Méréselmélet

6. fejezet: LS becslők rekurzív számítása

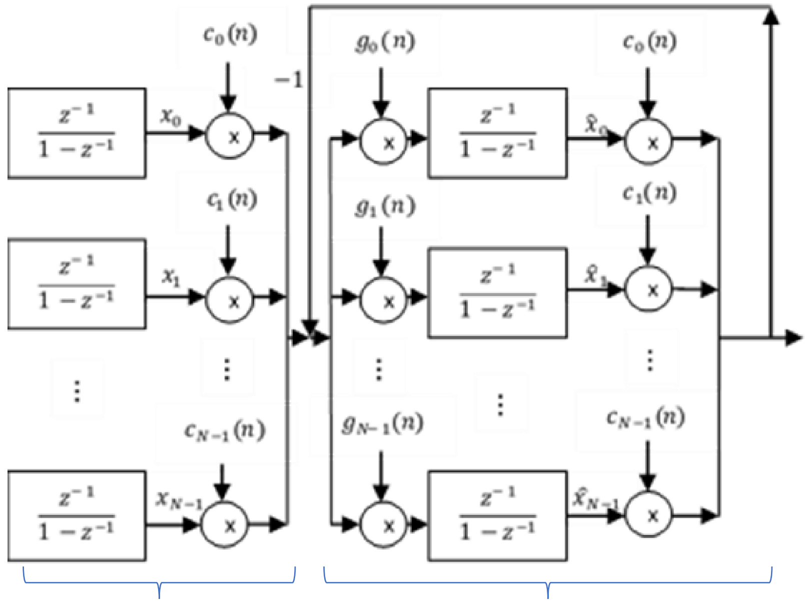
2022. április 6.

Némi útmutatás a 2. házi feladathoz:

1.1. Állítson elő $u(n), n = 0, 1, \ldots$ diszkrét értéksorozatot **multiszinusz generátor** „segítségével”, mégpedig úgy, hogy a diszkrét jel **$M = 100$ harmonikus komponensből álljon**, az egyes harmonikus komponensek amplitúdója egységnyi, kezdőfázisa véletlen, és a sorozat várható értéke pedig 1 legyen!

$$\boxed{u(n) = 1 + \sum_{m=1}^{M} \cos\left(\frac{2\pi}{N}mn + \varphi_m\right)} = 1 + \frac{1}{2}\sum_{m=1}^{M}\left(e^{j\left(\frac{2\pi}{N}mn+\varphi_m\right)} + e^{-j\left(\frac{2\pi}{N}mn+\varphi_m\right)}\right) \qquad \text{ahol } N = 2M + 1$$

A multiszinusz generátort a jegyzet **46. ábrának** megfelelően készítse el! Ügyeljen arra, hogy a generált jel a vizsgált frekvenciatartomány egészében megfelelő gerjesztést adjon (max. 3 pont)! Vizsgálja meg, hogy hogyan alakul a generált jel csúcsértéke, és hasonlítsa össze azzal az esettel, amikor a kezdőfázisok rendre nullák (max. 2 pont)!



$$e^{j\left(\frac{2\pi}{N}mn+\varphi_m\right)} + e^{-j\left(\frac{2\pi}{N}mn+\varphi_m\right)} = e^{j\frac{2\pi}{N}mn}e^{j\varphi_m} + e^{-j\frac{2\pi}{N}mn}e^{-j\varphi_m} =$$

$$\boxed{= c_m(n)x_m + c_{N-m}(n)x_{N-m}}$$

$$x_0 = 1, \quad x_m = x_{N-m}^* = \frac{1}{2}e^{j\varphi_m}, \qquad m = 1, \ldots, M$$

$$c_0 = 1, \; c_m(n) = c_{N-m}^*(n) = e^{j\frac{2\pi}{N}mn}, m = 1, \ldots, M$$

$$\boxed{c_m(n) = e^{j\frac{2\pi}{N}mn}, g_m(n) = \frac{1}{N}e^{-j\frac{2\pi}{N}mn}, m = 0,1, \ldots, N-1,}$$

1.2. Készítsen az előző pont szerinti generátor jelével gerjesztett rendszerek kimeneti jelének analizálására alkalmas rekurzív multiszinusz analizátort, ugyancsak a 46. ábrán látható elrendezést követve, amely a bemenetére kapcsolt jelből képes kiszámítani az abban levő harmonikus komponensek amplitúdóját és fázisát! Az 1.1. pont szerinti jel és az analizátor által rekonstruált jel különbségét ábrázolva mutassa be, hogy a generátor-analizátor pár megfelelően működik (max. 5 pont)!

Ismétlés

5. Szűréselmélet alapjai: Lényeges üzenetek

(Jegyzet 55-67. oldalak)

5.1. Optimális nemrekurzív becslő (skalár Wiener Szűrő)

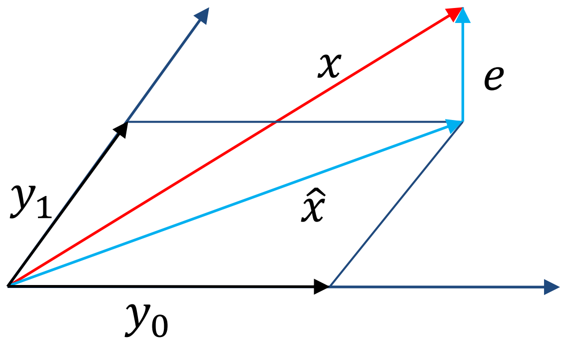
Keressük x legjobb becslőjét

$$\hat{x} = \sum_{k=0}^{N-1} a_k y_k$$

$$\frac{\partial E\{e^2\}}{\partial a_j} = 0, E\{e y_j\} = 0, \forall j - \text{re.}$$

Ez mátrix formában

A legjobb becslés geometriai interpretációja (29. ábra):



Legjobbnak azt a becslést tekintjük, amelyik legkisebb hibával jár: ez az x vektor altérre történő merőleges vetítésével áll elő.

Ez utóbbit ortogonalitási egyenletnek nevezzük, mert a vektorokkal történő geometriai interpretációban éppen azt fejezi ki, hogy az optimális beállítás esetén az e hibavektor merőleges valamennyi y_j vektorra.

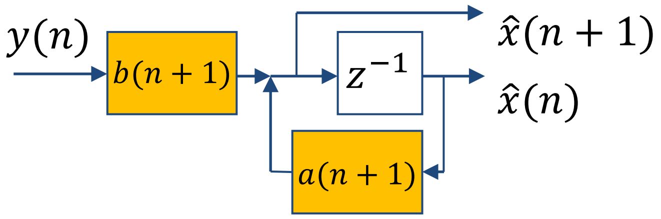
$$\boldsymbol{W} = \boldsymbol{R}_{yy}^{-1}\boldsymbol{P}_{xy} \quad (\text{ahol } \boldsymbol{W} = [a_0 \quad a_1 \quad \dots \quad a_{N-1}]^T) \qquad E\{e^2\}\Big|_{min} = E\{x^2\} - \boldsymbol{P}_{xy}^T\boldsymbol{R}_{yy}^{-1}\boldsymbol{P}_{xy} = E\{x^2\} - \boldsymbol{P}_{xy}^T\boldsymbol{W}$$

Rekurzív becslő az optimális nemrekurzív becslőből

A rekurzív eljárás nagy előnye, hogy nem kell megvárni az összes adatot: folyamatosan számolható az egyre jobb minőségű becslő.

$$\hat{x}(n+1) = a(n+1)\hat{x}(n) + b(n+1)y(n) =$$

$$= \hat{x}(n) + \underbrace{b(n+1)\big(y(n) - \hat{x}(n)\big)}_{\text{korrekciós tag}} \quad ?$$



Lineáris átlagoló: $\hat{x}(n+1) = \dfrac{n}{n+1}\hat{x}(n) + \dfrac{1}{n+1}y(n) = \hat{x}(n) + \dfrac{1}{n+1}\big(y(n) - \hat{x}(n)\big)$

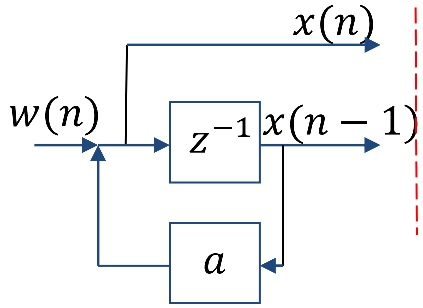
5.2. Optimális rekurzív becslő (skalár Kalman szűrő és prediktor)

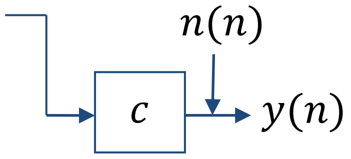
A modell, amit alkalmazunk a legegyszerűbb állapotváltozós modell, amit egy Gauss eloszlású fehér-zaj folyamat gerjeszt.

$$\boxed{x(n) = ax(n-1) + w(n)}$$

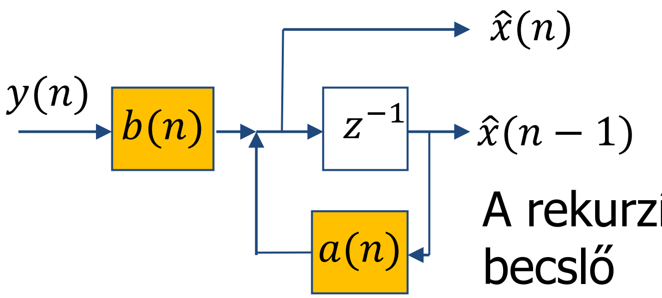
$$R_{xx}(j) = a^{|j|}R_{xx}(0)$$

$$E\{w(n)\} = 0, E\{w(n)w(j)\} = \begin{cases} 0 & n \neq j \\ \sigma_w^2 & n = j \end{cases}$$





A megfigyelés lineáris modellje



A rekurzív becslő

elsőrendű autoregresszív folyamat

$$\boxed{\hat{x}(n) = a(n)\hat{x}(n-1) + b(n)y(n)}$$

Keressük az optimális súlytényezőket: $\frac{\partial E\{e^2(n)\}}{\partial a(n)} = -2E\{e(n)\hat{x}(n-1)\} = 0,$ $\frac{\partial E\{e^2(n)\}}{\partial b(n)} - 2E\{e(n)y(n)\} = 0$

„ortogonalitás”! „ortogonalitás”!

Nagyon fontos eredmény:

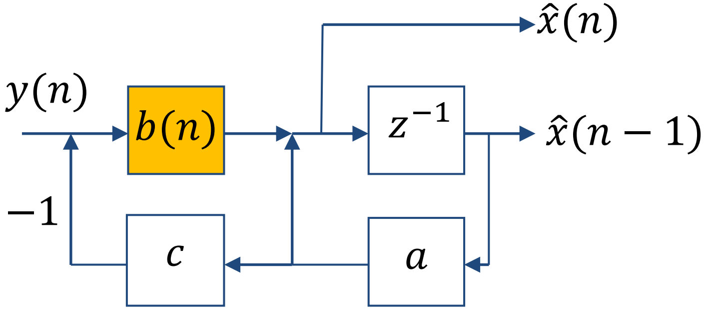
$$\boxed{a(n) = a[(1 - b(n)c)]} \quad \boxed{\hat{x}(n) = a\hat{x}(n-1) + b(n)(y(n) - ca\hat{x}(n-1))}$$

$b(n)$ meghatározása:

$$E\{e^2(n)\} = E\left\{\left[ae(n-1) + w(n) - b(n)\big(ace(n-1) + cw(n) + n(n)\big)\right]^2\right\} = p(n)$$

$$E\{e^2(n)\} = E\{[a[1 - b(n)c]e(n-1) + [1 - b(n)c]w(n) - b(n)n(n)]^2\}$$

$$E\{e(n-1)w(n)\} = 0 \quad E\{e(n-1)n(n)\} = 0 \quad E\{w(n)n(n)\} = 0$$



$$\boxed{b(n)\Big|_{opt} = \frac{a^2cp(n-1) + c\sigma_w^2}{a^2c^2p(n-1) + c^2\sigma_w^2 + \sigma_n^2}}$$

$$p(n) = a^2[1 - b(n)c]^2p(n-1) + [1 - b(n)c]^2\sigma_w^2 + b^2(n)\sigma_n^2$$

(1) $p_1(n) = a^2p(n-1) + \sigma_w^2$ (2) $b(n) = cp_1(n)[c^2p_1(n) + \sigma_n^2]^{-1}$

(3) $p(n) = [1 - b(n)c]p_1(n)$ $p(0)$ kell hozzá!

Optimális rekurzív becslő (skalár Kalman szűrő és prediktor):

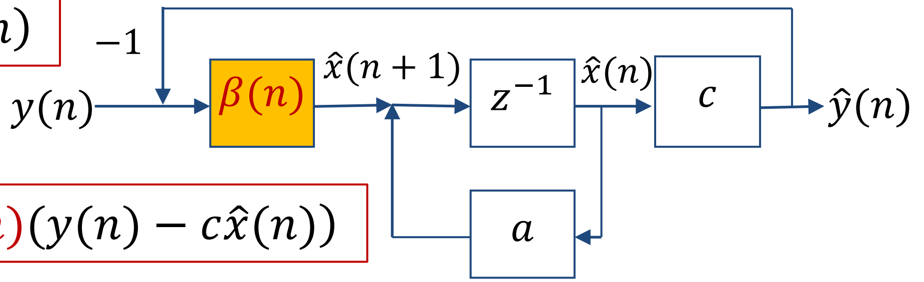
$\boxed{x(n) = ax(n-1) + w(n)}$

$y(n) = cx(n) + n(n)$

helyett

$\boxed{x(n+1) = ax(n) + w(n)}$

$y(n) = cx(n) + n(n)$



$\boxed{\hat{x}(n) = a\hat{x}(n-1) + b(n)(y(n) - ca\hat{x}(n-1))}$

(1) $p_1(n) = a^2 p(n-1) + \sigma_w^2$

(2) $b(n) = cp_1(n)[c^2 p_1(n) + \sigma_n^2]^{-1}$ $\quad n = 1, 2, \ldots$

(3) $p(n) = [1 - b(n)c]p_1(n)$ $\quad p(0)$ kell hozzá!

$\boxed{\hat{x}(n+1) = a\hat{x}(n) + \beta(n)(y(n) - c\hat{x}(n))}$

(1) $\beta(n) = acp(n)[c^2 p(n) + \sigma_n^2]^{-1}$,

(2) $p(n+1) = a(a - \beta(n)c)p(n) + \sigma_w^2$ $\quad n = 0, 1, 2, \ldots$

$p(0)$ kell hozzá!

Optimális rekurzív becslő (vektor Kalman szűrő és prediktor):

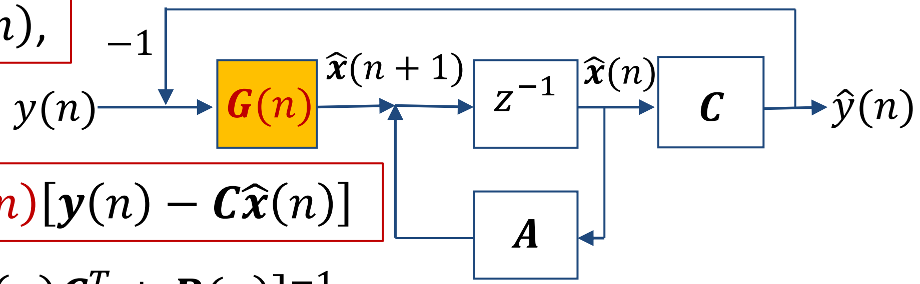
$\boxed{\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{x}(n-1) + \boldsymbol{w}(n)}$

$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{n}(n)$

helyett

$\boxed{\boldsymbol{x}(n+1) = \boldsymbol{A}\boldsymbol{x}(n) + \boldsymbol{w}(n),}$

$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{n}(n)$



$\boxed{\hat{\boldsymbol{x}}(n) = \boldsymbol{A}\hat{\boldsymbol{x}}(n-1) + \boldsymbol{K}(n)\left(\boldsymbol{y}(n) - \boldsymbol{C}\boldsymbol{A}\hat{\boldsymbol{x}}(n-1)\right)}$

(1) $\boldsymbol{P}_1(n) = [\boldsymbol{A}\boldsymbol{P}(n-1)\boldsymbol{A}^T + \boldsymbol{Q}(n)]$

(2) $\boldsymbol{K}(n) = \boldsymbol{P}_1(n)\boldsymbol{C}^T[\boldsymbol{C}\boldsymbol{P}_1(n)\boldsymbol{C}^T + \boldsymbol{R}(n)]^{-1}$

(3) $\boldsymbol{P}(n) = [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}]\boldsymbol{P}_1(n)$ $\quad n = 1, 2, \ldots$ $\quad \boldsymbol{P}(0)$ kell hozzá!

$\boldsymbol{P}(n) = E\{[\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)][\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)]^T\}$

$\boxed{tr\boldsymbol{P}(n) \rightarrow \text{minimum!}}$

$\boxed{\hat{\boldsymbol{x}}(n+1) = \boldsymbol{A}\hat{\boldsymbol{x}}(n) + \boldsymbol{G}(n)[\boldsymbol{y}(n) - \boldsymbol{C}\hat{\boldsymbol{x}}(n)]}$

(1) $\boldsymbol{G}(n) = \boldsymbol{A}\boldsymbol{P}(n)\boldsymbol{C}^T[\boldsymbol{C}\boldsymbol{P}(n)\boldsymbol{C}^T + \boldsymbol{R}(n)]^{-1}$

(2) $\boldsymbol{P}(n+1) = [\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{P}(n)\boldsymbol{A}^T + \boldsymbol{Q}(n)$ $\quad n = 0, 1, 2, \ldots$

$\boldsymbol{P}(0)$ kell hozzá!

$\boldsymbol{P}(n+1) = E\{[\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)][\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)]^T\}$

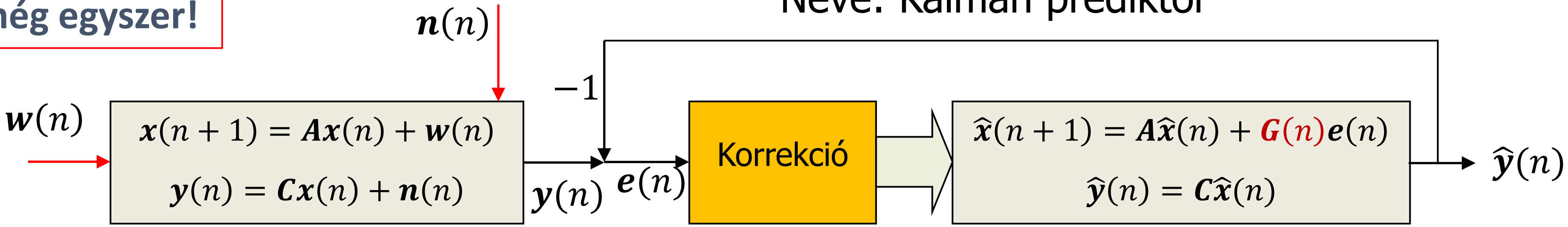
$\boxed{tr\boldsymbol{P}(n+1) \rightarrow \text{minimum!}}$

A Kalman prediktor levezetés még egyszer!

Neve: Kalman prediktor

$$\boldsymbol{x}(n+1) = \boldsymbol{A}\boldsymbol{x}(n) + \boldsymbol{w}(n)$$
$$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{n}(n),$$

$\boldsymbol{w}(n)$ az ún. rendszer-zaj,
$\boldsymbol{n}(n)$ az ún. megfigyelési-zaj.
Mind a rendszer-, mind a megfigyelési-zaj vektor egymástól és a rendszer állapotától független nulla várható értékű fehér Gauss folyamat.



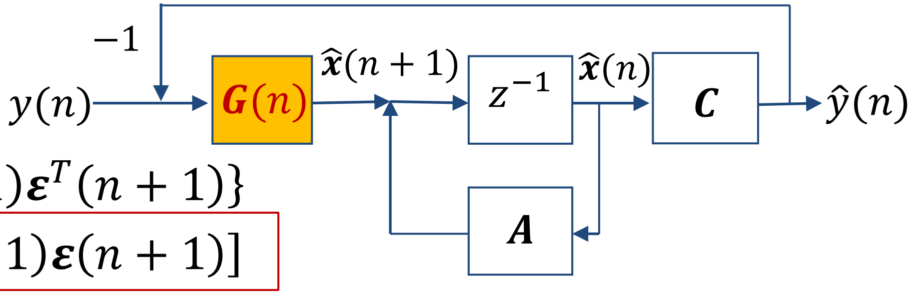
$$\boldsymbol{e}(n) = \boldsymbol{y}(n) - \hat{\boldsymbol{y}}(n)$$

A mérési eljárás, amit a kiegészített modellhez rendelünk egy megfigyelő:

$$\hat{\boldsymbol{x}}(n+1) = \boldsymbol{A}\hat{\boldsymbol{x}}(n) + \boldsymbol{G}(n)\boldsymbol{e}(n) = \boldsymbol{A}\hat{\boldsymbol{x}}(n) + \boldsymbol{G}(n)\big(\boldsymbol{y}(n) - \boldsymbol{C}\hat{\boldsymbol{x}}(n)\big)$$

$\boldsymbol{Q}(n) = E\{\boldsymbol{w}(n)\boldsymbol{w}^T(n)\}$
$\boldsymbol{R}(n) = E\{\boldsymbol{n}(n)\boldsymbol{n}^T(n)\}$

Keressük azt a $\boldsymbol{G}(n)$ mátrixot (az ún. prediktor erősítést), amely mellett a becslési hiba



$$\boldsymbol{P}(n+1) = E\{[\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)][\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)]^T\} = E\{\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^T(n+1)\}$$

kovariancia mátrixának nyoma minimális.

$$\boxed{tr\boldsymbol{P}(n+1) = E[\boldsymbol{\varepsilon}^T(n+1)\boldsymbol{\varepsilon}(n+1)]}$$

$$x(n+1) - \hat{x}(n+1) = \boldsymbol{A}[\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)] + \boldsymbol{w}(n) - \boldsymbol{G}(\boldsymbol{n})\{\boldsymbol{C}[\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)] + \boldsymbol{n}(n)\} =$$

$$\boldsymbol{\varepsilon}(n+1) = [\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{\varepsilon}(n) + \boldsymbol{w}(n) - \boldsymbol{G}(n)\boldsymbol{n}(n)$$

$$E[\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^T(n+1)] = E\left\{\big[[\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{\varepsilon}(n) + \boldsymbol{w}(n) - \boldsymbol{G}(n)\boldsymbol{n}(n)\big]\big[[\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{\varepsilon}(n) + \boldsymbol{w}(n) - \boldsymbol{G}(n)\boldsymbol{n}(n)\big]^T\right\}$$

$$= [\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]E[\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^T(n)][\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]^T + E[\boldsymbol{w}(n)\boldsymbol{w}^T(n)] + \boldsymbol{G}(n)E[\boldsymbol{n}(n)\boldsymbol{n}^T(n)]\boldsymbol{G}^T(n)$$

$$+[\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]E[\boldsymbol{\varepsilon}(n)\boldsymbol{w}^T(n)] + [\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]E[\boldsymbol{\varepsilon}(n)\boldsymbol{n}^T(n)]\,\boldsymbol{G}^T(n) + E[\boldsymbol{w}(n)\boldsymbol{\varepsilon}^T(n)][\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]^T - E[\boldsymbol{w}(n)\boldsymbol{n}^T(n)]\,\boldsymbol{G}^T(n)$$

$$-\boldsymbol{G}(n)E[\boldsymbol{n}(n)\varepsilon^T(n)][\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]^T - \boldsymbol{G}(n)E[\boldsymbol{n}(n)\boldsymbol{w}^T(n)]$$

$$= \boldsymbol{P}(n+1) = [\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{P}(n)[\boldsymbol{A} - \boldsymbol{G}(n)\boldsymbol{C}]^T + \boldsymbol{Q}(n) + \boldsymbol{G}(n)\boldsymbol{R}(n)\boldsymbol{G}^T(n)$$

$$E\{\boldsymbol{\varepsilon}(n)\boldsymbol{w}^T(n)\} = 0,$$
$$E\{\boldsymbol{\varepsilon}(n)\boldsymbol{n}^T(n)\} = 0$$
$$E\{\boldsymbol{w}(n)\boldsymbol{n}^T(n)\} = 0$$

$$\frac{\partial tr\boldsymbol{P}(n+1)}{\partial \boldsymbol{G}(n)} = \frac{\partial tr[(\boldsymbol{A}-\boldsymbol{G}(n)\boldsymbol{C})\boldsymbol{P}(n)(\boldsymbol{A}-\boldsymbol{G}(n)\boldsymbol{C})^T + \boldsymbol{Q}(n) + \boldsymbol{G}(n)\boldsymbol{R}(n)\boldsymbol{G}^T(n)]}{\partial \boldsymbol{G}(n)} = \boldsymbol{0}.$$

$$\frac{\partial tr\boldsymbol{P}(n+1)}{\partial \boldsymbol{G}(n)} = -2[\boldsymbol{A}-\boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{P}(n)\boldsymbol{C}^T + 2\boldsymbol{G}(n)\boldsymbol{R}(n) = 0.$$

$$\boxed{\boldsymbol{G}(n) = \boldsymbol{A}\boldsymbol{P}(n)\boldsymbol{C}^T[\boldsymbol{C}\boldsymbol{P}(n)\boldsymbol{C}^T + \boldsymbol{R}(n)]^{-1}}$$

A deriválásnál a következő szabályokat alkalmaztuk:

$$\frac{\partial tr[\boldsymbol{X}\boldsymbol{W}\boldsymbol{X}^T]}{\partial \boldsymbol{X}} = \boldsymbol{X}\boldsymbol{W}^T + \boldsymbol{X}\boldsymbol{W}; \quad \frac{\partial tr[\boldsymbol{X}\boldsymbol{W}]}{\partial \boldsymbol{X}} = \boldsymbol{W}^T; \quad \frac{\partial tr[\boldsymbol{W}\boldsymbol{X}^T]}{\partial \boldsymbol{X}} = \boldsymbol{W},$$

ahol a $\boldsymbol{W}$ mátrix az $\boldsymbol{X}$ mátrixtól független.

Optimális rekurzív becslő (vektor Kalman szűrő és prediktor) paraméterbecslésre: $\boldsymbol{A} = \boldsymbol{I}$ $\quad \boldsymbol{w}(n)$=0

$$\boxed{\boldsymbol{x}(n) = \boldsymbol{x}(n-1)}$$

$\boldsymbol{x}$ az ismeretlen paraméter

$$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{n}(n)$$

$\boldsymbol{C} = \boldsymbol{C}(n)$ a regressziós vektor

$$\boxed{\hat{\boldsymbol{x}}(n) = \hat{\boldsymbol{x}}(n-1) + \boldsymbol{K}(n)\big(\boldsymbol{y}(n) - \boldsymbol{C}\hat{\boldsymbol{x}}(n-1)\big)}$$

(1) $\boldsymbol{P}_1(n) = \boldsymbol{P}(n-1)$

(2) $\boldsymbol{K}(n) = \boldsymbol{P}_1(n)\boldsymbol{C}^T[\boldsymbol{C}\boldsymbol{P}_1(n)\boldsymbol{C}^T + \boldsymbol{R}(n)]^{-1}$

(3) $\boldsymbol{P}(n) = [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}]\boldsymbol{P}_1(n)$ $\quad n = 1, 2, \ldots$ $\quad \boldsymbol{P}(0)$ kell hozzá!

$$\boldsymbol{P}(n) = E\{[\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)][\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n)]^T\}$$

$\boxed{tr\boldsymbol{P}(n) \rightarrow \text{minimum!}}$

$$\boxed{\boldsymbol{x}(n+1) = \boldsymbol{x}(n)}$$

$\boldsymbol{x}$ az ismeretlen paraméter

$$\boldsymbol{y}(n) = \boldsymbol{C}\boldsymbol{x}(n) + \boldsymbol{n}(n)$$

$$\boxed{\hat{\boldsymbol{x}}(n+1) = \hat{\boldsymbol{x}}(n) + \boldsymbol{G}(n)[\boldsymbol{y}(n) - \boldsymbol{C}\hat{\boldsymbol{x}}(n)]}$$

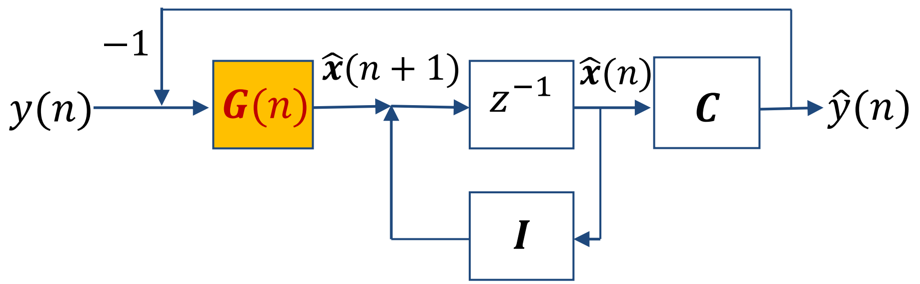
(1) $\boldsymbol{G}(n) = \boldsymbol{P}(n)\boldsymbol{C}^T[\boldsymbol{C}\boldsymbol{P}(n)\boldsymbol{C}^T + \boldsymbol{R}(n)]^{-1}$

(2) $\boldsymbol{P}(n+1)$= $[\boldsymbol{I} - \boldsymbol{G}(n)\boldsymbol{C}]\boldsymbol{P}(n)$

$n = 0, 1, 2, \ldots$ $\quad \boldsymbol{P}(0)$ kell hozzá!

$$\boldsymbol{P}(n+1) = E\{[\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)][\boldsymbol{x}(n+1) - \hat{\boldsymbol{x}}(n+1)]^T\}$$

$\boxed{tr\boldsymbol{P}(n+1) \rightarrow \text{minimum!}}$



6. LS becslők rekurzív számítása

(Jegyzet 71-75. oldalak) $\boxed{\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}}$ $\boxed{J(a, \hat{a}) = (\boldsymbol{z} - \boldsymbol{U}\boldsymbol{a})^T(\boldsymbol{z} - \boldsymbol{U}\boldsymbol{a})}$

6.1. A lineáris megfigyelési modellt alkalmazó LS becslők rekurzív számítása

A továbbiakban is n két dolgot jelöl: (1) az eddig figyelembe vett minták számát (az egyes minták indexelhetők például $k = 0,1, \dots, n-1$ jelöléssel); (2) Az n-edik „időpillanatban” vett mintát, ami már az $n+1$-edik mintavett adat. Ezzel a fenti összefüggés n minta feltételezésével: $\boldsymbol{z}(n) = \boldsymbol{U}(n)\boldsymbol{a}(n) + \boldsymbol{w}(n)$, amihez tartozóan tudjuk, hogy $\boxed{\hat{\boldsymbol{a}}(n) = [\boldsymbol{U}^T(n)\boldsymbol{U}(n)]^{-1}\boldsymbol{U}^T(n)\boldsymbol{z}(n) = \boldsymbol{P}(n)\boldsymbol{U}^T(n)\boldsymbol{z}(n).}$ $\boldsymbol{P}(n) = [\boldsymbol{U}^T(n)\boldsymbol{U}(n)]^{-1}$ Ha veszünk egy újabb mintát, akkor

$$\boldsymbol{z}(n+1) = \boldsymbol{U}(n+1)\boldsymbol{a}(n+1) + \boldsymbol{w}(n+1) = \begin{bmatrix} \boldsymbol{z}(n) \\ z(n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{U}(n) \\ \boldsymbol{u}(n) \end{bmatrix} \boldsymbol{a}(n+1) + \begin{bmatrix} \boldsymbol{w}(n) \\ w(n) \end{bmatrix},$$

ahol $z(n)$ skalár, az $n+1$-edik megfigyelési érték, $\boldsymbol{u}(n)$ sorvektor, a regressziós vektor $n+1$-edik sora, és végül $w(n)$ skalár, az $n+1$-edik zajkomponens.Ezekkel a jelölésekkel:

$$\boxed{\hat{\boldsymbol{a}}(n+1) = \left[[\boldsymbol{U}^T(n) \quad \boldsymbol{u}^T(n)] \begin{bmatrix} \boldsymbol{U}(n) \\ \boldsymbol{u}(n) \end{bmatrix} \right]^{-1} [\boldsymbol{U}^T(n) \quad \boldsymbol{u}^T(n)] \begin{bmatrix} \boldsymbol{z}(n) \\ z(n) \end{bmatrix},}$$

illetve a műveletek elvégzésével

$$\hat{\boldsymbol{a}}(n+1) = \left[[\boldsymbol{U}^T(n)\boldsymbol{U}(n) + \boldsymbol{u}^T(n)\boldsymbol{u}(n)] \right]^{-1} [\boldsymbol{U}^T(n)\boldsymbol{z}(n) + \boldsymbol{u}^T(n)z(n)] = \boldsymbol{P}(n+1)[\boldsymbol{U}^T(n)\boldsymbol{z}(n) + \boldsymbol{u}^T(n)z(n)].$$

Az ebben szereplő mátrix invertálást az ún. mátrix inverziós lemma felhasználásával végezzük:

$$\boxed{[\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D}]^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}\left[\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right]^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}}$$

$\boxed{[\boldsymbol{A}+\boldsymbol{BCD}]^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}\left[\boldsymbol{C}^{-1}+\boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B}\right]^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}}$ $\left[[\boldsymbol{U}^T(n)\boldsymbol{U}(n)+\boldsymbol{u}^T(n)\boldsymbol{u}(n)]\right]^{-1}=\boldsymbol{P}(n+1)$ ahol most

$\boldsymbol{A}=\boldsymbol{P}^{-1}(n); \boldsymbol{B}=\boldsymbol{u}^T(n); \boldsymbol{D}=\boldsymbol{u}(n); \boldsymbol{C}=1$, ezért $\boldsymbol{P}(n+1)=\boldsymbol{P}(n)-\boldsymbol{P}(n)\boldsymbol{u}^T(n)\underbrace{[1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)]}_{\text{Skalár!}}{}^{-1}\boldsymbol{u}(n)\boldsymbol{P}(n)=$

$=\boldsymbol{P}(n)-\dfrac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)\boldsymbol{u}(n)\boldsymbol{P}(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}$, $\boldsymbol{P}(0)$ ismeretében nincs további mátrix invertálás!

Fontos üzenet: Az egy diáddal módosított mátrix inverze mátrix szorzásokkal kiszámítható! Ennek felhasználásával:

$$\hat{\boldsymbol{a}}(n+1)=\boldsymbol{P}(n+1)[\boldsymbol{U}^T(n)\boldsymbol{z}(n)+\boldsymbol{u}^T(n)z(n)]=\hat{\boldsymbol{a}}(n)+\boldsymbol{P}(n)\boldsymbol{u}^T(n)z(n)-$$

$$-\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)\boldsymbol{u}(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}\hat{\boldsymbol{a}}(n)-\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)\boldsymbol{u}(n)\boldsymbol{P}(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}\boldsymbol{u}^T(n)z(n).$$

Itt a második és a negyedik tagot közös nevezőre hozva:

$$\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)z(n)+\boldsymbol{P}(n)\boldsymbol{u}^T(n)\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)z(n)-\boldsymbol{P}(n)\boldsymbol{u}^T(n)\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)z(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}=\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}z(n)=$$

$=\boldsymbol{G}(n)z(n)$ Ezzel: $\boxed{\hat{\boldsymbol{a}}(n+1)=\hat{\boldsymbol{a}}(n)+\boldsymbol{G}(n)[z(n)-\boldsymbol{u}(n)\hat{\boldsymbol{a}}(n)]}$ Itt $\boldsymbol{u}(n)\hat{\boldsymbol{a}}(n)=\hat{z}(n)$ az új mérés becslése!

$\boxed{\boldsymbol{G}(n)=\dfrac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)}{1+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}}$ $\boxed{\boldsymbol{P}(n+1)=[\boldsymbol{I}-\boldsymbol{G}(n)\boldsymbol{u}(n)]\boldsymbol{P}(n)}$ Az iteráció $n=0$-tól indul. $\hat{\boldsymbol{a}}(0)$=? $\boldsymbol{P}(0)$=?

Ezeket vagy meg tudjuk becsülni, vagy megoldjuk az LS becslést a paraméter vektor dimenziójának megfelelő mérési adat alapján, és annak eredményét használjuk az iteratív megoldás kiindulópontjaként.

$\boxed{J(a,\hat{a})=(\boldsymbol{z}-\boldsymbol{Ua})^T\boldsymbol{Q}(n)(\boldsymbol{z}-\boldsymbol{Ua})}$ ahol $\boldsymbol{Q}$ diagonális súlyozó mátrix, tehát $\boldsymbol{Q}(n)=diag\langle q_0,q_1,\dots,q_{n-1}\rangle$

ill. $Q(n+1)=diag\langle q_0,q_1,\dots,q_{n-1},q_n\rangle$

Ehhez a mátrix inverziós lemma $\boldsymbol{C}$ mátrixát $\boldsymbol{C}=1$ helyett $\boldsymbol{C}=q_n$ skalár értékre kell állítanunk. Ezzel:

$$\boldsymbol{G}(n)=\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)}{{}^1/_{q_n}+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}$$

Ha súlyozó mátrix egy diagonális kovariancia-mátrix inverze, azaz

$$\boldsymbol{Q}(n)=diag\left\langle {}^1/_{\sigma_0^2}, {}^1/_{\sigma_1^2},\dots,{}^1/_{\sigma_{n-1}^2}\right\rangle,\ \boldsymbol{Q}(n+1)=diag\left\langle {}^1/_{\sigma_0^2}, {}^1/_{\sigma_1^2},\dots,{}^1/_{\sigma_{n-1}^2},{}^1/_{\sigma_n^2}\right\rangle,$$

akkor :

$$\boldsymbol{G}(n)=\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)}{\sigma_n^2+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)}$$

A többi változatlan!

Ha a kovariancia mátrix nem lenne diagonális, akkor a megismert fehérítési eljárás alkalmazásával, azaz egy transzformáció közbeiktatásával juthatunk el a rekurzív forma alkalmazhatóságához.

Mindezekkel együtt rekurzív megoldást tudunk adni a Gauss-Markov, a BLUE és a súlyozott LS becslések esetére!

Egy további hasznos súlyozási forma lehet a $\boldsymbol{Q}(n)=diag\langle\beta^{n-1},\beta^{n-2},\dots,1\rangle$ $\boldsymbol{Q}(n+1)=diag\langle\beta^{n},\beta^{n-1},\beta^{n-2},\dots,1\rangle$

$0<\beta<1$. Itt a legutolsó mérési adattal képzett négyzetes hiba 1 súllyal szerepel, míg a korábbiak rendre kisebbel.

$$\boldsymbol{G}(n)=\frac{\boldsymbol{P}(n)\boldsymbol{u}^T(n)}{\beta+\boldsymbol{u}(n)\boldsymbol{P}(n)\boldsymbol{u}^T(n)},\ \boldsymbol{P}(n+1)=\frac{1}{\beta}[\boldsymbol{I}-\boldsymbol{G}(n)\boldsymbol{u}(n)]\boldsymbol{P}(n)$$

Ezzel egy felejtő hatást érvényesítünk, ami nemstacionárius folyamatok esetén előnyös.

6.2. LS becslők számítása kényszerfeltétel esetén

Tételezzük fel, hogy a becsülendő $\boldsymbol{a}$ paraméter kielégíti az $\boldsymbol{A}\boldsymbol{a}=\boldsymbol{b}$ kényszerfeltételt!

Ekkor feltételes szélsőértéket keresünk a Lagrange multiplikátoros technika segítségével:

$$J(\boldsymbol{a})=(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})+\boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{a}-\boldsymbol{b}),\quad \frac{\partial J(\boldsymbol{a})}{\partial\boldsymbol{a}}=-2\boldsymbol{U}^T(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})+\boldsymbol{A}^T\boldsymbol{\lambda}=0.$$

Ebből a kényszerfeltételt kielégítő $\hat{\boldsymbol{a}}_C$ megoldás: $\hat{\boldsymbol{a}}_C=[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{z}-[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T\frac{\boldsymbol{\lambda}}{2}.$

Mivel ez kielégíti a kényszerfeltételt ($\hat{\boldsymbol{a}}$ a kényszer nélküli becslőt jelöli):

$$\boldsymbol{A}\hat{\boldsymbol{a}}_C=\boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{z}-\boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T\frac{\boldsymbol{\lambda}}{2}=\boldsymbol{A}\hat{\boldsymbol{a}}-\boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T\frac{\boldsymbol{\lambda}}{2}=\boldsymbol{b}$$

$\boldsymbol{A}\hat{\boldsymbol{a}}_C = \boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{z} - \boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T\frac{\boldsymbol{\lambda}}{2} = \boldsymbol{A}\hat{\boldsymbol{a}} - \boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T\frac{\boldsymbol{\lambda}}{2} = \boldsymbol{b}$ Ebből: $\frac{\boldsymbol{\lambda}}{2} = [\boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T]^{-1}(\boldsymbol{A}\hat{\boldsymbol{a}} - \boldsymbol{b})$,

amivel $\boxed{\hat{\boldsymbol{a}}_C = \hat{\boldsymbol{a}} - [\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T[\boldsymbol{A}[\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{A}^T]^{-1}(\boldsymbol{A}\hat{\boldsymbol{a}} - \boldsymbol{b}).}$

6.3. Nemlineáris megfigyelési modellt alkalmazó LS becslők számítása

A megfigyelési modell: $\boldsymbol{s}(\boldsymbol{a}) \neq \boldsymbol{U}\boldsymbol{a}$. A kritériumfüggvény deriváltja:

$$\frac{\partial}{\partial \boldsymbol{a}}(\boldsymbol{z} - \boldsymbol{s}(\boldsymbol{a}))^T(\boldsymbol{z} - \boldsymbol{s}(\boldsymbol{a})) = -2\frac{\partial \boldsymbol{s}^T(\boldsymbol{a})}{\partial \boldsymbol{a}}(\boldsymbol{z} - \boldsymbol{s}(\boldsymbol{a})) = -2\boldsymbol{g}(\boldsymbol{a}) = 0.$$

A Newton módszer szerint a $\boldsymbol{g}(\boldsymbol{a}) = \boldsymbol{0}$ egyenlet iteratív megoldása:

$$\left.\frac{\partial \boldsymbol{g}(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} = \frac{\boldsymbol{g}(\boldsymbol{a}_0)}{\boldsymbol{a}_0 - \boldsymbol{a}_1}, \quad \boldsymbol{a}_1 = \boldsymbol{a}_0 - \frac{\boldsymbol{g}(\boldsymbol{a}_0)}{\left.\frac{\partial \boldsymbol{g}(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}}, \quad \boldsymbol{a}_2 = \boldsymbol{a}_1 - \frac{\boldsymbol{g}(\boldsymbol{a}_1)}{\left.\frac{\partial \boldsymbol{g}(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_1}} \cdots$$

De kiindulhatunk $\boldsymbol{s}(\boldsymbol{a})$ Taylor sorából is:

$\left.\frac{\partial \boldsymbol{s}(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} = \boldsymbol{U}(\boldsymbol{a}_0)$ jelöléssel, és a második deriváltat tartalmazó tag elhanyagolásával: $\boldsymbol{s}(\boldsymbol{a}) \cong \boldsymbol{s}(\boldsymbol{a}_0) + \boldsymbol{U}(\boldsymbol{a}_0)(\boldsymbol{a} - \boldsymbol{a}_0)$,

$$\left.\frac{\partial \boldsymbol{s}^T(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}(\boldsymbol{z} - \boldsymbol{s}(\boldsymbol{a})) = \boldsymbol{U}^T(\boldsymbol{a}_0)\big[\boldsymbol{z} - [\boldsymbol{s}(\boldsymbol{a}_0) + \boldsymbol{U}(\boldsymbol{a}_0)(\boldsymbol{a} - \boldsymbol{a}_0)]\big] = 0, \quad \boxed{\hat{\boldsymbol{a}} \cong \boldsymbol{a}_0 + [\boldsymbol{U}^T(\boldsymbol{a}_0)\boldsymbol{U}(\boldsymbol{a}_0)]^{-1}\boldsymbol{U}^T(\boldsymbol{a}_0)[\boldsymbol{z} - \boldsymbol{s}(\boldsymbol{a}_0)].}$$

De kiindulhatunk $J(\boldsymbol{a})$ Taylor sorából is: $J(\boldsymbol{a}) = J(\boldsymbol{a}_0) + \left.\frac{\partial J(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}(\boldsymbol{a} - \boldsymbol{a}_0) + \frac{1}{2}(\boldsymbol{a} - \boldsymbol{a}_0)^T\left.\frac{\partial^2 J(\boldsymbol{a})}{\partial^2 \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}(\boldsymbol{a} - \boldsymbol{a}_0) + \cdots$

Ha $J(\boldsymbol{a})|_{min} = 0$, akkor kiindulhatunk az első két tagból, és a többit elhanyagoljuk. $\left.\frac{\partial J(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} = \nabla J(\boldsymbol{a}_0)$ jelöléssel:

$$\nabla J^T(\boldsymbol{a}_0)\nabla J(\boldsymbol{a}_0)\hat{\boldsymbol{a}} = \nabla J^T(\boldsymbol{a}_0)\nabla J(\boldsymbol{a}_0)\boldsymbol{a}_0 - \nabla J^T(\boldsymbol{a}_0)J(\boldsymbol{a}_0), \quad \boxed{\hat{\boldsymbol{a}} \cong \boldsymbol{a}_0 - [\nabla J^T(\boldsymbol{a}_0)\nabla J(\boldsymbol{a}_0)]^{-1}\nabla J^T(\boldsymbol{a}_0)J(\boldsymbol{a}_0)}$$

Ha $J(\boldsymbol{a})|_{min} \neq 0$, akkor a $\nabla J(\boldsymbol{a}) = \mathbf{0}$ feltétel teljesülését keressük az első és a másodrendű tag figyelembevételével, a többi elhanyagolásával:

$$\nabla J(\boldsymbol{a}) = 0 = \left.\frac{\partial J(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} + \left.\frac{\partial^2 J(\boldsymbol{a})}{\partial^2 \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} (\boldsymbol{a} - \boldsymbol{a}_0),$$

$$\boxed{\hat{\boldsymbol{a}} = \boldsymbol{a}_0 - \left[\left.\frac{\partial^2 J(\boldsymbol{a})}{\partial^2 \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}\right]^{-1} \left.\frac{\partial J(\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0} = \boldsymbol{a}_0 - \left[\left.\frac{\partial^2 J(\boldsymbol{a})}{\partial^2 \boldsymbol{a}}\right|_{\boldsymbol{a}=\boldsymbol{a}_0}\right]^{-1} \nabla J(\boldsymbol{a}_0)}$$

amit Gauss-Newton módszernek is nevezünk.

Megjegyzés: Egyes feladatok alkalmas transzformációval visszavezethetők lineáris paraméter-függésre. Például a következő hullámforma

$$s(n) = A\cos(2\pi f_0 n + \varphi), \qquad n = 0,1,\dots,N-1$$

ismeretlen A és φ paramétereinek együttes meghatározása nemlineáris LS probléma.

Ha azonban az ezzel ekvivalens felírás

$$s(n) = A\cos(2\pi f_0 n + \varphi) = B\cos(2\pi f_0 n) + C\sin(2\pi f_0 n)$$

B és C paramétereit becsüljük, akkor a sokkal egyszerűbb lineáris LS problémát kell megoldanunk, és a keresett jellemzőket pedig transzformációval kapjuk:

$$A = \sqrt{B^2 + C^2}, \qquad \varphi = \arctan\left(\frac{-C}{B}\right)$$