Beágyazott információs rendszerek

5. Mennyiségek, változók valós idejű rendszerekben

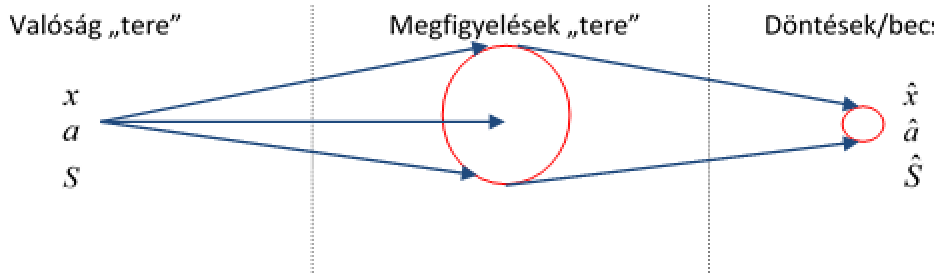
2020. október 22.

5. Mennyiségek, változók valós idejű rendszerekben

Előzmények: Real-time változó, real-time változó képe, időbeni pontosság, periodikus frissítés, megfigyelés: állapotmegfigyelés, eseménymegfigyelés, valós idejű objektum, állandóság, akciókésleltetés, idempotencia, hihetőség (bizánci típusú hibák), befolyásolhatóság, …

A befogadó környezet modellezése:

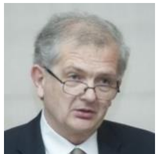
Ami nem megkerülhető: A befogadó környezet **megismerése**, majd a működtető **szoftverbe építése!** **Ennek eszköze a mérési eljárás:** a megismerési folyamat része, amelynek során a rendelkezésünkre álló ismereteinket pontosítjuk, ill. bővítjük.



Az ábra a folyamat interpretálását segíti. A mérés során a **valóság** jelenségeit szeretnénk **megragadni**. Ezt a „megragadást” előszeretettel végezzük olyan jellemzőkre építve, amelyek valamilyen értelemben **stabilitást** mutatnak. Ilyen jellemzőkhöz (is) **absztrakció** révén jutunk.

- az állapotváltozók ($\boldsymbol{x}$), amelyek változásai a kölcsönhatások révén fellépő energia-folyamatokhoz köthetők (feszültség, nyomás, hőmérséklet, sebesség, stb.)
- a paraméterek ($\boldsymbol{a}$), amelyek a kölcsönhatások intenzitásviszonyait ragadják meg, és
- a struktúrák ($\boldsymbol{S}$), amelyek a rendszer-komponensek kapcsolatait írják le.

A valóság „tere” egy olyan absztrakció, amelyben a vizsgált jellemzők konkrét értékei a tér egy pontjának felelnek meg.



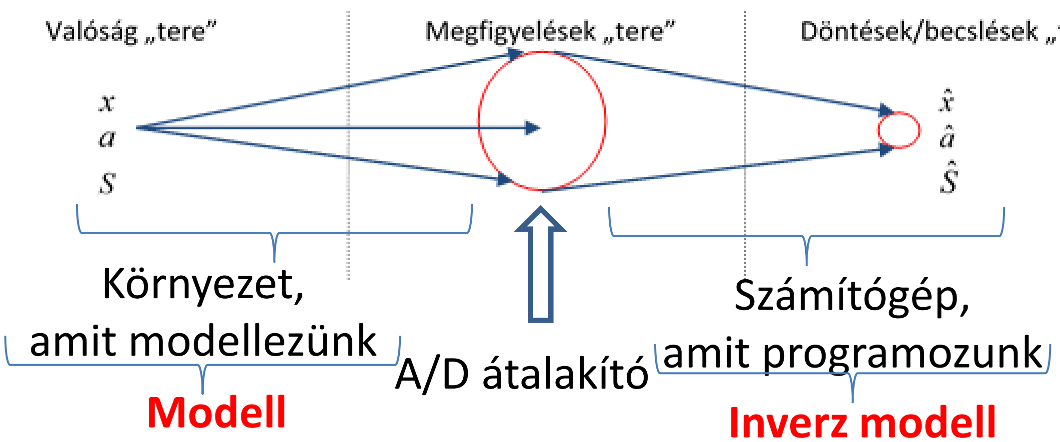
A mérés előtt a pont koordinátáit **nem ismerjük**.

A mérések során egy-egy ilyen pont koordinátáinak meghatározására (megmérésére) törekszünk, ami – ismert módon – **csak közelítőleg** lehetséges (a mérés hibával terhelt).

További nehézség, hogy a **mérendő mennyiséghez** sok esetben nem férünk közvetlenül hozzá, ezért többnyire csak valamilyen leképzéséből tudunk kiindulni.

Ezt a leképzést nevezzük **megfigyelés**nek.

A mérendő és a megfigyelt érték közötti út a **mérési/jelátviteli csatorna**.



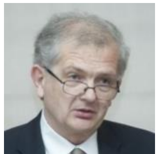
Megfigyelés determinisztikus csatorna esetén:

A megfigyelt “valóságot” autonóm rendszerként képzeljük el, és diszkrét modellel írjuk le. A „valóságot” és a megfigyelést leíró állapot, ill. megfigyelési egyenletek:

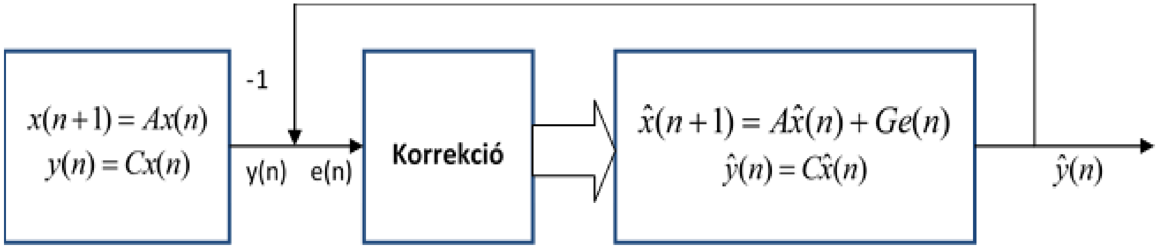
$$\boxed{x(n+1) = Ax(n),} \quad dim[x(n)] = N, \quad dim[A] = N * N$$

$$\boxed{y(n) = Cx(n)} \quad dim[y(n)] = M, \quad M \le N, \; dim[C] = M * N$$

Tudunk egy ilyen modellt invertálni? **Általában nem!** Mikor tudunk? Ha C^{-1} létezik!



Ha **nem létezik** C^{-1}, akkor mást kell csinálni! A megoldás egy **szimulátor!**
Ehhez a modellnek **be kell épülnie** a számítógépbe! Ezt vezéreljük a **megfigyelt értékekkel**!



Célunk az $x(n)$ állapotvektor becslése.

Ennek az eszköznek a neve: **megfigyelő**, amely a „valóság" **másolata** igyekszik lenni azáltal, hogy egy korrekciós/tanuló/adaptáló mechanizmus eredményeképpen **követi** azt.

A követés bekövetkeztével a mérés „eredménye" $\hat{x}(n)$ a megfigyelőből olvasható ki.

A megfigyelőben megvalósuló „**másolat**" állapot, ill. megfigyelési egyenletei:

$$\boxed{\hat{x}(n+1) = A\hat{x}(n) + Ge(n)}$$

$$\boxed{\hat{y}(n) = C\hat{x}(n)}$$

G: **korrekciós mátrix**; $dim[G] = N * M$, $e(n) = y(n) - \hat{y}(n)$

A G mátrixot úgy tervezzük meg, hogy $\boxed{\hat{x}(n) \rightarrow x(n).}$

A minimalizálandó állapothiba az állapotegyenletek különbségeként:

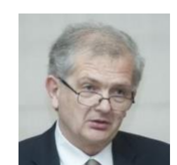
$$\underbrace{x(n+1) - \hat{x}(n+1)}_{\varepsilon(n+1)} = Ax(n) - A\hat{x}(n) - Ge(n) = \underbrace{(A - GC)}_{F}\underbrace{(x(n) - \hat{x}(n))}_{\varepsilon(n)}.$$

$$\boxed{\varepsilon(n+1) = F\varepsilon(n)}$$

az ún. hibarendszer állapotátmenet mátrixa.

A G mátrix tervezése: $\boxed{\varepsilon(n) \xrightarrow[n\to\infty]{} 0,}$ aminek érdekében $\varepsilon(n+1) < \varepsilon(n)$, lehetőleg $\forall\ n$-re.

F csökkenti $\varepsilon(n)$ hosszát minden lépésben: „**kontraktív**". Az $\varepsilon(n)$ hibavektorral kapcsolatos egyenlőtlenség a vektor hosszára (normájára) értelmezendő, skalár esetben pedig a hiba abszolút értékére. **Elvárás:** a hibarendszer a belső energiáját **disszipálja.**



Emlékeztető

Esetek:

1. $F = A - GC = 0$. Ebben az esetben $G = AC^{-1}$. Ez akkor lehetséges, ha C négyzetes, azaz a megfigyelés éppen annyi komponensű, mint maga az állapotvektor.

Az egyenlet explicit formában megoldható! A konvergencia egyetlen lépéses!

2. $F^N = (A - GC)^N = 0$. Ebben az esetben a hibarendszer N lépésben konvergál:

$$\boxed{x(N) - \hat{x}(N) = (A - GC)^N (x(0) - \hat{x}(0)) = 0}$$

Az $F^N = 0$ tulajdonságú mátrixok, az ún. **nemderogatórius nilpotens** mátrixok, amelyek sajátja, hogy **valamennyi sajátértékük nulla**. Az ilyen tulajdonságú állapotátmenet mátrixszal jellemezhető rendszerek **véges impulzusválaszúak** (ún. ***FIR*** rendszerek), hiszen a kezdeti hiba **véges lépésben** eltűnik.

3. Ha $F^N = (A - GC)^N \neq 0$, akkor a stabilra tervezett hibarendszer állapotvektorának hossza **exponenciális jelleggel** fog csökkenni. Egy ilyen hibarendszer akkor lesz **stabil**, ha összes **sajátértéke** az egységsugarú körön **belül** helyezkedik el. Az ilyen tulajdonságú állapotátmenet mátrixszal jellemezhető rendszerek **végtelen impulzusválaszúak** (ún. ***IIR*** rendszerek), mert a kezdeti hiba csak **végtelen lépésben** tűnik el.

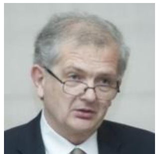
Megfigyelés zajos csatorna esetén: Ebben az esetben nem $\varepsilon(n) \xrightarrow[n\to\infty]{} 0$ az elvárásunk, hanem az, hogy $trace\{E[\varepsilon(n)\varepsilon^T(n)]\} \xrightarrow[n\to\infty]{} min$ legyen.

$$\varepsilon(n) = \begin{bmatrix} \varepsilon_0(n) \\ \varepsilon_1(n) \\ \vdots \\ \varepsilon_{N-1}(n) \end{bmatrix}$$

ezért $trace\{E[\varepsilon(n)\varepsilon^T(n)]\} = \sum_{k=0}^{N-1} \varepsilon_k^2(n)$. $\boxed{\varepsilon(n+1) = F\varepsilon(n)}$ helyett

$$\boxed{E[\varepsilon(n+1)\varepsilon^T(n+1)] = FE[\varepsilon(n)\varepsilon^T(n)]F^T}$$

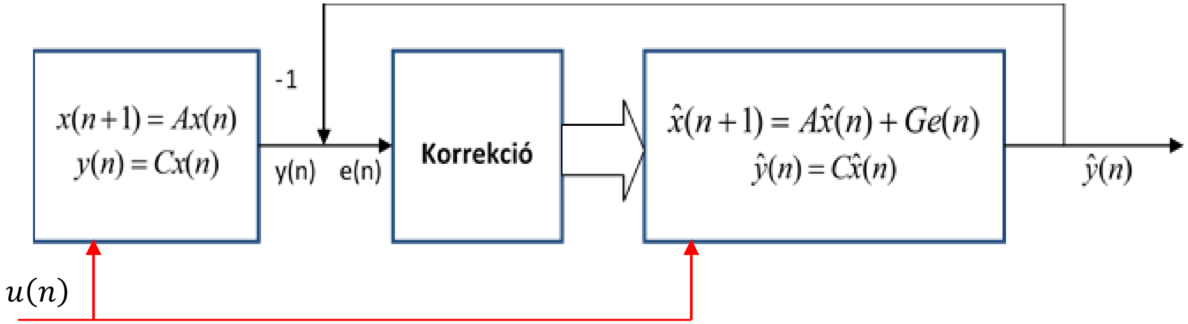
hibarendszer jellemzést használunk.



Emlékeztető

Ez a $\boxed{P=E\left[\varepsilon(n)\varepsilon^{T}(n)\right]}$ hiba-mátrix központi szerepet kap a híres Kalman prediktor, ill. szűrő esetében. **Megjegyzések:**

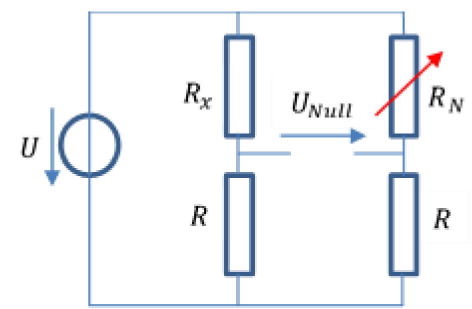
1. A megfigyelő elrendezés mindkét modellje „gerjeszthető” egy közös gerjesztéssel.



Mivel a modellek lineárisak, a **szuperpozíció** értelmében a megfigyelő konvergenciája változatlanul megvalósul.

2. Az ábra szerinti megfigyelőt **Luenberger megfigyelő**nek nevezzük. Luenberger szerint **majdnem minden rendszer** megfigyelő. A megfigyelő tulajdonság feltétele, hogy a megfigyelő legyen „gyorsabb”, mint a megfigyelt rendszer, különben nem képes követni a változásokat.

3. Egy ellenállás- vagy impedancia-mérő híd ismeretlen elemet tartalmazó hídága a valóság **fizikai modellje**, a kiegyenlítő elemet tartalmazó ága pedig a megfigyelőben felépülő, **beállítható/hangolható modell**.



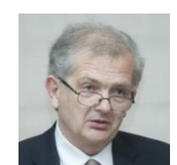
A hídágak osztópontján megjelenő feszültségek különbsége vezérli a hangolást, és a végén a két feszültség megegyezik, a beállítható elemről leolvasott érték segítségével meghatározható az ismeretlen. Ez az áramkör, a hangolást végző operátor részvételével megvalósítja a megfigyelőt.

Példa:

Adott $A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$; $C=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

Hogyan állítsuk be G-t?

$$\boxed{G=AC^{-1}=A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}}$$



Emlékeztető

Példa: Adott $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; C = [1 \quad 1]$. Hogyan állítsuk be G-t? $\boxed{G = \begin{bmatrix} g_0 \\ g_1 \end{bmatrix} = ?}$

$GC = \begin{bmatrix} g_0 \\ g_1 \end{bmatrix} [1 \quad 1] = \begin{bmatrix} g_0 & g_0 \\ g_1 & g_1 \end{bmatrix}$, $[A - GC] = \begin{bmatrix} 1 - g_0 & -g_0 \\ -g_1 & -1 - 1g_1 \end{bmatrix}$, $[A - GC]^2 = 0$ alapján.

$$\begin{bmatrix} 1 - g_0 & -g_0 \\ -g_1 & -1 - g_1 \end{bmatrix} \begin{bmatrix} 1 - g_0 & -g_0 \\ -g_1 & -1 - g_1 \end{bmatrix} = \begin{bmatrix} 1 - 2g_0 + g_0^2 + g_0 g_1 & -g_0 + g_0^2 + g_0 + g_0 g_1 \\ -g_1 + g_1^2 + g_1 + g_0 g_1 & 1 + 2g_1 + g_1^2 + g_0 g_1 \end{bmatrix} =$$

$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ A mellékátló kifejezéseit a főátló kifejezéseibe behelyettesítve kapjuk:

$\boxed{1 - 2g_0 = 0,}$ $\boxed{1 + 2g_1 = 0,}$ amiből: $g_0 = 0.5$ és $g_1 = -0.5$.

Ellenőrzésképpen: $\begin{bmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Példa: Határozzuk meg $[A - GC]$ sajátértékeit erre az esetre: $det[\lambda I - A + GC] = 0$

$$det \begin{bmatrix} \lambda - 0.5 & 0.5 \\ -0.5 & \lambda + 0.5 \end{bmatrix} = (\lambda - 0.5)(\lambda + 0.5) + 0.25 = \lambda^2 - 0.25 + 0.25 = 0.$$

Mindkét sajátérték nulla.

Megjegyzések:

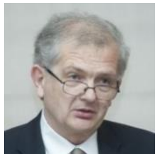
1. Ez a tulajdonság **általánosan igaz** véges lépésben konvergálni képes rendszerek esetében.
2. Az ilyen rendszerek átviteli függvénye olyan (elfajuló) racionális törtfüggvény, amelynek **valamennyi pólusa az origóban** van:

$$H(z) = a_1 z^{-1} + a_2 z^{-2} + \ldots + a_N z^{-N} = \frac{a_N + a_{N-1} z + a_{N-2} z^2 + \ldots + a_1 z^{N-1}}{z^N}$$

Ezek az ún. **véges impulzusválaszú** (*FIR*) szűrők.

Időtartománybeli megfelelője: $y(n) = a_1 x(n-1) + a_2 x(n-2) + \ldots + a_N x(n-N)$,

ahol a valós idejű **kiszámíthatóság** miatt csak $x(n)$ korábbi mintái szerepelhetnek.



Példa: A sajátértékekre vonatkozó feltétel felhasználható a g_0 és a g_1 értékek meghatározására:

$$det[\lambda I - A + GC] = 0,\ \ det\begin{bmatrix}\lambda - 1 + g_0 & g_0 \\ g_1 & \lambda + 1 + g_1\end{bmatrix} = \lambda^2 + \lambda(g_0 + g_1) + g_0 - g_1 - 1 =$$

$= \lambda^2 = 0.$ Ebből: $g_0 + g_1 = 0,$ ill. $g_0 - g_1 = 1,$ amiből: $g_0 = 0.5$ és $g_1 = -0.5.$

Van, hogy a befogadó környezet „dinamikáját” közvetlenül nem ismerjük:
Az állapotegyenlet nem ismert vagy $\boxed{A = I.}$ Ekkor csak megfigyeléseink vannak.

Legkisebb négyzetes hibájú (LS) becslők: Nincs előzetes ismeretünk sem a mérendő paraméterről, sem a csatorna karakterisztikáról (a zajról).

Tegyük fel, hogy a megfigyelési egyenlet lineáris: $y(n) = Cx(n) + w(n);$ $w(n)$ megfigyelési zaj
Feltételezzük, hogy az $x(n)$ ismeretlen $\hat{x}(n)$ értéket vesz fel, és felállítjuk a **megfigyelés modelljét**: $C\hat{x}(n)$. A megfigyelést ezzel összevetve keressük $\hat{x}(n)$ legjobb beállítását négyzetes hibafüggvény feltételezésével: $J\big(x(n), \hat{x}(n)\big) = \big(y(n) - C\hat{x}(n)\big)^T \big(y(n) - C\hat{x}(n)\big) =$

$= y(n)^T y(n) - y(n)^T C\hat{x}(n) - \hat{x}(n)^T C^T y(n) + \hat{x}(n)^T C^T C\hat{x}(n) =$

$= y(n)^T y(n) - 2\hat{x}(n)^T C^T y(n) + \hat{x}(n)^T C^T C\hat{x}(n)$ melynek szélsőértékét (minimumát) keressük:

$$\boxed{\left.\frac{\partial J\big(x(n), \hat{x}(n)\big)}{\partial \hat{x}(n)}\right|_{\hat{x}(n) = \hat{x}_{LS}} = 0} \quad -2C^T y(n) + 2C^T C\hat{x}(n) = 0, \quad \boxed{\hat{x}(n) = [C^T C]^{-1} C^T y(n)}$$

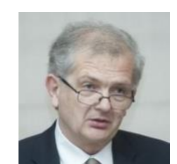
Itt $A = I$ miatt $x(n)$ valójában konstans, tipikusan paraméter.

Többször is megfigyeljük, és a megfigyelt (zajos) értékeket $y(n)$ vektorba gyűjtjük.

Ha $C = [1 \quad 1 \quad \cdots \quad 1]^T$, akkor

$$\boxed{\hat{x}_{LS} = \frac{1}{N}\sum_{k=0}^{N-1} y_k(n)}$$

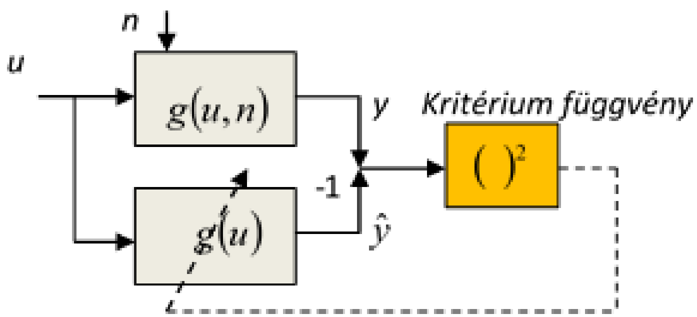
Azaz az egyszerű átlagolásra jutunk.



Modellillesztés

A legkisebb négyzetes hibájú becslők esetén nincs előzetes ismeretünk, valójában modellt illesztettünk.

A modellillesztés problémája meglehetősen szerteágazó.
Egyik klasszikus válfaja a **regresszió számítás**: függő és független változók közötti közvetlen, determinisztikus kapcsolat meghatározása, a modellillesztés egy speciális esete.



Az ábrán látható elrendezésben a modellezendő $y = g(u,n)$ függvény **kétfajta független változóval** rendelkezik: az egyiket $u(n)$ jelöli, **amelyet ismerünk** és „kézben tudunk tartani”, a másik, amelyiket $n(n)$ jelöli, **amely ismeretlen**, kézben nem tartható, tipikusan zajfolyamatnak elképzelt/modellezett folyamat.

A modellezéshez egy általunk kézben tartott, tipikusan paraméterek segítségével módosítható („hangolható”) $\hat{y} = \hat{g}(u)$ függvényt használunk. A cél egy olyan „beállítás” elérése, amely valamilyen értelemben optimális. Tipikusan négyzetes kritériumot használunk:

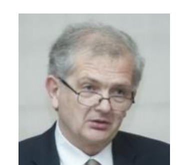
$$\boxed{J = E\{(y - \hat{y})^T (y - \hat{y})\}}$$

Lineáris regresszió: Az illesztendő függvény a $\hat{g}(u) = a_0 + a_1 u$ skalár lineáris függvény, melynek paraméterei úgy választandók meg, hogy $E\left\{\left(y - \hat{g}(u)\right)^2\right\}$ minimális legyen.

Minimalizálandó a $J = E\{(y - a_0 - a_1 u)^2\}$ összefüggés a_0 és a_1 szerint:

$$\frac{\partial J}{\partial a_0} = -2E\{(y - a_0 - a_1 u)\} = -2(E\{y\} - a_0 - a_1 E\{u\}) = 0$$

$$\frac{\partial J}{\partial a_1} = -2E\{u(y - a_0 - a_1 u)\} = -2(E\{uy\} - a_0 E\{u\} - a_1 E\{u^2\}) = 0$$



Polinomiális regresszió: Fontos tulajdonsága, hogy paramétereiben lineáris.

$$\hat{g}(u) = \sum_{k=0}^{N} a_k \, u^k$$

A paramétereiben lineáris modellek előnyösek, mert négyzetes hibakritérium esetén a szélsőérték-keresés lineáris egyenletrendszer megoldására vezet: négyzetes kifejezések paraméterek szerinti deriválása lineáris összefüggést eredményez.

Lineáris regresszió mérési adatok alapján:

A fentieket végigvihetjük akkor is, ha nincsen előzetes információnk.

$\boxed{\text{Ilyenkor } y_n = a_0 + a_1 u_n + w_n,}$ ami az n-edik időpillanatban végzett megfigyelés modellje.

Végzünk N megfigyelést. A „bemenet” és a megfigyelt értékeket vektorba rendezzük.

$$\boldsymbol{z} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix}}_{y(n)} = \underbrace{\begin{bmatrix} 1 & u_0 \\ 1 & u_1 \\ \vdots & \vdots \\ 1 & u_{N-1} \end{bmatrix}}_{C} * \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{x(n)} + \underbrace{\begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix}}_{w(n)}$$

Felismerjük, hogy ez a struktúra ugyanaz, mint a **legkisebb négyzetes hibájú becslők**nél megismert!

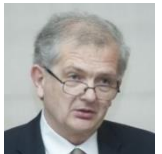
$$\boxed{\hat{x}(n) = [C^T C]^{-1} C^T y(n)}$$

$$[\boldsymbol{C}^T \boldsymbol{C}] = \begin{bmatrix} N & \sum_{n=0}^{N-1} u_n \\ \sum_{n=0}^{N-1} u_n & \sum_{n=0}^{N-1} u_n^2 \end{bmatrix},$$

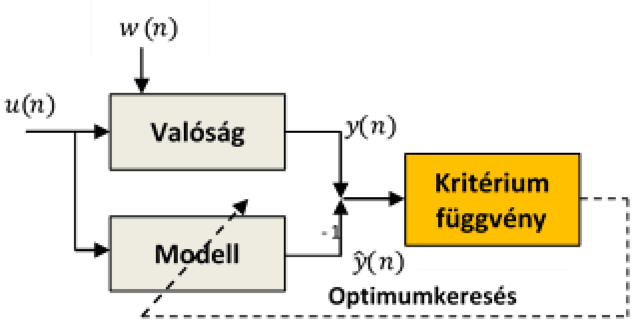
$$\boldsymbol{C}^T \boldsymbol{z} = \begin{bmatrix} \sum_{n=0}^{N-1} y_n \\ \sum_{n=0}^{N-1} u_n y_n \end{bmatrix},$$

$$\boxed{\begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} = \frac{1}{\frac{1}{N}\sum_{n=0}^{N-1} u_n^2 - \left(\frac{1}{N}\sum_{n=0}^{N-1} u_n\right)^2} \begin{bmatrix} \frac{1}{N}\sum_{n=0}^{N-1} u_n^2 & -\frac{1}{N}\sum_{n=0}^{N-1} u_n \\ -\frac{1}{N}\sum_{n=0}^{N-1} u_n & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{N}\sum_{n=0}^{N-1} y_n \\ \frac{1}{N}\sum_{n=0}^{N-1} u_n y_n \end{bmatrix},}$$

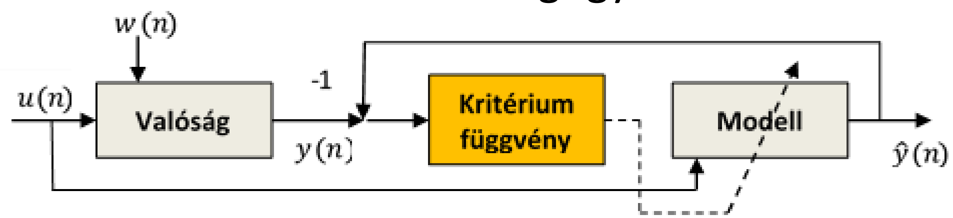
Felismerjük a statisztikai jellemzők közelítő összefüggéseit!



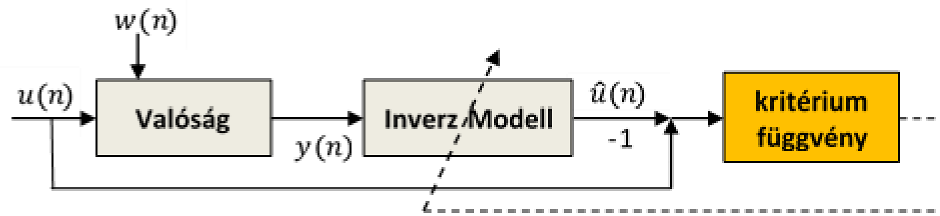
A regressziós séma általánosítása



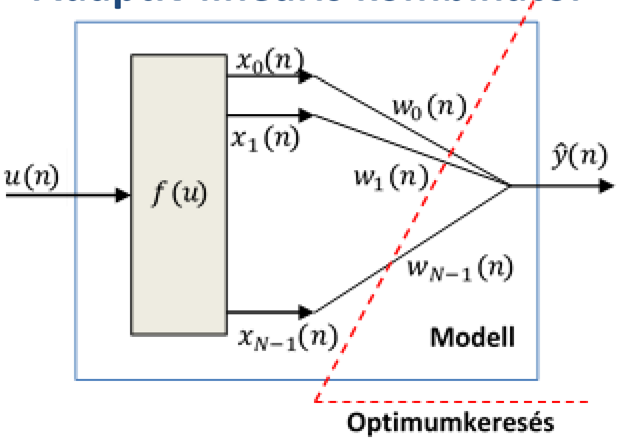
A modell-illesztési feladat megfigyelőként ábrázolva



Inverz modell illesztése



Adaptív lineáris kombinátor



Az általánosított regressziós sémához kapcsolódó egyik gyakran használt modell-család. Ebben az $u(n)$ diszkrét értéksorozatból egy $\boldsymbol{X}^T(n) = [x_o(n) \quad x_1(n) \quad \ldots \quad x_{N-1}(n)]$ értéksorozatot állítunk elő először, majd ezen értékek lineáris kombinációjaként állítjuk elő az $\hat{y}(n)$ értéket.

A $\mathbf{W}^T(n) = [w_0(n) \quad w_1(n) \quad \ldots \quad w_{N-1}(n)]$ paraméterek legkedvezőbb, minimális négyzetes hibát eredményező beállítására törekszünk.

$$J(\boldsymbol{W}(n)) = E\{[y(n) - \boldsymbol{X}^T(n)\boldsymbol{W}(n)]^T[y(n) - \boldsymbol{X}^T(n)\boldsymbol{W}(n)]\}=$$

$$= E\{y^T(n)y(n)\} - 2\boldsymbol{W}^T(n)E\{\boldsymbol{X}(n)y(n)\} + \boldsymbol{W}^T(n)E\{\boldsymbol{X}(n)\boldsymbol{X}^T(n)\}\boldsymbol{W}(n).$$

$$\boxed{E\{\boldsymbol{X}(n)y(n)\} = \boldsymbol{P}}$$

$$\boxed{E\{\boldsymbol{X}(n)\boldsymbol{X}^T(n)\} = \boldsymbol{R}}$$

$$\boxed{\frac{\partial J(\boldsymbol{W}(n))}{\partial \boldsymbol{W}(n)} = -2\boldsymbol{P} + 2\boldsymbol{R}\boldsymbol{W}(n) = 0}$$

$$\boxed{\boldsymbol{W}^* = \boldsymbol{R}^{-1}\boldsymbol{P}}$$

Wiener-Hopf egyenlet.