy shift) operátor

Vezessük be az i identitás és q, q^{-1} ún. shift-, vagy eltolás operátorokat az ℓ_2 vektortérben:

$$\begin{aligned} q &: \ell_2 \rightarrow \ell_2 & q \cdot x(t) &= x(t+1) \\ q^{-1} &: \ell_2 \rightarrow \ell_2 & q^{-1} \cdot x(t) &= x(t-1) \\ i &: \ell_2 \rightarrow \ell_2 & i \cdot x(t) &= x(t). \end{aligned}$$

Definíció - $A : \ell_2 \rightarrow \ell_2$ **korlátos, lineáris** operátor

$\exists c > 0$, hogy $\forall x \in \ell_2$ esetén $\|A \cdot x\| \leq c\|x\|$.

A korlátos lineáris operátorok halmazának jelölése: $K(\ell_2 \rightarrow \ell_2)$

Az A korlátos lineáris operátor normája:

$$\begin{aligned} \|A\| &= \inf \{c > 0 : \|Ax\| \leq c\|x\|, \forall x \in \ell_2\} < \infty \\ \|A\| &= \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| < \infty \end{aligned}$$

q, q^{-1} és i korlátos lineáris operátorok ($c = 1$)

Kauzális, korlátos, lineáris operátor

Számunkra csak a **kauzális** korlátos, lineáris operátorok elfogadhatóak

$$H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + h_3 q^{-3} + \dots,$$

ahol q^{-k} a q^{-1} operátor k darab egymást követő alkalmazását jelenti.



Átviteli függvények és lineáris operátorok kapcsolata

- Kiindulópont a diszkrét átviteli függvény:

$$H(z) = \frac{B(z)}{A(z)}, \quad \text{gr}B(z) \leq \text{gr}A(z)$$

és $A(z)$ gyökei az egységkörön **belül** vannak azaz **$A(z)$ polinom stabil**.

- $H(z)$ részlettörtekre bontása után mértani sor összegeként írható fel:

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots$$

Konvergens $\forall |z| \geq 1$ választás mellett, ha

- $A(z)$ gyökei az egységkörön belül vannak
- $\sum_{n=1}^{\infty} |h_n| < \infty$
- Most rendeljük hozzá $H(z)$ racionális törtfüggvényhez egy $H(q)$ operátort:

$$H(z) \longrightarrow H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + \dots,$$



Átviteli függvények és lineáris operátorok kapcsolata

- Kiindulópont a diszkrét átviteli függvény:

$$H(z) = \frac{B(z)}{A(z)}, \quad \text{gr}B(z) \leq \text{gr}A(z)$$

és $A(z)$ gyökei az egységkörön **belül** vannak azaz **$A(z)$ polinom stabil**.

- $H(z)$ részlettörtekre bontása után mértani sor összegeként írható fel:

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots$$

Konvergens $\forall |z| \geq 1$ választás mellett, ha

- $A(z)$ gyökei az egységkörön belül vannak
- $\sum_{n=1}^{\infty} |h_n| < \infty$
- Most rendeljünk hozzá $H(z)$ racionális törtfüggvényhez egy $H(q)$ operátort:

$$H(z) \longrightarrow H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + \dots,$$



Átviteli függvények és lineáris operátorok kapcsolata

- Kiindulópont a diszkrét átviteli függvény:

$$H(z) = \frac{B(z)}{A(z)}, \quad \text{gr}B(z) \leq \text{gr}A(z)$$

és $A(z)$ gyökei az egységkörön **belül** vannak azaz **$A(z)$ polinom stabil**.

- $H(z)$ részlettörtekre bontása után mértani sor összegeként írható fel:

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + \dots$$

Konvergens $\forall |z| \geq 1$ választás mellett, ha

- $A(z)$ gyökei az egységkörön belül vannak
- $\sum_{n=1}^{\infty} |h_n| < \infty$
- Most rendeljünk hozzá $H(z)$ racionális törtfüggvényhez egy $H(q)$ operátort:

$$H(z) \longrightarrow H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + \dots,$$



Átviteli függvények és lineáris operátorok kapcsolata

$$H(z) \longrightarrow H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + \dots,$$

- **A $H(z)$ hozzárendelés egyértelmű**
- **$H(q)$ korlátos**, mivel $\sum_{n=1}^{\infty} |h_n| < \infty$ miatt

$$H(q) \cdot x(t) = h_0 x(t) + h_1 x(t-1) + h_2 x(t-2) + \dots,$$

$$\|H(q) \cdot x(t)\| \leq |h_0| \|x\| + |h_1| \|x\| + |h_2| \|x\| + \dots = \sum_{n=0}^{\infty} |h_n| \|x\| < \infty$$

Következmények

- minden stabil racionális törtfüggvényhez egy és csak egy korlátos lineáris operátor tartozik
stabil racionális törtfüggvény $\equiv$ stabil A polinom
- ha $B(z)$ polinom is stabil, akkor létezik a $H(q)$ inverz operátora
 $H^{-1}(q) \in K(\ell_2 \rightarrow \ell_2)$

Átviteli függvények és lineáris operátorok kapcsolata

$$H(z) \longrightarrow H(q) = h_0 + h_1 q^{-1} + h_2 q^{-2} + \dots,$$

- **A $H(z)$ hozzárendelés egyértelmű**
- **$H(q)$ korlátos**, mivel $\sum_{n=1}^{\infty} |h_n| < \infty$ miatt

$$H(q) \cdot x(t) = h_0 x(t) + h_1 x(t-1) + h_2 x(t-2) + \dots,$$

$$\|H(q) \cdot x(t)\| \leq |h_0| \|x\| + |h_1| \|x\| + |h_2| \|x\| + \dots = \sum_{n=0}^{\infty} |h_n| \|x\| < \infty$$

Következmények

- minden stabil racionális törtfüggvényhez egy és csak egy korlátos lineáris operátor tartozik

stabil racionális törtfüggvény $\equiv$ stabil A polinom

- ha $B(z)$ polinom is stabil, akkor létezik a $H(q)$ inverz operátora
 $H^{-1}(q) \in K(\ell_2 \rightarrow \ell_2)$

Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségyszámítási alapjai**
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszervei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Identifikáció valószínűségszámítási alapjai

◀ Megoldás lépései

▶▶ Fejezet kihagyása



Amiért szükség van rá...

- Az identifikálandó rendszerre ható zaj: Egy **sztochasztikus folyamat** realizációja
- A zaj mérési időpontokban felvett értéke: **valószínűségi változó**



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségyszámítási alapjai**
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alaplódszerei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Alapfogalmak egyváltozós esetben

Valószínűségi mező

$(\Omega, \mathcal{A}, P)$ rendezett hármas a jelek leírására alkalmas valószínűségi mező

- Ω az eseménytér
- $\mathcal{A}$ tartalmazza az eseménytér valószínűséggel rendelkező eseményeit ($\mathcal{A}$ az Ω részhalmazainak egy σ -algebrája)
- P egy normált mérték $\mathcal{A}$ felett ($P(\Omega) = 1$)

Legyen $\xi = \xi(\omega)$ egy valószínűségi változó Ω felett

E - Várható érték operátor

$$E\xi = \int_{\Omega} \xi(\omega) dP.$$

A ξ valószínűségi változó n -ik momentuma: $E\xi^n$



Alapfogalmak egyváltozós esetben

Valószínűségi mező

$(\Omega, \mathcal{A}, P)$ rendezett hármas a jelek leírására alkalmas valószínűségi mező

- Ω az eseménytér
- $\mathcal{A}$ tartalmazza az eseménytér valószínűséggel rendelkező eseményeit ($\mathcal{A}$ az Ω részhalmazainak egy σ -algebrája)
- P egy normált mérték $\mathcal{A}$ felett ($P(\Omega) = 1$)

Legyen $\xi = \xi(\omega)$ egy valószínűségi változó Ω felett

E - Várható érték operátor

$$E\xi = \int_{\Omega} \xi(\omega) dP.$$

A ξ valószínűségi változó n -ik momentuma: $E\xi^n$



Alapfogalmak egyváltozós esetben

D operátor - A ξ valószínűségi változó szórása

$$D\xi = +\sqrt{E((\xi - E\xi)^2)}.$$

Variancia: $(D\xi)^2$ (A szórás négyzete)

$F(x)$ - ξ valószínűségi változó eloszlásfüggvénye

$$F(x) = P(\xi < x) \stackrel{\text{def}}{=} P(\{\omega : \xi(\omega) < x\})$$

$f(x)$ sűrűségfüggvény: az $F(x)$ eloszlásfüggvény deriváltja.

egydimenziós Gauss-féle vagy normális eloszlás

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right),$$

ahol m a várható érték és σ a szórás



Alapfogalmak egyváltozós esetben

D operátor - A ξ valószínűségi változó szórása

$$D\xi = +\sqrt{E((\xi - E\xi)^2)}.$$

Variancia: $(D\xi)^2$ (A szórás négyzete)

$F(x)$ - ξ valószínűségi változó eloszlásfüggvénye

$$F(x) = P(\xi < x) \stackrel{\text{def}}{=} P(\{\omega : \xi(\omega) < x\})$$

$f(x)$ sűrűségfüggvény: az $F(x)$ eloszlásfüggvény deriváltja.

egydimenziós Gauss-féle vagy normális eloszlás

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right),$$

ahol m a várható érték és σ a szórás



Alapfogalmak egyváltozós esetben

D operátor - A ξ valószínűségi változó szórása

$$D\xi = +\sqrt{E((\xi - E\xi)^2)}.$$

Variancia: $(D\xi)^2$ (A szórás négyzete)

$F(x)$ - ξ valószínűségi változó eloszlásfüggvénye

$$F(x) = P(\xi < x) \stackrel{\text{def}}{=} P(\{\omega : \xi(\omega) < x\})$$

$f(x)$ sűrűségfüggvény: az $F(x)$ eloszlásfüggvény deriváltja.

egydimenziós Gauss-féle vagy normális eloszlás

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right),$$

ahol m a várható érték és σ a szórás



Alapfogalmak többváltozós esetben

Többváltozós eloszlásfüggvény, sűrűségfüggvény

Legyenek $\xi_1, \xi_2, \dots, \xi_n$ valószínűségi változók, ezeket vektorba rendezve:
 $\xi = (\xi_1, \xi_2, \dots, \xi_n)$.

Az együttes eloszlásfüggvény:

$$F(x_1, x_2, \dots, x_n) = P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n)$$

Ennek sűrűségfüggvénye

$$f(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}.$$



Alapfogalmak többváltozós esetben

Legyenek

- ξ és η valószínűségi **vektor**változók
- ξ és η várható értéke: $m_\xi = E\xi$ és $m_\eta = E\eta$

Keresztkovariancia mátrix

$$R_{\xi,\eta} = \text{cov}[\xi\eta] = E[(\xi - m_\xi)(\eta - m_\eta)^T]$$

Keresztkorrelációs mátrix

$$\sigma_{\xi\eta} = \frac{\text{cov}[\xi, \eta]}{D\xi D\eta}$$

Speciálisan

- ξ (auto-) kovariancia mátrixa

$$R_{\xi\xi} = \text{cov}[\xi, \xi]$$

- ξ (auto-) korrelációs mátrixa

$$\sigma_{\xi\xi} = \frac{R_{\xi\xi}}{D^2\xi}$$

Megjegyzés

$R_{\xi,\eta}$ keresztkovariancia mátrix és $\sigma_{\xi\eta}$ keresztkorrelációs mátrix a ξ és η valószínűségi változók statisztikai „függőségét” fejezik ki.

Alapfogalmak többváltozós esetben

Legyenek

- ξ és η valószínűségi **vektor**változók
- ξ és η várható értéke: $m_\xi = E\xi$ és $m_\eta = E\eta$

Keresztkovariancia mátrix

$$R_{\xi,\eta} = \text{cov}[\xi\eta] = E[(\xi - m_\xi)(\eta - m_\eta)^T]$$

Keresztkorrelációs mátrix

$$\sigma_{\xi\eta} = \frac{\text{cov}[\xi, \eta]}{D\xi D\eta}$$

Speciálisan

- ξ (auto-) kovariancia mátrixa

$$R_{\xi\xi} = \text{cov}[\xi, \xi]$$

- ξ (auto-) korrelációs mátrixa

$$\sigma_{\xi\xi} = \frac{R_{\xi\xi}}{D^2\xi}$$

Megjegyzés

$R_{\xi,\eta}$ keresztkovariancia mátrix és $\sigma_{\xi\eta}$ keresztkorrelációs mátrix a ξ és η valószínűségi változók statisztikai „függőségét” fejezik ki.

Alapfogalmak többváltozós esetben

Legyenek

- ξ és η valószínűségi **vektor**változók
- ξ és η várható értéke: $m_\xi = E\xi$ és $m_\eta = E\eta$

Keresztkovariancia mátrix

$$R_{\xi,\eta} = \text{cov}[\xi\eta] = E[(\xi - m_\xi)(\eta - m_\eta)^T]$$

Keresztkorrelációs mátrix

$$\sigma_{\xi\eta} = \frac{\text{cov}[\xi, \eta]}{D\xi D\eta}$$

Speciálisan

- ξ (auto-) kovariancia mátrixa

$$R_{\xi\xi} = \text{cov}[\xi, \xi]$$

- ξ (auto-) korrelációs mátrixa

$$\sigma_{\xi\xi} = \frac{R_{\xi\xi}}{D^2\xi}$$

Megjegyzés

$R_{\xi,\eta}$ keresztkovariancia mátrix és $\sigma_{\xi\eta}$ keresztkorrelációs mátrix a ξ és η valószínűségi változók statisztikai „függőségét” fejezik ki.

Alapfogalmak többváltozós esetben

Többdimenziós Gauss- vagy normális eloszlás

Az (n elemű) ξ valószínűségi vektorváltozót n -dimenziós Gauss- vagy normális eloszlásúnak nevezzük, ha sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det(R_{\xi\xi})}} \exp \left(-\frac{(x - m_{\xi})^T R_{\xi\xi}^{-1} (x - m_{\xi})}{2} \right)$$



Korrelálatlanság, függetlenség

Definíció - Korrelálatlanság

A ξ és η valószínűségi vektorváltozókat korrelálatlannak nevezzük, ha teljesül

$$E[\xi\eta^T] = E\xi E\eta^T = m_\xi m_\eta^T,$$

vagy, ami ezzel ekvivalens

$$R_{\xi\eta} = E[(\xi - m_\xi)(\eta - m_\eta)^T] = E[\xi\eta^T] - m_\xi m_\eta^T = 0$$

A korrelálatlanságnál erősebb követelmény a függetlenség:

Definíció - Függetlenség

$\xi_1, \xi_2, \dots, \xi_n$ valószínűségi változók függetlenek ha

$$P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n) = \prod_{i=1}^n P(\xi_i < x_i),$$

vagy vektorváltozók esetén

$$P(\xi < x, \eta < y) = P(\xi < x)P(\eta < y).$$

Korrelálatlanság, függetlenség

Definíció - Korrelálatlanság

A ξ és η valószínűségi vektorváltozókat korrelálatlannak nevezzük, ha teljesül

$$E[\xi\eta^T] = E\xi E\eta^T = m_\xi m_\eta^T,$$

vagy, ami ezzel ekvivalens

$$R_{\xi\eta} = E[(\xi - m_\xi)(\eta - m_\eta)^T] = E[\xi\eta^T] - m_\xi m_\eta^T = 0$$

A korrelálatlanságnál erősebb követelmény a függetlenség:

Definíció - Függetlenség

$\xi_1, \xi_2, \dots, \xi_n$ valószínűségi változók függetlenek ha

$$P(\xi_1 < x_1, \xi_2 < x_2, \dots, \xi_n < x_n) = \prod_{i=1}^n P(\xi_i < x_i),$$

vagy vektorváltozók esetén

$$P(\xi < x, \eta < y) = P(\xi < x)P(\eta < y).$$

Korrelálatlanság, függetlenség

Megjegyzések

- Független valószínűségi változók együttes eloszlásfüggvénye a komponens valószínűségi eloszlásfüggvényeinek szorzata
- Ugyanez fennáll a sűrűségfüggvényekre is
- Ha ξ és η függetlenek, akkor korrelálatlanok is, de a fordítottja általában nem igaz



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségszámítási alapjai**
 - Valószínűségi változók
 - **Sztochasztikus folyamatok**
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszabványai
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Alapfogalmak

Sztochasztikus folyamat

Ha az $x(t)$ (diszkrét idejű) jel megfigyelései az $\omega \in \Omega$ eseménytől is függenek, azaz különböző ω eseményekhez különböző jelsorozatok tartoznak.

Megjegyzések:

- Precíz jelölés: $x(t, \omega)$
Ha nem okoz félreértést, akkor röviden: $x(t)$
- $x(t)$ több jelet is tartalmazó vektor is lehet.



Alapfogalmak

Valószínűségi változó származtatása a folyamatból

Ha lerögzítjük az idő értékét $t = t_1$ -re. Akkor $\xi_1 \stackrel{\text{def}}{=} x(t_1, \omega)$ valószínűségi változóhoz jutunk

- amelynek $x(t_1, \omega_1), x(t_1, \omega_2), x(t_1, \omega_3), \dots$ a realizációi
- amely jellemezhető az $F(\xi_1 < x_1; t_1)$ eloszlásfüggvénnyel.
Rövid jelölés : $F(\xi_1, t_1) = F(x(t_1) < x_1)$



Alapfogalmak

Valószínűségi változó származtatása a folyamatból

Ha lerögzítjük az idő értékét $t = t_1$ -re. Akkor $\xi_1 \stackrel{\text{def}}{=} x(t_1, \omega)$ valószínűségi változóhoz jutunk

- amelynek $x(t_1, \omega_1), x(t_1, \omega_2), x(t_1, \omega_3), \dots$ a realizációi
- amely jellemezhető az $F(\xi_1 < x_1; t_1)$ eloszlásfüggvénnyel.
Rövid jelölés : $F(\xi_1, t_1) = F(x(t_1) < x_1)$

Valószínűségi **vektor**változó származtatása a folyamatból

Ha véges sok $t_1, t_2, \dots, t_n$ időpontot rögzítünk le, akkor az $\xi_i \stackrel{\text{def}}{=} x(t_i, \omega)$ n darab valószínűségi (vektor) változót definiál: $\xi_1, \xi_2, \dots, \xi_n$



Alapfogalmak

Probléma

Egy $x(t)$ folyamat kimenetének minden $-\infty < t < +\infty$ időpontban való jellemzése végtelen dimenziós eloszlásfüggvényt igényel .



Alapfogalmak

Probléma

Egy $x(t)$ folyamat kimenetének minden $-\infty < t < +\infty$ időpontban való jellemzése végtelen dimenziós eloszlásfüggvényt igényel .

Ehelyett...

véges számú, mondjuk n darab $(t_1, t_2, \dots, t_n)$ időpontot választunk ki, így egy *végesdimenziós* eloszlásfüggvényhez jutunk:

$$F(\xi_1, \dots, \xi_n; t_1, \dots, t_n) = P(x(t_1) < \xi_1, \dots, x(t_n) < \xi_n).$$



Alapfogalmak

Probléma

Egy $x(t)$ folyamat kimenetének minden $-\infty < t < +\infty$ időpontban való jellemzése végtelen dimenziós eloszlásfüggvényt igényel.

Ehelyett...

véges számú, mondjuk n darab $(t_1, t_2, \dots, t_n)$ időpontot választunk ki, így egy *végesdimenziós* eloszlásfüggvényhez jutunk:

$$F(\xi_1, \dots, \xi_n; t_1, \dots, t_n) = P(x(t_1) < \xi_1, \dots, x(t_n) < \xi_n).$$

Definíció - Gauss folyamat

Gauss folyamatról van szó, ha az $F(\xi_1, \dots, \xi_n; t_1, \dots, t_n)$ eloszlásfüggvény deriváltja a korábban definiált többdimenziós Gauss (normális) eloszlás sűrűségfüggvénye

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det(R_{\xi\xi})}} \exp \left(-\frac{(x - m_\xi)^T R_{\xi\xi}^{-1} (x - m_\xi)}{2} \right)$$

Sztochasztikus folyamatok osztályozása

A sztochasztikus folyamat t időpontban vett **várható értéke**

$$m(t) = Ex(t) = \int_{-\infty}^{\infty} \xi dF(\xi; t)$$



Sztochasztikus folyamatok osztályozása

A sztochasztikus folyamat t időpontban vett **várható értéke**

$$m(t) = Ex(t) = \int_{-\infty}^{\infty} \xi dF(\xi; t)$$

Az $x(t)$ és $y(t)$ folyamatok **keresztkovariancia függvénye**

$$\begin{aligned} R_{xy}(s, t) &= \text{cov}[x(s), y(t)] = E \left[(x(s) - m_x(s))(y(t) - m_y(t))^T \right] = \\ &= \int \int (\xi_x(s) - m_x(s))(\xi_y(t) - m_y(t))^T dF(\xi_x, \xi_y; s, t) \end{aligned}$$



Sztochasztikus folyamatok osztályozása

A sztochasztikus folyamat t időpontban vett **várható értéke**

$$m(t) = Ex(t) = \int_{-\infty}^{\infty} \xi dF(\xi; t)$$

Az $x(t)$ és $y(t)$ folyamatok **keresztkovariancia függvénye**

$$\begin{aligned} R_{xy}(s, t) &= \text{cov}[x(s), y(t)] = E \left[(x(s) - m_x(s))(y(t) - m_y(t))^T \right] = \\ &= \int \int (\xi_x(s) - m_x(s))(\xi_y(t) - m_y(t))^T dF(\xi_x, \xi_y; s, t) \end{aligned}$$

Egy sztochasztikus folyamat **autokovariancia függvénye**
(azaz saját magával vett keresztkovariancia függvénye)

$$R_{xx}(s, t) = \text{cov}[x(s), x(t)]$$

Sztochasztikus folyamatok jellemzői

Erős értelemben vett stacionárius sztochasztikus folyamat - Definíció

Ha egy sztochasztikus folyamat minden véges dimenziójú eloszlásfüggvénye érzéketlen az időbeli eltolásra, azaz $\forall \tau, n, t_1, \dots, t_n$ választásra teljesül, hogy

$$F(\xi_1, \dots, \xi_n; t_1, \dots, t_n) = F(\xi_1, \dots, \xi_n; t_1 + \tau, \dots, t_n + \tau),$$

akkor *erős értelemben stacionárius* sztochasztikus folyamatról beszélünk.



Sztochasztikus folyamatok jellemzői

Gyenge értelemben vett stacionárius sztochasztikus folyamat - Definíció

Azok a sztochasztikus folyamatok, ahol csak az első és második momentumok érzéketlen az időbeli eltolásra. Azaz

- gyengén stacionárius folyamatok várható értéke minden t -re megegyezik:

$$Ex(t) = Ex(t + \tau) = \textit{konstans}$$

- a kovariancia függvények kizárólag az eltolástól (időkülönbségtől) függenek:

$$E[x(t_1)y(t_2)] = E[x(t_1 + \tau)y(t_2 + \tau)] = E[x(t_1 - t_2), y(0)]$$

$$E[x(t_1)x(t_2)] = E[x(t_1 + \tau)x(t_2 + \tau)] = E[x(t_1 - t_2), x(0)]$$

↓

$$R_{xy}(s, t) = R_{xy}(s - t) = R_{xy}(\tau)$$

$$R_{xy}(\tau) = \text{cov}[x(t + \tau), y(t)]$$

$$R_{xx}(\tau) = \text{cov}[x(t + \tau), x(t)]$$

Kovariancia függvény tulajdonságai

A legtöbb szemléletes információt egy gyengén stacionárius folyamatról annak kovariancia függvénye hordozza, amelynek legfontosabb tulajdonságai:

- 1 $R_{xx}(0)$ jellemzi a folyamat fluktuációját,
- 2 $\sqrt{R_{xx}(0)} = \sigma$ a folyamat szórása, ahonnan a folyamat korrelációs függvénye $\rho_{xx} = R_{xx}(\tau)/R_{xx}(0)$,
- 3 az $\rho_{xx}(\tau)$ korrelációs függvény a τ távolságra lévő értékek közötti függőséget jellemzi:

$$\rho_{xx}(\tau) = \begin{cases} \approx 1 : \text{erős koreláció} \\ \approx 0 : \text{nincs koreláció} \\ \approx -1 : \text{erős negatív koreláció} \end{cases}$$

- 4 Skalárfolyamat esetén $R_{xx}(\tau) = R_{xx}(-\tau)$ (páros függvény).



Fehér zaj

Definíció - Diszkrétidejű fehér zaj

Egy $x(t, \omega)$ sztochasztikus folyamatot diszkrétidejű fehér zajnak nevezzük, ha $x(t, \omega)$

- 1 független,
- 2 egyforma eloszlású,
- 3 nulla várható értékű, és
- 4

$$R_{xx}(\tau) = \begin{cases} \sigma^2, & \text{ha } \tau = 0 \\ 0, & \text{különben.} \end{cases}$$



Fehér zaj

Definíció - Diszkrétidejű fehér zaj

Egy $x(t, \omega)$ sztochasztikus folyamatot diszkrétidejű fehér zajnak nevezzük, ha $x(t, \omega)$

- 1 független,
- 2 egyforma eloszlású,
- 3 nulla várható értékű, és
- 4

$$R_{xx}(\tau) = \begin{cases} \sigma^2, & \text{ha } \tau = 0 \\ 0, & \text{különben.} \end{cases}$$

Definíció - Gauss-féle fehér zaj

- 1 Diszkrétidejű fehér zaj +
- 2 minden t esetén a $\xi_t \stackrel{\text{def}}{=} x(t, \omega)$ valószínűségi változó normális eloszlású σ szórással és nulla várható értékkel: $\mathcal{N}(0, \sigma)$.

Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alaplódszerei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Diszkrétidejű lineáris rendszermodellek zajos környezetben

◀ Megoldás lépései

▶▶ Fejezet kihagyása



Modellosztályok

Megjegyzések

- 1 A legáltalánosabb a PEM modellosztály, amelynek az összes többi speciális esete.
- 2 Minden modellosztály más más G és H operátort generál az identifikációhoz
(▶ modell kimenet)
- 3 A fenti modellosztályokat használja a **Matlab IDENT Toolbox**a

Identifikáció feladata

A megfelelő modellosztály paramétereit illeszteni a rendszerhez, azaz ugyanazt a bemenet-kimenet közötti viselkedést szeretnénk elérni a modellenél is, mint ami a rendszerénél van.



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségszámítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója**
- 7 A paraméterbecslés alaplódszerei
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



A modell kimenetének optimális predikciója

◀ Megoldás lépései

▶▶ Fejezet kihagyása



A predikciós feladat

Legyen

- A modell kimenete újra: $\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$ (modell kimenet)
- A kimenet t időpillanatban ismert értékei: $y_{-\infty}^t = \{\dots, y(t-2), y(t-1), y(t)\}$
- A bemenet t időpillanatban ismert értékei:
 $u_{-\infty}^{t+k} = \{\dots, u(t+k-2), u(t+k-1), u(t+k)\}$
 (A bemenetet mi adjuk ki, annak értékeit előre is ismertnek tételezzük fel.)



A predikciós feladat

Legyen

- A modell kimenete újra: $\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$ (modell kimenet)
- A kimenet t időpillanatban ismert értékei: $y_{-\infty}^t = \{\dots, y(t-2), y(t-1), y(t)\}$
- A bemenet t időpillanatban ismert értékei:
 $u_{-\infty}^{t+k} = \{\dots, u(t+k-2), u(t+k-1), u(t+k)\}$
 (A bemenetet mi adjuk ki, annak értékeit előre is ismertnek tételezzük fel.)

k lépéssel előre tartó optimális prediktor

Előállítja $y(t+k)$ becslését az ismert információk birtokában:

$$\hat{y}(t+k|y_{-\infty}^t, u_{-\infty}^{t+k}) = \hat{y}(t+k|t)$$

miközben a becslési hiba várható értékét minimalizálja (LS becslő)

$$E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$$



A predikciós feladat

Legyen

- A modell kimenete újra: $\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$ (modell kimenet)
- A kimenet t időpillanatban ismert értékei: $y_{-\infty}^t = \{\dots, y(t-2), y(t-1), y(t)\}$
- A bemenet t időpillanatban ismert értékei:
 $u_{-\infty}^{t+k} = \{\dots, u(t+k-2), u(t+k-1), u(t+k)\}$
 (A bemenetet mi adjuk ki, annak értékeit előre is ismertnek tételezzük fel.)

k lépéssel előre tartó optimális prediktor

Előállítja $y(t+k)$ becslését az ismert információk birtokában:

$$\hat{y}(t+k|y_{-\infty}^t, u_{-\infty}^{t+k}) = \hat{y}(t+k|t)$$

miközben a becslési hiba várható értékét minimalizálja (LS becslő)

$$E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$$

Feladat

$\hat{y}(t+k|t)$ meghatározása.

Levezetés - k lépéssel előre tartó optimális prediktor

1. lépés

Kiindulás:

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$$

Kikötés:

- ❶ $G(q)$, $H(q)$ és $H^{-1}(q)$ szűrők (operátorok) stabilak
- ❷ Külső zajforrás van



Levezetés - k lépéssel előre tartó optimális prediktor

2. lépés

A $H(q) = 1 + h_1 q^{-1} + h_2 q^{-2} + \dots$ szűrő tagjait egy véges és egy végtelen összegre bontjuk fel:

$$\begin{aligned} H(q) &= H_k(q) + q^{-k} \tilde{H}_k(q) \\ H_k(q) &= 1 + h_1 q^{-1} + h_2 q^{-2} + \dots + h_{k-1} q^{-(k-1)} \\ \tilde{H}_k(q) &= h_k + h_{k+1} q^{-1} + h_{k+2} q^{-2} + \dots \end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

3. lépés

A rendszer kimenete a $t + k$ időpontban

$$\begin{aligned}
 y(t+k) &\stackrel{(1)}{=} G(q)u(t+k) + H(q)e(t+k) \\
 &\stackrel{(2)}{=} G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)e(t) \\
 &= G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]
 \end{aligned}$$

1. lépés felhasználása

$$\hat{y}(t) = G(\theta)u(t) + H(\theta)e(t)$$



Levezetés - k lépéssel előre tartó optimális prediktor

3. lépés

A rendszer kimenete a $t + k$ időpontban

$$\begin{aligned}
 y(t+k) &\stackrel{(1)}{=} G(q)u(t+k) + H(q)e(t+k) \\
 &\stackrel{(2)}{=} G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)e(t) \\
 &= G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]
 \end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

3. lépés

A rendszer kimenete a $t + k$ időpontban

$$\begin{aligned}
 y(t+k) &\stackrel{(1)}{=} G(q)u(t+k) + H(q)e(t+k) \\
 &\stackrel{(2)}{=} G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)e(t) \\
 &= G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]
 \end{aligned}$$

2. lépés felhasználása

A $H(q) = 1 + h_1q^{-1} + h_2q^{-2} + \dots$ szűrő tagjait egy véges és egy végtelen összegre bontjuk fel:

$$\begin{aligned}
 H(q) &= H_k(q) + q^{-k}\tilde{H}_k(q) \\
 H_k(q) &= 1 + h_1q^{-1} + h_2q^{-2} + \dots + h_{k-1}q^{-(k-1)} \\
 \tilde{H}_k(q) &= h_k + h_{k+1}q^{-1} + h_{k+2}q^{-2} + \dots
 \end{aligned}$$

Levezetés - k lépéssel előre tartó optimális prediktor

3. lépés

A rendszer kimenete a $t + k$ időpontban

$$\begin{aligned}
 y(t+k) &\stackrel{(1)}{=} G(q)u(t+k) + H(q)e(t+k) \\
 &\stackrel{(2)}{=} G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)e(t) \\
 &= G(q)u(t+k) + H_k(q)e(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]
 \end{aligned}$$

2. lépés felhasználása

A $H(q) = 1 + h_1q^{-1} + h_2q^{-2} + \dots$ szűrő tagjait egy véges és egy végtelen összegre bontjuk fel:

$$\begin{aligned}
 H(q) &= H_k(q) + q^{-k}\tilde{H}_k(q) \\
 H_k(q) &= 1 + h_1q^{-1} + h_2q^{-2} + \dots + h_{k-1}q^{-(k-1)} \\
 \tilde{H}_k(q) &= h_k + h_{k+1}q^{-1} + h_{k+2}q^{-2} + \dots
 \end{aligned}$$

Levezetés - k lépéssel előre tartó optimális prediktor

4. lépés

Az alábbi feltevésekkel élhetünk

- 1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek
- 2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,
- 3 $Ee(t+k) = 0$

1. lépés felhasználása

- $H_k(q)e(t+k) = e(t+k) + h_1e(t+k-1) + \dots + h_{k-1}e(t+1)$
összeg csak az e valószínűségi változó t utáni értékeit tartalmazza
- $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$
csak t -beli, vagy azt megelőző értékeket tartalmaz

Külső zajforrás esetén nincs kapcsolat/csatolás a két diszjunkt időintervallum között
 $\Rightarrow$ A két mennyiség független.

Levezetés - k lépéssel előre tartó optimális prediktor

4. lépés

Az alábbi feltevésekkel élhetünk

- 1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek
- 2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,
- 3 $Ee(t+k) = 0$

$H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek

Nyílt körben teljesül.

Zárt körben végrehajtott identifikációra nem teljesül, mert a kimeneti zaj vissza van csatolva a bemenetre.



Levezetés - k lépéssel előre tartó optimális prediktor

4. lépés

Az alábbi feltevésekkel élhetünk

- 1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek
- 2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,
- 3 $Ee(t+k) = 0$

$H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek

Nyílt körben teljesül.

Zárt körben végrehajtott identifikációra nem teljesül, mert a kimeneti zaj vissza van csatolva a bemenetre.

$Ee(t+k) = 0$

Fehér zajra teljesül.



Levezetés - k lépéssel előre tartó optimális prediktor

5. lépés

Később $E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$ számításához szükséges lesz $E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\}$ számításához.

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} \stackrel{(4.1),(4.2)}{=}$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}E\{H_k(q)e(t+k)\} \stackrel{(4.3)}{=} 0$$

4. lépés felhasználása

Az alábbi feltevésekkel élünk

4.1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek

4.2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,

4.3 $Ee(t+k) = 0$



Levezetés - k lépéssel előre tartó optimális prediktor

5. lépés

Később $E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$ számításához szükséges lesz $E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\}$ számításához.

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} \stackrel{(4.1),(4.2)}{=}$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}E\{H_k(q)e(t+k)\} \stackrel{(4.3)}{=} 0$$

4. lépés felhasználása

Az alábbi feltevésekkel élünk

4.1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek

4.2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,

4.3 $Ee(t+k) = 0$



Levezetés - k lépéssel előre tartó optimális prediktor

5. lépés

Később $E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$ számításához szükséges lesz $E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\}$ számításához.

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} \stackrel{(4.1),(4.2)}{=}$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}E\{H_k(q)e(t+k)\} \stackrel{(4.3)}{=} 0$$

4. lépés felhasználása

Az alábbi feltevésekkel élünk

4.1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek

4.2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,

4.3 $Ee(t+k) = 0$



Levezetés - k lépéssel előre tartó optimális prediktor

5. lépés

Később $E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$ számításához szükséges lesz $E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\}$ számításához.

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} \stackrel{(4.1),(4.2)}{=}$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}E\{H_k(q)e(t+k)\} \stackrel{(4.3)}{=} 0$$

4. lépés felhasználása

Az alábbi feltevésekkel élünk

4.1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek

4.2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,

4.3 $Ee(t+k) = 0$



Levezetés - k lépéssel előre tartó optimális prediktor

5. lépés

Később $E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$ számításához szükséges lesz $E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\}$ számításához.

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} \stackrel{(4.1),(4.2)}{=}$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}E\{H_k(q)e(t+k)\} \stackrel{(4.3)}{=} 0$$

4. lépés felhasználása

Az alábbi feltevésekkel élünk

4.1 $H_k(q)e(t+k)$ és $\tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]$ függetlenek

4.2 $H_k(q)e(t+k)$ és $G(q)u(t+k)$ függetlenek,

4.3 $Ee(t+k) = 0$



Levezetés - k lépéssel előre tartó optimális prediktor

6. lépés

Most pedig tekintsük a $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálendő kifejezést.

$$\begin{aligned}
 E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\
 &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}] \cdot [H_k(q)e(t+k)]\} + \\
 &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \\
 &E[H_k(q)e(t+k)]^2 + \\
 &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]] \cdot [H_k(q)e(t+k)]\} + \\
 &2E\hat{y} \cdot [H_k(q)e(t+k)]\} + \\
 &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2
 \end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

6. lépés

Most pedig tekintsük a $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálendő kifejezést.

$$\begin{aligned}
 E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\
 &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}] \cdot [H_k(q)e(t+k)]\} + \\
 &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \\
 &E[H_k(q)e(t+k)]^2 + \\
 &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]] \cdot [H_k(q)e(t+k)]\} + \\
 &2E\hat{y} \cdot [H_k(q)e(t+k)] + \\
 &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2
 \end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]] \cdot [H_k(q)e(t+k)]\} + \\ &2E\hat{y} \cdot [H_k(q)e(t+k)] + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

Magyarázat

$\hat{y}$ választása nem befolyásolja. Nem tudunk vele mit kezdeni a minimalizálás érdekében.



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &2E\{[G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]] \cdot [H_k(q)e(t+k)]\} + \\ &2E\hat{y} \cdot [H_k(q)e(t+k)] + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

5. lépés felhasználása

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} = 0$$



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &0 + \\ &2E\hat{y} \cdot [H_k(q)e(t+k)] + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

5. lépés felhasználása

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)]\}\{H_k(q)e(t+k)\} = 0$$



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &0 + \\ &2E\hat{y} \cdot [H_k(q)e(t+k)] + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

4. lépés felhasználása

$$4.3 \quad Ee(t+k) = 0$$



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &0 + \\ &0 + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

4. lépés felhasználása

$$4.3 \quad Ee(t+k) = 0$$



Levezetés - k lépéssel előre tartó optimális prediktor

7. lépés

Az $E[y(t+k) - \hat{y}(t+k|t)]^2$ minimalizálandó kifejezés új alakja.

$$\begin{aligned} E[y(t+k) - \hat{y}]^2 &= E[H_k(q)e(t+k)]^2 + \\ &0 + \\ &0 + \\ &+ E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \end{aligned}$$

Magyarázat

Ezt a tagot kell minimalizálni, azaz alkalmas $\hat{y}$ választással kinullázni.



Levezetés - k lépéssel előre tartó optimális prediktor

8. lépés

$$E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$$

$$\Updownarrow$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \rightarrow \min$$

$$\Updownarrow$$

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



Levezetés - k lépéssel előre tartó optimális prediktor

8. lépés

$$E[y(t+k) - \hat{y}(t+k|t)]^2 \rightarrow \min$$

$$\Updownarrow$$

$$E\{G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] - \hat{y}\}^2 \rightarrow \min$$

$$\Updownarrow$$

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



Levezetés - k lépéssel előre tartó optimális prediktor

9. lépés

$$\begin{aligned}
 \hat{y}(t+k|t) &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] \\
 &= G(q)u(t+k) - q^{-k}\tilde{H}_k(q)H^{-1}(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \\
 &= [1 - q^{-k}\tilde{H}_k(q)H^{-1}(q)]G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t)
 \end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

9. lépés

$$\begin{aligned}\hat{y}(t+k|t) &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] \\ &= G(q)u(t+k) - q^{-k}\tilde{H}_k(q)H^{-1}(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \\ &= [1 - q^{-k}\tilde{H}_k(q)H^{-1}(q)]G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t)\end{aligned}$$

Kihasználva, hogy...

$$u(t) = q^{-k}u(t+k)$$



Levezetés - k lépéssel előre tartó optimális prediktor

9. lépés

$$\begin{aligned}\hat{y}(t+k|t) &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] \\ &= G(q)u(t+k) - q^{-k}\tilde{H}_k(q)H^{-1}(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \\ &= [1 - q^{-k}\tilde{H}_k(q)H^{-1}(q)]G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t)\end{aligned}$$

Kihasználva, hogy...

$$u(t) = q^{-k}u(t+k)$$



Levezetés - k lépéssel előre tartó optimális prediktor

9. lépés

$$\begin{aligned}\hat{y}(t+k|t) &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] \\ &= G(q)u(t+k) - q^{-k}\tilde{H}_k(q)H^{-1}(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \\ &= [1 - q^{-k}\tilde{H}_k(q)H^{-1}(q)]G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t)\end{aligned}$$



Levezetés - k lépéssel előre tartó optimális prediktor

9. lépés

$$\begin{aligned}\hat{y}(t+k|t) &= G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)[y(t) - G(q)u(t)] \\ &= G(q)u(t+k) - q^{-k}\tilde{H}_k(q)H^{-1}(q)G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t) \\ &= [1 - q^{-k}\tilde{H}_k(q)H^{-1}(q)]G(q)u(t+k) + \tilde{H}_k(q)H^{-1}(q)y(t)\end{aligned}$$

A k lépéssel előre tartó optimális prediktor

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



k lépéssel előre tartó optimális prediktor

k lépéssel előre tartó optimális prediktor - Kompakt alak

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



1 lépéssel előre tartó optimális prediktor

1 lépéssel előre tartó optimális prediktor

$$k = 1$$

$$H_k(q) = 1,$$

$$\tilde{H}_k(q) = q[H(q) - 1],$$

$$\tilde{H}_k(q)H^{-1}(q) = q[1 - H^{-1}],$$

$$q^{-k}\tilde{H}_k(q)H^{-1}(q) = [1 - H^{-1}],$$

$$W_k(q) = H^{-1}(q)$$

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

◀ LS módszer ARX modellre



Az 1 lépéssel előre tartó predikció hibája

Az 1 lépéssel előre tartó predikció hibájának levezetése

$$\begin{aligned}\varepsilon(t) &= y(t) - \hat{y}(t|t-1) = \\ &= y(t) - H^{-1}(q)G(q)u(t) - [1 - H^{-1}(q)]y(t) = \\ &= H^{-1}(q)[y(t) - G(q)u(t)].\end{aligned}$$

Az 1 lépéssel előre tartó predikció hibája

$$\varepsilon(t) = H^{-1}(q)[y(t) - G(q)u(t)].$$



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszerei**
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modell esetén



A paraméterbecslés alapmódszerei

◀ Megoldás lépései



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapmódszerei**
 - **A legkisebb négyzetek módszere (LS) ARX modellen**
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



1 lépéses predikció ARX modellen

ARX modell

▶ ARX modell:

$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t)$$

Mindezt megfeleltetve a ▶ Modell kimenetnek

$$G(q) = \frac{B(q)}{A(q)}, \quad H(q) = \frac{1}{A(q)}.$$

▶ A kimenet optimális predikciója

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



1 lépéses predikció ARX modellen

ARX modell

▶ ARX modell:

$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t)$$

Mindezt megfeleltetve a **▶ Modell kimenetnek**

$$G(q) = \frac{B(q)}{A(q)}, \quad H(q) = \frac{1}{A(q)}.$$

▶ A kimenet optimális predikciója

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



1 lépéses predikció ARX modellen

ARX modell

▶ ARX modell:

$$y(t) = \frac{B(q)}{A(q)} u(t) + \frac{1}{A(q)} e(t)$$

Mindezt megfeleltetve a ▶ Modell kimenetnek

$$G(q) = \frac{B(q)}{A(q)}, \quad H(q) = \frac{1}{A(q)}.$$

▶ A kimenet optimális predikciója

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



1 lépéses predikció ARX modellen

Predikció ARX modellre

Az eddigi eredmény újra:

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

Legyen a mérési regisztrátumok vektora (n_a , n_b az A ill. B polinóm fokszáma):

$$\phi(t) = [-y(t-1) \dots - (y(t-n_a) \ u(t) \dots u(t-n_b+1))]^T$$

Legyen a paramétereket tartalmazó vektor:

$$\theta = (a_1 \dots a_{n_a} \ b_1 \dots b_{n_b})^T$$

Akkor a predikció új alakja:

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



1 lépéses predikció ARX modellen

Predikció ARX modellre

Az eddigi eredmény újra:

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

Legyen a mérési regisztrátumok vektora (n_a , n_b az A ill. B polinóm fokszáma):

$$\phi(t) = [-y(t-1) \dots - (y(t-n_a) \ u(t) \dots u(t-n_b+1))]^T$$

Legyen a paramétereket tartalmazó vektor:

$$\theta = (a_1 \dots a_{n_a} \ b_1 \dots b_{n_b})^T$$

Akkor a predikció új alakja:

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



1 lépéses predikció ARX modellen

Megjegyzés

$\hat{y}(t|\theta)$ felírásánál a paraméterekben való linearitás nem szükségképpen áll fenn minden rendszermodell esetén, ami [numerikus módszerek](#) alkalmazását követelheti meg.



1 lépéses predikció ARX modellen

Megjegyzés

$\hat{y}(t|\theta)$ felírásánál a paraméterekben való linearitás nem szükségképpen áll fenn minden rendszermodell esetén, ami [▶ numerikus módszerek](#) alkalmazását követelheti meg.



1 lépéses predikció ARX modellen

Megjegyzés

$\hat{y}(t|\theta)$ felírásánál a paraméterekben való linearitás nem szükségképpen áll fenn minden rendszermodell esetén, ami [numerikus módszerek](#) alkalmazását követelheti meg.



Optimalizálási feladat

Predikciós hiba (ϵ) és költségfüggvény (V_N)

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

$$\epsilon(t|\theta) = y(t) - \phi^T(t)\theta \quad V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \left[\frac{1}{2} (y(t) - \phi^T(t)\theta)^2 \right]$$

A minimum helyén a költségfüggvény helye θ szerinti deriváltja nullává válik.



Optimalizálási feladat

Predikciós hiba (ϵ) és költségfüggvény (V_N)

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

$$\epsilon(t|\theta) = y(t) - \phi^T(t)\theta \quad V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \left[\frac{1}{2} (y(t) - \phi^T(t)\theta)^2 \right]$$

A minimum helyén a költségfüggvény helye θ szerinti deriváltja nullává válik.

Szumman belüli tagot nézve csak (rögzített t)

$$\begin{aligned} f(\theta) &\equiv \left[\frac{1}{2} (y - \phi^T \theta) \right]^2 = \langle y(t) - \phi^T \theta, y(t) - \phi^T \theta \rangle \\ &= \langle y, y \rangle - 2 \langle y, \phi^T \theta \rangle + \langle \phi^T \theta, \phi^T \theta \rangle = \langle y, y \rangle - 2 \langle \phi y, \phi^T \rangle + \langle \phi \phi^T \theta, \theta \rangle \end{aligned}$$

$$\frac{\partial}{\partial \theta} f(\theta) = -2\phi y + 2\phi \phi^T \theta$$

Optimalizálási feladat

Predikciós hiba (ϵ) és költségfüggvény (V_N)

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

$$\epsilon(t|\theta) = y(t) - \phi^T(t)\theta \quad V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \left[\frac{1}{2} (y(t) - \phi^T(t)\theta)^2 \right]$$

A minimum helyén a költségfüggvény helye θ szerinti deriváltja nullává válik.

A költségfüggvény deriváltja

Az optimum helyén $V_N(\theta, Z^N)$ gradiense nulla:

$$\frac{\partial}{\partial \theta} V_N(\theta, Z^N) = -\frac{1}{N} \sum_{t=1}^N \phi(t)y(t) + \frac{1}{N} \sum_{t=1}^N \phi(t)\phi(t)^T \theta = \mathbf{0}$$

Ebből...



Optimalizálási feladat megoldása

Az optimális paramétervektor

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

Az optimális paramétervektor (alternatív felírás)

$$R(N) = \frac{1}{N} \sum_{t=1}^N \phi(t) \phi(t)^T \quad h(N) = \frac{1}{N} \sum_{t=1}^N \phi(t) y(t)$$

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N)$$

Elvárások az LS becslővel szemben

A valódi és optimális paramétervektor viszonya

Ha az adatokat θ_0 valódi paraméterhez tartozó zajos

$$y(t) = \phi^T(t)\theta_0 + \nu_0(t)$$

rendszer generálta (azaz ez lenne az optimális megoldás), akkor a valódi becslés eredménye:

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) [\phi^T(t)\theta_0 + \nu_0(t)] = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t)\nu_0$$

Elvárások a becslőtől

A $\hat{\theta}_N^{LS}$ becslőtől elvárjuk, hogy

- 1 legyen θ_0 közelében
- 2 konvergáljon θ_0 -hoz, ha N tart a végtelenhez.

A következőkben a konvergenciát vizsgáljuk...

Elvárások az LS becslővel szemben

A valódi és optimális paramétervektor viszonya

Ha az adatokat θ_0 valódi paraméterhez tartozó zajos

$$y(t) = \phi^T(t)\theta_0 + \nu_0(t)$$

rendszer generálta (azaz ez lenne az optimális megoldás), akkor a valódi becslés eredménye:

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) [\phi^T(t)\theta_0 + \nu_0(t)] = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t)\nu_0$$

Elvárások a becslőtől

A $\hat{\theta}_N^{LS}$ becslőtől elvárjuk, hogy

- 1 legyen θ_0 közelében
- 2 konvergáljon θ_0 -hoz, ha N tart a végtelenhez.

A következőkben a konvergenciát vizsgáljuk...

Elvárások az LS becslővel szemben

A valódi és optimális paramétervektor viszonya

Ha az adatokat θ_0 valódi paraméterhez tartozó zajos

$$y(t) = \phi^T(t)\theta_0 + \nu_0(t)$$

rendszer generálta (azaz ez lenne az optimális megoldás), akkor a valódi becslés eredménye:

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \left[\phi^T(t)\theta_0 + \nu_0(t) \right] = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t)\nu_0$$

Elvárások a becslőtől

A $\hat{\theta}_N^{LS}$ becslőtől elvárjuk, hogy

- 1 legyen θ_0 közelében
- 2 konvergáljon θ_0 -hoz, ha N tart a végtelenhez.

A következőkben a konvergenciát vizsgáljuk...

Feltevés a θ_0 határértékű konvergenciához

Feltevés

Feltesszük, hogy a jelek ergodikusak (a jel bármelyik elég hosszú reprezentációjának átlagai jól közelítik a valószínűségi átlagokat), akkor

$$R(N) \xrightarrow{N \rightarrow \infty} R^* = R_{\phi\phi} \qquad h(N) \xrightarrow{N \rightarrow \infty} h^* = R_{\phi\nu_0}$$

Miért fontos a feltevés?



Feltevés a θ_0 határértékű konvergenciához

Feltevés

Feltesszük, hogy a jelek ergodikusak (a jel bármelyik elég hosszú reprezentációjának átlagai jól közelítik a valószínűségi átlagokat), akkor

$$R(N) \xrightarrow{N \rightarrow \infty} R^* = R_{\phi\phi} \qquad h(N) \xrightarrow{N \rightarrow \infty} h^* = R_{\phi\nu_0}$$

Miért fontos a feltevés?



Konzisztencia feltételei

Emlékeztető: valódi és becült paramétervektor kapcsolata

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (1)$$

Ergodikus jelek esetén az LS becslés akkor konzisztens, azaz $\hat{\theta}_N \rightarrow \theta_0$, ha



Konzisztencia feltételei

Emlékeztető: valódi és becsült paramétervektor kapcsolata

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (1)$$

Ergodikus jelek esetén az LS becslés akkor konzisztens, azaz $\hat{\theta}_N \rightarrow \theta_0$, ha

LS1) R^* nonszinguláris

$R(N)$ csak ekkor invertálható az (3) egyenletben



Konzisztencia feltételei

Emlékeztető: valódi és becsült paramétervektor kapcsolata

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (1)$$

Ergodikus jelek esetén az LS becslés akkor konzisztens, azaz $\hat{\theta}_N \rightarrow \theta_0$, ha

LS2) $h^* = 0$

Ebben az esetben (3) egyenlet jobb oldalának második tagja eltűnik, ami szükséges $\hat{\theta}_N \rightarrow \theta_0$ feltétel teljesüléséhez. Mindez elérhető, ha

LS2.1) $\nu_0(t)$ fehér zaj nulla várható értékkel (ekkor $E\phi(t)\nu_0(t) = 0$ teljesül (3) kifejezésben), vagy



Konzisztencia feltételei

Emlékeztető: valódi és becsült paramétervektor kapcsolata

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (1)$$

Ergodikus jelek esetén az LS becslés akkor konzisztens, azaz $\hat{\theta}_N \rightarrow \theta_0$, ha

LS2) $h^* = 0$

Ebben az esetben (3) egyenlet jobb oldalának második tagja eltűnik, ami szükséges $\hat{\theta}_N \rightarrow \theta_0$ feltétel teljesüléséhez. Mindez elérhető, ha

LS2.2) $u(t)$ független $\nu_0(t)$ -tol, $E\nu_0 = 0$ és $na = 0$ (Ekkor $\phi(t)$ csak u tagokat tartalmaz.) A feltétel azért szükséges, mert ha $na > 0$ lenne, akkor $\phi(t)$ tartalmazza $y(t-1)$ -t, az pedig $\nu_0(t-1)$ értékét, ezért ha $\nu_0(t)$ nem fehér zaj, akkor általában $E\phi(t)\nu_0(t) \neq 0$, mivel $\nu_0(t-1)$ és $\nu_0(t)$ korreláltak. $E\phi(t)\nu_0(t) \neq 0$ pedig azt jelenti, hogy $\hat{\theta}_N \rightarrow \theta_0$ nem teljesül (3) -ben.

LS becslés problémája zárt körben

Emlékeztető: valódi és becsült paramétervektor kapcsolata

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (2)$$

Megjegyzés

Konzisztencia LS2.2) feltételnél $E\phi(t)\nu_0(t) = 0$ érdekében $\phi(t)$ csak u tagokat tartalmaz. Zárt körben végrehajtott identifikáció esetében azonban $\nu_0(t)$ a visszacsatolás révén hathat az u bemenetre ami $E\phi(t)\nu_0(t) \neq 0$ eredményezhet, azaz

$$\hat{\theta}_N^{LS} = [R(N)]^{-1} h(N) = \theta_0 + [R(N)]^{-1} \frac{1}{N} \sum_{t=1}^N \phi(t) \nu_0 \quad (3)$$

második tagja nem tűnik el és ezért $\hat{\theta}_N \rightarrow \theta_0$ nem teljesül.

LS becslés rossz numerikus kondicionáltsága a korreláció miatt

Megjegyzés

Mivel az LS becslésben szereplő $R(N)$ mátrix nagyméretű és rosszul kondicionált is lehet pl.

- állandósult állapotban végrehajtott identifikáció vagy
- a rosszul becsült struktúra fokszám miatt),

ezért $R(N)^{-1}$ számítása numerikus problémákat vethet fel, ezért gyakran használják a

$$\begin{bmatrix} \phi^T(1) \\ \vdots \\ \phi^T(N) \end{bmatrix}$$

mátrix különféle transzformációit (Cholesky-felbontás, QR felbontás, Householder transzformáció, SVD - szinguláris érték felbontás), amelyek numerikusan stabilak.

Hogyan csökkentsük a korrelációt? **IV4 módszerrel!**

LS becslés rossz numerikus kondicionáltsága a korreláció miatt

Megjegyzés

Mivel az LS becslésben szereplő $R(N)$ mátrix nagyméretű és rosszul kondicionált is lehet pl.

- állandósult állapotban végrehajtott identifikáció vagy
- a rosszul becsült struktúra foksza miatt),

ezért $R(N)^{-1}$ számítása numerikus problémákat vethet fel, ezért gyakran használják a

$$\begin{bmatrix} \phi^T(1) \\ \vdots \\ \phi^T(N) \end{bmatrix}$$

mátrix különféle transzformációit (Cholesky-felbontás, QR felbontás, Householder transzformáció, SVD - szinguláris érték felbontás), amelyek numerikusan stabilak.

Hogyan csökkentjük a korrelációt? **IV4 módszerrel!**

LS becslés rossz numerikus kondicionáltsága a korreláció miatt

Megjegyzés

Mivel az LS becslésben szereplő $R(N)$ mátrix nagyméretű és rosszul kondicionált is lehet pl.

- állandósult állapotban végrehajtott identifikáció vagy
- a rosszul becsült struktúra fokszám miatt),

ezért $R(N)^{-1}$ számítása numerikus problémákat vethet fel, ezért gyakran használják a

$$\begin{bmatrix} \phi^T(1) \\ \vdots \\ \phi^T(N) \end{bmatrix}$$

mátrix különféle transzformációit (Cholesky-felbontás, QR felbontás, Householder transzformáció, SVD - szinguláris érték felbontás), amelyek numerikusan stabilak.

Hogyan csökkentsük a korrelációt? **IV4 módszerrel!**



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapmódszerei**
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)**
 - Numerikus optimalizálás
 - Illusztráció: A gradiens meghatározása ARMAX modellen



IV4 módszer motivációja

- Cél: $E\phi(t)\nu_0(t) = 0$ azaz ϕ és ν_0 ne legyen korrelált (LS becslésnél könnyen korrelált lehet)

Tekintsük újra a $\hat{y}(t|\theta) = \phi^T(t)\theta$ lineáris regressziós modellt, amely esetén az LS becslés kifejezhető a

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

alakban ahol 'sol' a megoldás (solution) szóra utal.

Tegyük fel, hogy a korábbiakhoz hasonlóan $y(t) = \phi^T(t)\theta_0 + \nu_0(t)$, azonban most egy ξ segédváltozót (Instrumental Variables - IV) tartalmazó vektort vezetünk be, így az IV becslés alapja:

$$\begin{aligned} \theta_N^{IV} &= \text{sol} \left\{ \frac{1}{N} \sum_{i=1}^N \xi(t) \left[y(t) - \phi^T(t)\theta \right] = 0 \right\} \\ \hat{\theta}_N^{IV} &= \left[\frac{1}{N} \sum_{t=1}^N \xi(t)\phi(t)^T \right]^{-1} \frac{1}{N} \sum_{t=1}^N \xi(t)y(t) \end{aligned} \quad (4)$$



IV4 módszer motivációja

- Cél: $E\phi(t)\nu_0(t) = 0$ azaz ϕ és ν_0 ne legyen korrelált
(LS becslésnél könnyen korrelált lehet)

Tekintsük újra a $\hat{y}(t|\theta) = \phi^T(t)\theta$ lineáris regressziós modellt, amely esetén az LS becslés kifejezhető a

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

alakban ahol 'sol' a megoldás (solution) szóra utal.

Tegyük fel, hogy a korábbiakhoz hasonlóan $y(t) = \phi^T(t)\theta_0 + \nu_0(t)$, azonban most egy ξ segédváltozót (Instrumental Variables - IV) tartalmazó vektort vezetünk be, így az IV becslés alapja:

$$\begin{aligned} \theta_N^{IV} &= \text{sol} \left\{ \frac{1}{N} \sum_{i=1}^N \xi(t) \left[y(t) - \phi^T(t)\theta \right] = 0 \right\} \\ \hat{\theta}_N^{IV} &= \left[\frac{1}{N} \sum_{t=1}^N \xi(t)\phi(t)^T \right]^{-1} \frac{1}{N} \sum_{t=1}^N \xi(t)y(t) \end{aligned} \quad (4)$$



IV4 módszer motivációja

- Cél: $E\phi(t)\nu_0(t) = 0$ azaz ϕ és ν_0 ne legyen korrelált (LS becslésnél könnyen korrelált lehet)

Tekintsük újra a $\hat{y}(t|\theta) = \phi^T(t)\theta$ lineáris regressziós modellt, amely esetén az LS becslés kifejezhető a

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

alakban ahol 'sol' a megoldás (solution) szóra utal.

Tegyük fel, hogy a korábbiakhoz hasonlóan $y(t) = \phi^T(t)\theta_0 + \nu_0(t)$, azonban most egy ξ segédváltozót (Instrumental Variables - IV) tartalmazó vektort vezetünk be, így az IV becslés alapja:

$$\begin{aligned} \theta_N^{IV} &= \text{sol} \left\{ \frac{1}{N} \sum_{i=1}^N \xi(t) \left[y(t) - \phi^T(t)\theta \right] = 0 \right\} \\ \hat{\theta}_N^{IV} &= \left[\frac{1}{N} \sum_{t=1}^N \xi(t)\phi(t)^T \right]^{-1} \frac{1}{N} \sum_{t=1}^N \xi(t)y(t) \end{aligned} \quad (4)$$



IV4 módszer motivációja

Az LS becslés (LS1)-LS2) feltételeihez hasonlóan, $\hat{\theta}_N \rightarrow \theta_0$ eléréséhez két feltételnek kell teljesülnie:

IV1) $E\xi(t)\phi^T(t)$ nonszinguláris,

IV2) $E\xi(t)\nu_0(t) = 0$ (ami az LS becslés IV módszerre alkalmazva a $1/N \sum_{i=1}^N \xi(t)\nu_0(t) \rightarrow 0$ feltétel teljesítéséhez szükséges).

Tekintsük ezután a segédváltozók megválasztását ARX modell esetén, amikor

$$y(t) + a_1 \cdot y(t-1) + \dots + a_{na} \cdot y(t-na) = b_1 \cdot u(t) + \dots + b_{nb} \cdot u(t-nb+1) + \nu(t),$$

$$\theta = (a_1, \dots, a_{na}, b_1, \dots, b_{nb})^T,$$

$$\phi(t) = (-y(t-1), \dots, -y(t-na), u(t), \dots, u(t-nb+1))^T,$$

$$y(t) = \phi^T(t)\theta + \nu.$$



IV4 módszer motivációja

Az IV módszer koncepciója. A valódi rendszerhez tartozzon θ_0 és ν_0 .

A koncepció ötlete, hogy alkalmasan választott $K(q)$, $N(q)$, $M(q)$ szűrőkkel olyan ξ segédváltozókat definiálunk, amely képes kielégíti IV1) és IV2) feltételeket.

A segédváltozók előállítása a következőképp történik:

$$N(q) = 1 + n_1 \cdot q^{-1} + \dots + n_{na} \cdot q^{-na} \quad (5)$$

$$M(q) = m_1 + m_2 \cdot q^{-1} + \dots + m_{nb} \cdot q^{-nb+1} \quad (6)$$

$$N(q)x(t) = M(q)u(t) \quad (7)$$

$$\xi(t) = K(q)[x(t), \dots, x(t-na), u(t), \dots, u(t-nb+1)]^T \quad (8)$$

ahol $N(q)x(t) = M(q)u(t)$ egy differenciaegyenlet, amelyből $x(t)$ kifejezhető és rekurzívan számítható.

A koncepcióban $\xi(t) = \xi(t, u(\cdot))$, ezért felnyitott (!) körben generált $u(t)$ és külső gerjesztésű zaj esetén teljesül IV2) feltétel: $E\xi(t)\nu_0(t) = 0$.

Mivel ϕ és ξ ugyanabból a bemenetből lett generálva (azonban ξ már nem tartalmazza ν_0 hatását), ezért elvárható, hogy IV1) feltétel is teljesül, azaz $E\xi(t)\phi^T(t)$ nonsinguláris és épp ezért invertálható.



IV4 módszer motivációja

Az IV módszer koncepciója. A valódi rendszerhez tartozzon θ_0 és ν_0 .

A koncepció ötlete, hogy alkalmasan választott $K(q)$, $N(q)$, $M(q)$ szűrőkkel olyan ξ segédváltozókat definiálunk, amely képes kielégíti IV1) és IV2) feltételeket.

A segédváltozók előállítása a következőképp történik:

$$N(q) = 1 + n_1 \cdot q^{-1} + \dots + n_{na} \cdot q^{-na} \quad (5)$$

$$M(q) = m_1 + m_2 \cdot q^{-1} + \dots + m_{nb} \cdot q^{-nb+1} \quad (6)$$

$$N(q)x(t) = M(q)u(t) \quad (7)$$

$$\xi(t) = K(q)[x(t), \dots, x(t-na), u(t), \dots, u(t-nb+1)]^T \quad (8)$$

ahol $N(q)x(t) = M(q)u(t)$ egy differenciaegyenlet, amelyből $x(t)$ kifejezhető és rekurzívan számítható.

A koncepcióban $\xi(t) = \xi(t, u(\cdot))$, ezért felnyitott (!) körben generált $u(t)$ és külső gerjesztésű zaj esetén teljesül IV2) feltétel: $E\xi(t)\nu_0(t) = 0$.

Mivel ϕ és ξ ugyanabból a bemenetből lett generálva (azonban ξ már nem tartalmazza ν_0 hatását), ezért elvárható, hogy IV1) feltétel is teljesül, azaz $E\xi(t)\phi^T(t)$ nonszinguláris és épp ezért invertálható.



IV4 módszer motivációja

Az IV módszer koncepciója. A valódi rendszerhez tartozzon θ_0 és ν_0 .

A koncepció ötlete, hogy alkalmasan választott $K(q)$, $N(q)$, $M(q)$ szűrőkkel olyan ξ segédváltozókat definiálunk, amely képes kielégíti IV1) és IV2) feltételeket.

A segédváltozók előállítása a következőképp történik:

$$N(q) = 1 + n_1 \cdot q^{-1} + \dots + n_{na} \cdot q^{-na} \quad (5)$$

$$M(q) = m_1 + m_2 \cdot q^{-1} + \dots + m_{nb} \cdot q^{-nb+1} \quad (6)$$

$$N(q)x(t) = M(q)u(t) \quad (7)$$

$$\xi(t) = K(q)[x(t), \dots, x(t - na), u(t), \dots, u(t - nb + 1)]^T \quad (8)$$

ahol $N(q)x(t) = M(q)u(t)$ egy differenciaegyenlet, amelyből $x(t)$ kifejezhető és rekurzívan számítható.

A koncepcióban $\xi(t) = \xi(t, u(\cdot))$, ezért felnyitott (!) körben generált $u(t)$ és külső gerjesztésű zaj esetén teljesül IV2) feltétel: $E\xi(t)\nu_0(t) = 0$.

Mivel ϕ és ξ ugyanabból a bemenetből lett generálva (azonban ξ már nem tartalmazza ν_0 hatását), ezért elvárható, hogy IV1) feltétel is teljesül, azaz $E\xi(t)\phi^T(t)$ nonsinguláris és épp ezért invertálható.



IV4 módszer motivációja

Az IV módszer koncepciója. A valódi rendszerhez tartozzon θ_0 és ν_0 .

A koncepció ötlete, hogy alkalmasan választott $K(q)$, $N(q)$, $M(q)$ szűrőkkel olyan ξ segédváltozókat definiálunk, amely képes kielégíteni IV1) és IV2) feltételeket.

A segédváltozók előállítása a következőképp történik:

$$N(q) = 1 + n_1 \cdot q^{-1} + \dots + n_{na} \cdot q^{-na} \quad (5)$$

$$M(q) = m_1 + m_2 \cdot q^{-1} + \dots + m_{nb} \cdot q^{-nb+1} \quad (6)$$

$$N(q)x(t) = M(q)u(t) \quad (7)$$

$$\xi(t) = K(q)[x(t), \dots, x(t-na), u(t), \dots, u(t-nb+1)]^T \quad (8)$$

ahol $N(q)x(t) = M(q)u(t)$ egy differenciaegyenlet, amelyből $x(t)$ kifejezhető és rekurzívan számítható.

A koncepcióban $\xi(t) = \xi(t, u(\cdot))$, ezért felnyitott (!) körben generált $u(t)$ és külső gerjesztésű zaj esetén teljesül IV2) feltétel: $E\xi(t)\nu_0(t) = 0$.

Mivel ϕ és ξ ugyanabból a bemenetből lett generálva (azonban ξ már nem tartalmazza ν_0 hatását), ezért elvárható, hogy IV1) feltétel is teljesül, azaz $E\xi(t)\phi^T(t)$ nonsinguláris és épp ezért invertálható.



IV4 módszer motivációja

Az IV módszer koncepciója. A valódi rendszerhez tartozzon θ_0 és ν_0 .

A koncepció ötlete, hogy alkalmasan választott $K(q)$, $N(q)$, $M(q)$ szűrőkkel olyan ξ segédváltozókat definiálunk, amely képes kielégíteni IV1) és IV2) feltételeket.

A segédváltozók előállítása a következőképp történik:

$$N(q) = 1 + n_1 \cdot q^{-1} + \dots + n_{na} \cdot q^{-na} \quad (5)$$

$$M(q) = m_1 + m_2 \cdot q^{-1} + \dots + m_{nb} \cdot q^{-nb+1} \quad (6)$$

$$N(q)x(t) = M(q)u(t) \quad (7)$$

$$\xi(t) = K(q)[x(t), \dots, x(t-na), u(t), \dots, u(t-nb+1)]^T \quad (8)$$

ahol $N(q)x(t) = M(q)u(t)$ egy differenciaegyenlet, amelyből $x(t)$ kifejezhető és rekurzívan számítható.

A koncepcióban $\xi(t) = \xi(t, u(\cdot))$, ezért felnyitott (!) körben generált $u(t)$ és külső gerjesztésű zaj esetén teljesül IV2) feltétel: $E\xi(t)\nu_0(t) = 0$.

Mivel ϕ és ξ ugyanabból a bemenetből lett generálva (azonban ξ már nem tartalmazza ν_0 hatását), ezért elvárható, hogy IV1) feltétel is teljesül, azaz $E\xi(t)\phi^T(t)$ nonszinguláris és épp ezért invertálható.



IV4 módszer motivációja

Az IV módszer elve

$$\begin{aligned}\xi(t, \theta) &= K_u(q, \theta)u(t) \\ \epsilon_F(t, \theta) &= L(q)[y(t) - \phi^T(t, \theta)\theta], \\ f_N(\theta, Z^N) &= \frac{1}{N} \sum_{t=1}^N \xi(t, \theta)\epsilon_F(t, \theta), \\ \theta_N^{IV} &= \text{sol}_{\theta \in D_M} [f_N(\theta, Z^N = 0)]\end{aligned}\tag{9}$$



IV4 módszer motivációja

Megjegyzések

- 1 A becslés elve hasonló az LS becsléshez, azonban a szükséges helyen $\phi(t)$ -t a $\xi(t, \theta)$ segédváltozóra cseréltük
- 2 az $u(t)$ bemeneti jelet és az $y(t)$ kimeneti jelet ugyanazon $L(q)$ szűrőn küldjük át (vagyis a zajmodellt az általános összefüggésben lecseréljük $H(q, \theta)$ -ról $L^{-1}(q)H(q)$ -ra). Emögött az húzódik meg, hogy a jelek „kisimításával” a zaj kevésbé szól bele $u(t)$ és $y(t)$ közötti kapcsolatba, azaz a két jel jobban korrelál, a zaj nélküli rendszer bemenete és kimenete közötti kapcsolat jobban érvényesül azonosítás során, tehát az $y(t)$ kimenetnek zaj által okozott ingadozása kevésbé fogja „rángatni” a paramétereket.



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapszerei**
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás**
 - Illusztráció: A gradiens meghatározása ARMAX modellen



Optimalizálás kvadratikus közelítéssel

- Az identifikáció felfogható egy optimalizálási problémának is a paramétertérben
- Az optimalizálandó célfüggvény az optimum közelében jól közelíthető egy konvex kvadratikus kifejezéssel
 θ helyett x , V_N helyett $f(x_k)$ függvényt tekintjük.
- Az $f(x)$ konvex kvadratikus függvény optimumkereséséhez keresünk módszert

Kvadratikus függvény optima

Legyen az optimum helye x_0

$$f(x) = \frac{1}{2} \langle A(x - x_0, x - x_0) \rangle, \quad A \geq 0,$$

$$x_0 = \arg \min f(x),$$

$$x_0 = \text{sol } [f(x) = 0]$$

Optimalizálás kvadratikus közelítéssel

- Az identifikáció felfogható egy optimalizálási problémának is a paramétertérben
- Az optimalizálandó célfüggvény az optimum közelében jól közelíthető egy konvex kvadratikus kifejezéssel
 θ helyett x , V_N helyett $f(x_k)$ függvényt tekintjük.
- Az $f(x)$ konvex kvadratikus függvény optimumkereséséhez keresünk módszert

Kvadratikus függvény optima

Legyen az optimum helye x_0

$$f(x) = \frac{1}{2} \langle A(x - x_0, x - x_0) \rangle, \quad A \geq 0,$$

$$x_0 = \arg \min f(x),$$

$$x_0 = \text{sol } [f(x) = 0]$$

Optimalizálás kvadratikus közelítéssel

- Az identifikáció felfogható egy optimalizálási problémának is a paramétertérben
- Az optimalizálandó célfüggvény az optimum közelében jól közelíthető egy konvex kvadratikus kifejezéssel
 θ helyett x , V_N helyett $f(x_k)$ függvényt tekintjük.
- Az $f(x)$ konvex kvadratikus függvény optimumkereséséhez keresünk módszert

Kvadratikus függvény optima

Legyen az optimum helye x_0

$$f(x) = \frac{1}{2} \langle A(x - x_0, x - x_0) \rangle, \quad A \geq 0,$$

$$x_0 = \arg \min f(x),$$

$$x_0 = \text{sol } [f(x) = 0]$$

Optimalizálás kvadratikus közelítéssel

- Az identifikáció felfogható egy optimalizálási problémának is a paramétertérben
- Az optimalizálandó célfüggvény az optimum közelében jól közelíthető egy konvex kvadratikus kifejezéssel
 θ helyett x , V_N helyett $f(x_k)$ függvényt tekintjük.
- Az $f(x)$ konvex kvadratikus függvény optimumkereséséhez keresünk módszert

Kvadratikus függvény optima

Legyen az optimum helye x_0

$$f(x) = \frac{1}{2} \langle A(x - x_0, x - x_0) \rangle, \quad A \geq 0,$$

$$x_0 = \arg \min f(x),$$

$$x_0 = \text{sol } [f(x) = 0]$$

Newton-módszer

Az optimum hely, az első derivált (gradiens) és a második derivált (Hess-mátrix) kapcsolata

Az optimumhoz kell a derivált:

$$f'(x) = A(x - x_0)$$

$$f''(x) = A.$$

amiből láthatóan tetszőleges x esetén fennáll

$$x_0 = x - A^{-1} [A(x - x_0)]$$

Azaz

Newton-módszer

$$x_0 = x - [f''(x)]^{-1} f'(x)$$

Newton-módszer

Az optimum hely, az első derivált (gradiens) és a második derivált (Hess-mátrix) kapcsolata

Az optimumhoz kell a derivált:

$$f'(x) = A(x - x_0)$$

$$f''(x) = A.$$

amiből láthatóan tetszőleges x esetén fennáll

$$x_0 = x - A^{-1} [A(x - x_0)]$$

Azaz

Newton-módszer

$$x_0 = x - [f''(x)]^{-1} f'(x)$$

Newton-módszer

Az optimum hely, az első derivált (gradiens) és a második derivált (Hess-mátrix) kapcsolata

Az optimumhoz kell a derivált:

$$f'(x) = A(x - x_0)$$

$$f''(x) = A.$$

amiből láthatóan tetszőleges x esetén fennáll

$$x_0 = x - A^{-1} [A(x - x_0)]$$

Azaz

Newton-módszer

$$x_0 = x - [f''(x)]^{-1} f'(x)$$

Iteratív Newton-módszer

Újra: Newton-módszer

$$x_0 = x - [f''(x)]^{-1} f'(x)$$

Megjegyzés

Newton-módszer szerint tetszőleges konvex és kvadratikus valós függvény globális minimumát tetszőleges helyről indítva egy lépésben meg lehet határozni.

Ha a függvény nem kvadratikus, akkor a módszer csak egy közelítő (iteratív) eljárás:

Iteratív Newton-módszer

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$



Iteratív Newton-módszer

Újra: Newton-módszer

$$x_0 = x - [f''(x)]^{-1} f'(x)$$

Megjegyzés

Newton-módszer szerint tetszőleges konvex és kvadratikus valós függvény globális minimumát tetszőleges helyről indítva egy lépésben meg lehet határozni.

Ha a függvény nem kvadratikus, akkor a módszer csak egy közelítő (iteratív) eljárás:

Iteratív Newton-módszer

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$



Kvázi Newton-módszer

Újra: Iteratív Newton-módszer

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$

Probléma: Newton módszer igényli a Hess-féle mátrix meghatározását is, ami gyakorlatban sokszor nehézkes vagy nem kivitelezhető, ezért

Kvázi Newton-módszer

Az Hess-féle mátrixnak csak egy $H(x_k)$ közelítését használjuk

$$x_{k+1} = x_k - [H(x_k)]^{-1} f'(x_k)$$



Kvázi Newton-módszer

Újra: Iteratív Newton-módszer

$$x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$$

Probléma: Newton módszer igényli a Hess-féle mátrix meghatározását is, ami gyakorlatban sokszor nehézkes vagy nem kivitelezhető, ezért

Kvázi Newton-módszer

Az Hess-féle mátrixnak csak egy $H(x_k)$ közelítését használjuk

$$x_{k+1} = x_k - [H(x_k)]^{-1} f'(x_k)$$



Levenberg-Marquardt-eljárás

Újra: Kvázi Newton-módszer

$$x_{k+1} = x_k - [H(x_k)]^{-1} f'(x_k)$$

Problémát jelenthet, ha $H(x)$ (vagy $f''(x)$) szingulárisává válik (azaz nem invertálható). Ennek kiküszöbölésére gyakran használt regularizációs technika a...

Levenberg-Marquardt-eljárás

$$H(x_k) := H(x_k) + \delta I$$

ahol δ kis pozitív szám biztosítja az invertálhatóságot az optimum közelében is.



Levenberg-Marquardt-eljárás

Újra: Kvázi Newton-módszer

$$x_{k+1} = x_k - [H(x_k)]^{-1} f'(x_k)$$

Problémát jelenthet, ha $H(x)$ (vagy $f''(x)$) szingulárisává válik (azaz nem invertálható). Ennek kiküszöbölésére gyakran használt regularizációs technika a...

Levenberg-Marquardt-eljárás

$$H(x_k) := H(x_k) + \delta I$$

ahol δ kis pozitív szám biztosítja az invertálhatóságot az optimum közelében is.



Összefoglalva a numerikus módszereket

A numerikus minimalizálási módszerek láthatóan az $x_{k+1} = x_k + \alpha f_k$ alakban írhatók fel, ahol α a lépésköz, f_k a keresési irány.

Az f_k a keresési irány meghatározásának módjától függően a következőképp csoportosíthatók:

- 1 Függvényértéket, gradienst és Hess-féle (második derivált) mátrixot használó módszerek. Ennek speciális esete a Newton-módszer, ahol $f_k = [f''(x_k)]^{-1} f'(x_k)$.
- 2 Függvényértéket és gradienst használó módszerek. Ennek speciális és fontos esete a kvázi Newton-módszer, ahol a Hess-féle mátrixot csak közelítjük.
- 3 Csak függvényértéket használó módszerek. Ezek az algoritmusok többnyire differencia módszerrel közelítik a gradienst és valamilyen kvázi Newton-módszerrel folytatják vagy pedig speciális keresési mintákat használnak.

A lépésköz nagy megválasztásánál hamar eljutunk az optimum közelébe, azonban ott oszcillálni kezd a keresés, míg kis lépésköz mellett sok időt vesz igénybe az optimum megtalálása.



Tartalom

- 1 Az előző részek tartalmából
- 2 Bevezetés
- 3 Stabil, lineáris operátorok a sorozatok terében
- 4 Identifikáció valószínűségi számítási alapjai
 - Valószínűségi változók
 - Sztochasztikus folyamatok
- 5 Diszkrétidejű lineáris rendszermodellek zajos környezetben
- 6 A modell kimenetének optimális predikciója
- 7 A paraméterbecslés alapmódszerei**
 - A legkisebb négyzetek módszere (LS) ARX modellen
 - Segédváltozók módszere (IV4)
 - Numerikus optimalizálás
 - **Illusztráció: A gradiens meghatározása ARMAX modell esetén**



Illusztráció: A gradiens meghatározása ARMAX modell esetén

A jelölések egyszerűsítése érdekében a θ paraméter kiírását átmenetileg szüneteltetjük. Az ARMAX modell esetén a korábbiak alapján:

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

$$y(t) = \frac{B(q)}{A(q)}u(t) + \frac{C(q)}{A(q)}e(t) = G(q)u(t) + H(q)e(t),$$

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

$$H^{-1}(q)G(q) = \frac{A(q)}{C(q)} \cdot \frac{B(q)}{A(q)} = \frac{B(q)}{C(q)},$$

$$1 - H^{-1}(q) = 1 - \frac{A(q)}{C(q)}$$

$$\hat{y}(t) = \frac{B(q)}{C(q)}u(t) + \left[1 - \frac{A(q)}{C(q)}\right] y(t),$$



Illusztráció: A gradiens meghatározása ARMAX modell esetén

$$\hat{y}(t) = \frac{B(q)}{C(q)}u(t) + 1 - \left[1 - \frac{A(q)}{C(q)}\right]y(t),$$

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

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

$$C(q)\frac{\partial}{\partial b_k}\hat{y}(t) = -q^{-(k-1)}u(t),$$

A deriválás szorzatszabálya alapján

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

$$C(q)\frac{\partial}{\partial c_k}\hat{y}(t) = q^{-k}[y(t) - \hat{y}(t)] = q^{-k}\epsilon(t)$$



Illusztráció: A gradiens meghatározása ARMAX modell esetén

Bevezetve a

$$\phi = [-y(t-1), \dots, -y(t-na), u(t), \dots, u(t-nb+1), \epsilon(t-1), \dots, \epsilon(t-nc)]^T \quad (10)$$

regressziós vektort, a gradiens a numerikus optimalizáláshoz a korábbi

$$C(q) \frac{\partial}{\partial ak} \hat{y}(t) = -q^{-k} y(t)$$

$$C(q) \frac{\partial}{\partial bk} \hat{y}(t) = -q^{-(k-1)} u(t), q^{-k} \hat{y}(t) + C(q) \frac{\partial}{\partial ck} \hat{y}(t) = q^{-k} y(t) \Rightarrow$$

$$C(q) \frac{\partial}{\partial ck} \hat{y}(t) = q^{-k} [y(t) - \hat{y}(t)] = q^{-k} \epsilon(t)$$

egyenletekből:

$$\psi(t) := \frac{1}{C(q)} \phi(t) \quad (11)$$



Illusztráció: A gradiens meghatározása ARMAX modell esetén

Összefoglalva: A gradiens számítása a

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

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

$$\phi = [-y(t-1), \dots, -y(t-na), u(), \dots, u(t-nb+1), \epsilon(t-1), \dots, \epsilon(t-nc)]^T$$

$$\psi(t) := \frac{1}{C(q)}\phi(t) \rightarrow \psi(t, \theta)$$

sorrendben történik az aktuális lépésben érvényes θ paraméter értékénél. Az $\hat{y}(t|\theta)$ és $\psi(t, \theta)$ számításához differenciaegyenleteket kell megoldani.

