Beágyazott információs rendszerek

5. Mennyiségek, változók valós idejű rendszerekben

2020. október 28.

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:

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

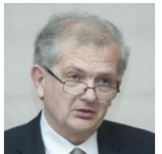
Itt $A = I$ miatt $x(n)$ valójában konstans, tipikusan paraméter.

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.

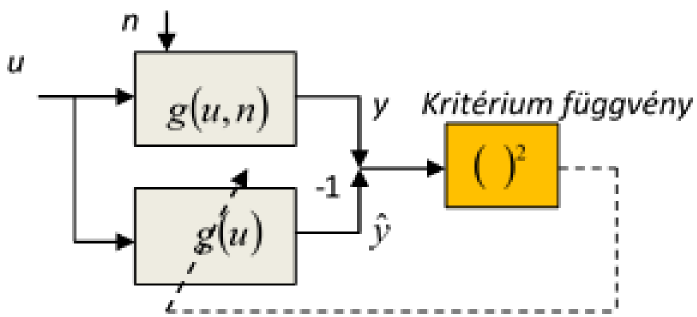


Modellillesztés

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

Emlékeztető

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:

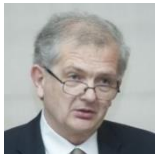
$$\boxed{J = E\{(y-\hat{y})^T(y-\hat{y})\}}$$

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:

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

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



Emlékeztető

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.

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!

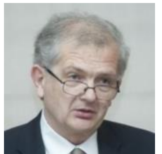
$$[\boldsymbol{C}^T \boldsymbol{C}] = \begin{bmatrix} N & \sum_{n=0}^{N-1} u_n \\ \sum_{n=0}^{N-1} u_n & \sum_{n=0}^{N-1} u_n^2 \end{bmatrix},$$

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

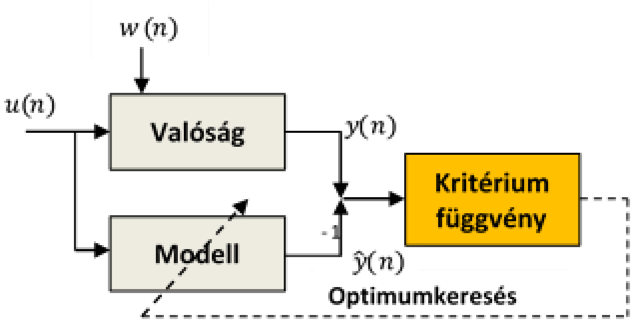
$$\boxed{\hat{x}(n) = [C^T C]^{-1} C^T y(n)}$$

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

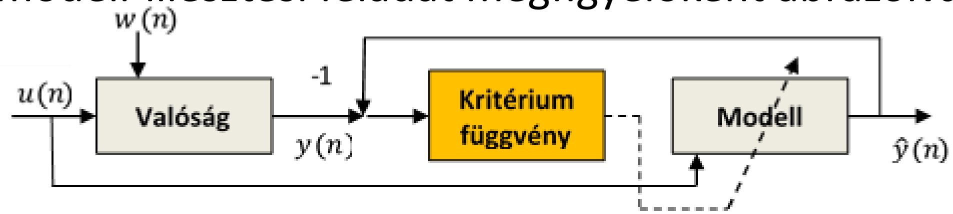
Felismerjük a statisztikai jellemzők közelítő összefüggéseit!



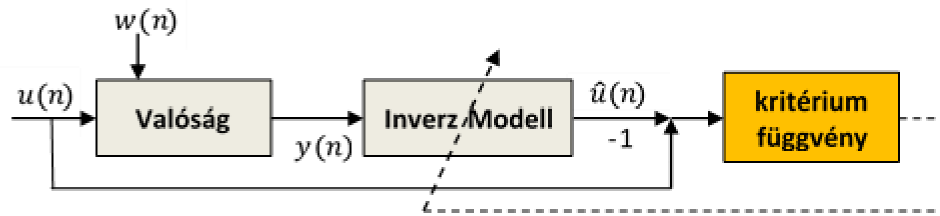
A regressziós séma általánosítása



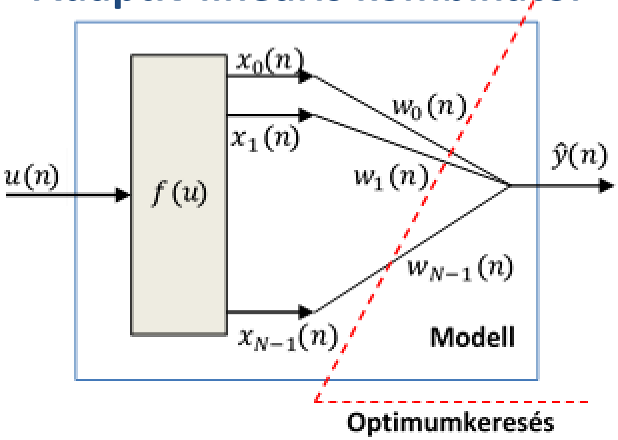
A modell-illesztési feladat megfigyelőként ábrázolva



Inverz modell illesztése



Adaptív lineáris kombinátor



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

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

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

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

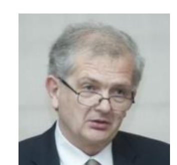
$$\boxed{E\{\boldsymbol{X}(n)y(n)\} = \boldsymbol{P}}$$

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

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

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

Wiener-Hopf egyenlet.

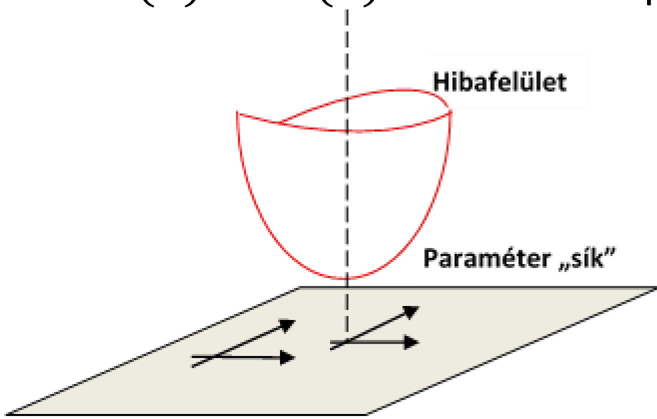


Az optimális beállítás minimális hibát eredményez, amit visszahelyettesítéssel kapunk:

$J_{min} = E\{y^T(n)y(n)\} - \boldsymbol{P}^T\boldsymbol{R}^{-1}\boldsymbol{P} = E\{y^T(n)y(n)\} - \boldsymbol{P}^T\boldsymbol{W}^*$ Ennek felhasználásával:

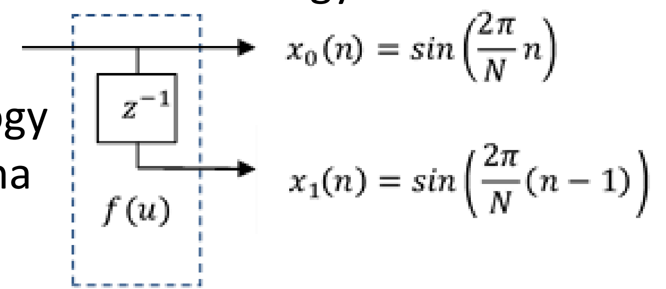
$$\boxed{J\big(\boldsymbol{W}(n)\big) = J_{min} + (\boldsymbol{W}(n) - \boldsymbol{W}^*)^T\boldsymbol{R}(\boldsymbol{W}(n) - \boldsymbol{W}^*) = J_{min} + \boldsymbol{V}^T(n)\boldsymbol{R}\boldsymbol{V}(n)}$$

ahol $\boldsymbol{V}(n) = \boldsymbol{W}(n) - \boldsymbol{W}^*$ az ún. paraméterhiba vektora. **Példa:**



Legyen $\boldsymbol{X}^T(n) = [sin(2\pi n/N) \quad sin(2\pi(n-1)/N)]$, azaz egy szinuszos hullámforma két egymás utáni mintája.

$u(n) = x_0(n) = sin\left(\frac{2\pi}{N}n\right)$



Itt most N azt jelöli, hogy a szinuszos hullámforma egy periódusa hány mintából áll.

A jel, amihez az illesztést végezzük: $y(n) = 2\,cos(2\pi n/N)$. $\quad \boldsymbol{W}^T(n) = [w_0(n) \quad w_1(n)]$

Az $\boldsymbol{R}$ és a $\boldsymbol{P}$ mátrixok a szinuszos, ill. koszinuszos hullámformák teljes ($N > 2$) periódusra történő átlagolásával származtathatók:

$E\{x_0^2(n)\} = E\{x_1^2(n)\} = E\{(sin^2(2\pi/N)\} = E\{sin^2(2\pi(n-1)/N)\} = 0.5,$

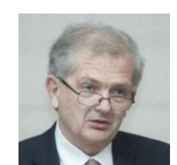
$E\{x_0(n)x_1(n)\} = E\{sin(2\pi n/N)\,sin(2\pi(n-1)/N)\} = 0.5\,cos(2\pi/N)$

$E\{x_0(n)y(n)\} = E\{2\,sin(2\pi n/N)\,cos(2\pi n/N)\} = 0,$

$E\{x_1(n)y(n)\} = E\{2\,sin(2\pi(n-1))/N\,cos(2\pi n/N)\} = -sin(2\pi/N).$

$$\boxed{\boldsymbol{R} = \begin{bmatrix} 0.5 & 0.5\,cos\left(\frac{2\pi}{N}\right) \\ 0.5\,cos\left(\frac{2\pi}{N}\right) & 0.5 \end{bmatrix}}$$

$$\boxed{\boldsymbol{P} = \begin{bmatrix} 0 \\ -sin\left(\frac{2\pi}{N}\right) \end{bmatrix}.}$$



Képezve $\boldsymbol{R}$ inverzét:

$$\boldsymbol{R}^{-1} = \frac{1}{0.25\, sin^2\left(\frac{2\pi}{N}\right)} \begin{bmatrix} 0.5 & -0.5\, cos\left(\frac{2\pi}{N}\right) \\ -0.5\, cos\left(\frac{2\pi}{N}\right) & 0.5 \end{bmatrix}$$

$$\boxed{\boldsymbol{R} = \begin{bmatrix} 0.5 & 0.5\, cos\left(\frac{2\pi}{N}\right) \\ 0.5\, cos\left(\frac{2\pi}{N}\right) & 0.5 \end{bmatrix}}$$

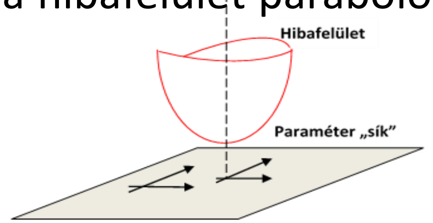
$$\boxed{\boldsymbol{W}^* = \boldsymbol{R}^{-1}\boldsymbol{P} = \begin{bmatrix} \dfrac{2}{tan(2\,\pi/N)} \\ -\dfrac{2}{sin(2\,\pi/N)} \end{bmatrix}}$$

Ezzel a paraméter-beállítással az illesztett modell kimenete:

$$\boxed{\boldsymbol{P} = \begin{bmatrix} 0 \\ -\, sin\left(\frac{2\pi}{N}\right) \end{bmatrix}}.$$

$$\hat{y}(n) = X^T(n)W^* = 2\frac{sin(2\pi\, n/N)}{tan(2\,\pi/N)} - 2\frac{sin(2\pi(n-1)/N)}{sin(2\,\pi/N)} = 2cos(2\pi n/N)$$

Mivel szinuszos minták lineáris kombinációjával hiba nélkül elő lehet állítani koszinuszos hullámformák mintáit, ezért a példa szerinti esetben J_{min}, azaz a hibafelület paraboloid legalsó pontja érinti a paraméterek síkját.



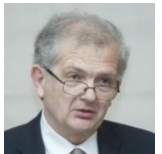
Út az adaptív eljárásokhoz

A hibafelület tetszőleges pontjában a hiba csökkenés fajlagos mértékét a felület meredekségével (gradiensével) mérhetjük:

$$\boxed{\boldsymbol{\nabla}(n) = \frac{\partial J(\boldsymbol{W}(n))}{\partial \boldsymbol{W}(n)} = 2\boldsymbol{R}[\boldsymbol{W}(n) - \boldsymbol{W}^*]}$$

Ez utóbbi mindkét oldalát megszorozva balról az $\frac{1}{2}\boldsymbol{R}^{-1}$ mátrixszal:

$\boldsymbol{W}^* = \boldsymbol{W}(n) - \frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{\nabla}(n),$ ahol felhasználtuk: $\boxed{\boldsymbol{W}^* = \boldsymbol{R}^{-1}\boldsymbol{P}}$



Feltételezve, hogy nincs tökéletes ismeretünk az $\boldsymbol{R}$ mátrixról, és ebből adódóan a gradiensről, ez átírható egy iteratív formára:

$$\boldsymbol{W}^* = \boldsymbol{W}(n) - \frac{1}{2}\boldsymbol{R}^{-1}\boldsymbol{\nabla}(n) \quad \Longrightarrow \quad \boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \frac{1}{2}\widehat{\boldsymbol{R}}^{-1}\widehat{\boldsymbol{\nabla}}(n),$$

illetve a konvergencia sebességet meghatározó $\boldsymbol{0 < \mu < 1}$ „bátorsági” tényező bevezetésével, visszaírva a „tökéletes” $\boldsymbol{R}$ mátrixot és gradienst:

$$\boxed{\boldsymbol{W}(n+1) = \boldsymbol{W}(n) - \mu\boldsymbol{R}^{-1}\boldsymbol{\nabla}(n)} \quad (*)$$

Megjegyzések:

1. Ha pontosan ismerjük az $\boldsymbol{R}$ mátrixot és gradienst, akkor $\mu = \frac{1}{2}$ egylépéses konvergenciát biztosít tetszőleges $\boldsymbol{W}(n)$ kezdőpontból.
2. Mivel $\boldsymbol{\nabla}(n) = 2\boldsymbol{R}[\boldsymbol{W}(n) - \boldsymbol{W}^*]$, ezért ezt a (*) összefüggésbe behelyettesítve, és az egyenlet mindkét oldalából levonva W^* értékét:

$$\boxed{\boldsymbol{W}(n+1) - \boldsymbol{W}^* = (1-2\mu)(\boldsymbol{W}(n) - \boldsymbol{W}^*) = \boldsymbol{V}(n+1) = (1-2\mu)^{n+1}\boldsymbol{V}(0),} \quad (**)$$

   vagyis a kezdeti hiba exponenciális jelleggel csökken, ha $\mu \neq \frac{1}{2}$.
   Ha $0 < \mu < 0.5$, akkor monoton csökkenő hibával, ellenkező esetben pedig monoton csökkenő amplitúdójú, de lengő jellegű hibával közelítjük meg.
3. A modell-illesztés gradiens módszereit a szerint különböztetjük meg, hogy a (*) szerinti összefüggés (közelítő) alkalmazásához milyen előzetes ismeretek állnak rendelkezésünkre. Amennyiben az $\boldsymbol{R}$ és a $\boldsymbol{P}$ mátrixok pontosan ismertek, akkor az adaptív lineáris kombinátor működését a (*) és (**) egyenletek írják le.
4. Figyeljük meg, hogy az $\boldsymbol{R}$ mátrix „globális” információt hordoz a hibafelületről, a $\boldsymbol{\nabla}(n)$ gradiens pedig a hibafelület „lokális” jellemzése. Ezen lokális ismeret alapján „ereszkedünk” a hibafelületen, hogy minél közelebb kerüljünk az optimumhoz.

A replikátum determinizmusa

Ha a megbízhatóságot aktív redundanciával, azaz fizikai többszörözéssel javítjuk, akkor a párhuzamosan működő egységeknél meg kell követeljük, hogy

(1) a kívülről látható RAM állapotuk **ugyanaz legyen**, és

(2) a **kimenetek azonosak** legyenek, maximum *d* időbeni eltéréssel.

A *d* értékét a rendszer dinamikai tulajdonságai alapján határozhatjuk meg: kell, hogy maradjon idő a hibás vagy hiányzó adat pótlására a replikátumból.

Példa: Három csatornás repülés-irányító rendszer többségi szavazással. Mindegyik csatorna önálló szenzorokkal és számítógépekkel rendelkezik, hogy az ún. közös-módusú hibák valószínűségét minimalizáljuk.

A “felszállás kezdete” eseményt követően egy előírt időn belül a vezérlő rendszernek ellenőriznie kell, hogy a repülőgép **elérte-e a felszálláshoz szükséges sebességet**.

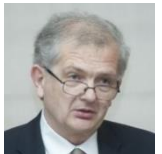
Ha igen, akkor kezdeményezi az emelkedést és **a motorokat tovább gyorsítja**.

Ha nem, akkor a felszállási folyamat megszakítandó, és **a motorokat le kell állítani**.

Az alábbi táblázat egy olyan helyzetet ír le, ahol a replikátum determinizmusa feltétel nem teljesül, és a hibás csatorna érvényesül a döntésben:

Csatorna	Döntés	Akció
1. csatorna	Felszállás	Motor gyorsítása
2. csatorna	Megszakítás	Motor leállítása
3. csatorna	Megszakítás	Motor gyorsítása

A táblázat szerinti első két csatorna helyesen működik, csak nem teljesül a replikátum determinizmusa feltétel. Véletlen hatások eredményeképpen (eltérés a szenzor kalibrációban, digitalizálási hiba, a sebességmérés időpontjában kicsi eltérés) a két csatorna eltérő következtetésre jut. A harmadik csatorna hibásan működik, mert megszakítást dönt, és gyorsítja a motort.



Az akcióra vonatkozó **többségi szavazás a hibás csatorna által javasolt eredményt hoz** a replikátum determinizmusára vonatkozó feltétel teljesülésének hiányában!

Mintavételezés és lekérdezés:

Mintavételezés (sampling) szóhasználattal élünk, ha az adatot a szenzor egységnél írjuk memóriába:

A memória a számítógép befolyásolhatósági tartományán kívül helyezkedik el. A számítógép leállása, újraindulása esetén a memóriatartalom nem vész el.

Lekérdezés (polling) szóhasználattal élünk, ha az adatot a számítógép memóriájába helyezzük:

Funkcionális szempontból a két megoldás nem tér el egymástól, de hiba esetén **a mintavételezés robusztusabb.**

Megjegyzés: Az **interrupt** mechanizmus a **polling-**ot bemutató ábrával jellemezhető.

Súlyos problémája, hogy külső eszköz befolyásolhatósági tartományába helyezi a számítógépet, ezért fokozott körültekintéssel kell alkalmazni, mert hiba esetén oly mértékben túlterhelheti a processzort, hogy az képtelen lesz feladatait (időre) ellátni.