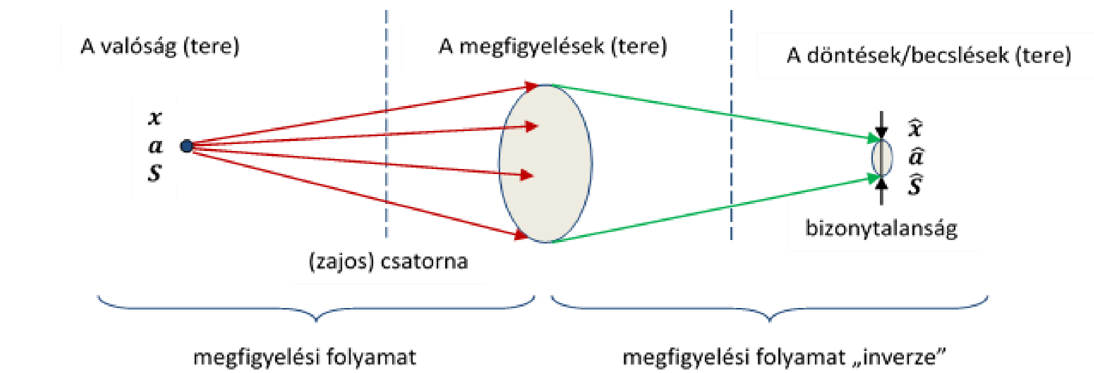
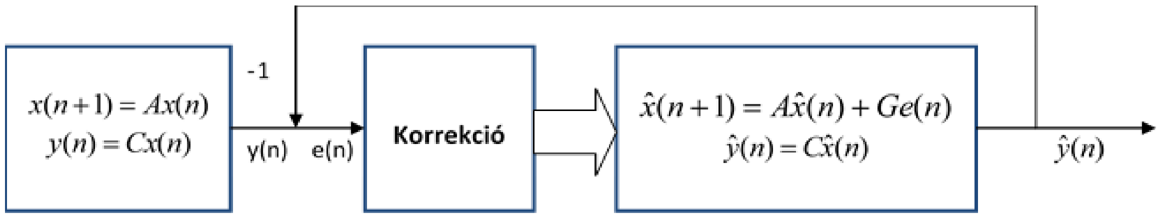
Méréselmélet

3. fejezet: A becsléselmélet alapjai

2022. március 9.

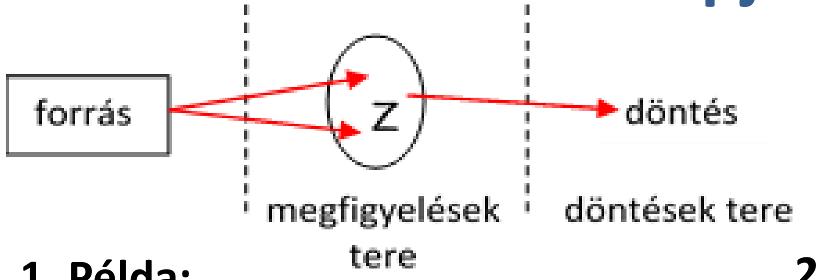
Ismétlés

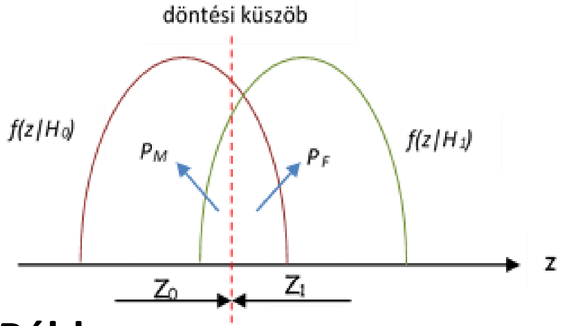


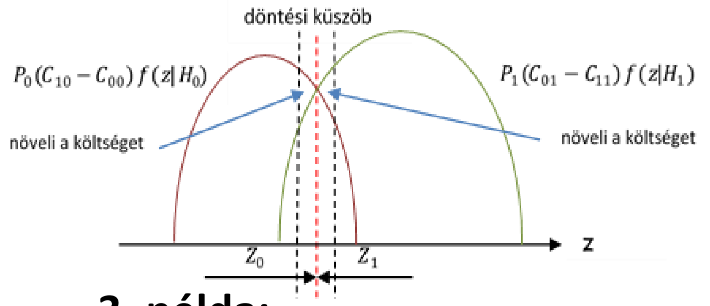


$$P_0(C_{10} - C_{00})f(z|H_0) = P_1(C_{01} - C_{11})f(z|H_1).$$

2. A döntéselmélet alapjai

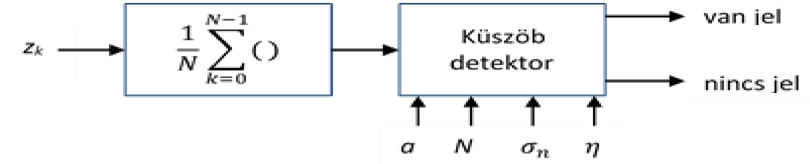




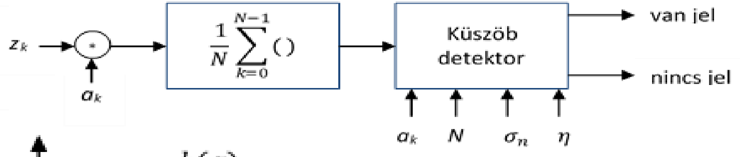


$$\Lambda(z) \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

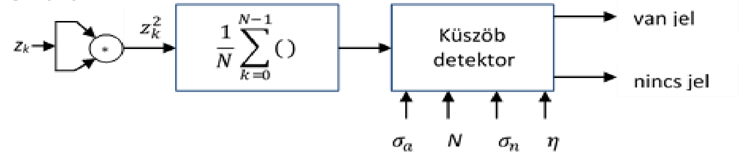
1. Példa:



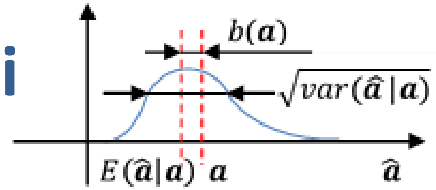
2. Példa:



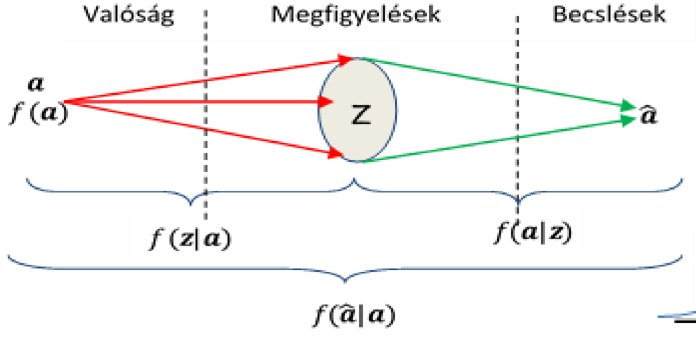
3. példa:

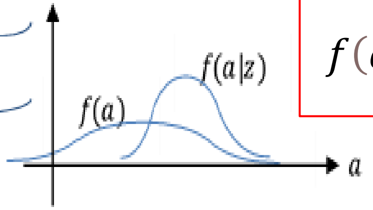


3. A becsléselmélet alapjai



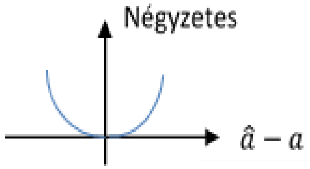
$$R(\hat{a},a) = E\{C(\hat{a},a)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(\hat{a},a)f(a,z)\,da\,dz$$

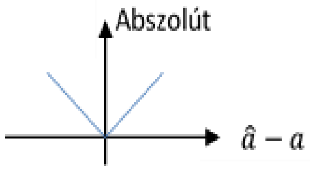


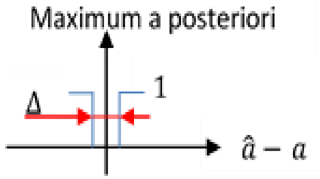


3.1. Bayes becslők

$$f(a|z) = \frac{f(z|a)f(a)}{f(z)},$$







$$\hat{a}_{MS} = \int_{-\infty}^{\infty} af(a|z)da$$

$\hat{a}_{ABS} = f(a|z)$ **mediánja.**

$\hat{a}_{MAP} = f(a|z)$ **maximumhelye.**

Bayes becslő Gauss eloszlások esetén A megfigyelés modellje legyen lineáris! $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{n}$

$$\hat{\boldsymbol{a}}_{MS} = \boldsymbol{\mu}_{\boldsymbol{a}|\boldsymbol{z}} = \underbrace{\boldsymbol{\mu}_{\boldsymbol{a}}}_{\text{apriori ismeret}} + \underbrace{[\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U} + \boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}(\boldsymbol{z} - \boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}})}_{\text{korrekció } \boldsymbol{z}-\boldsymbol{U}\boldsymbol{\mu}_{\boldsymbol{a}} \text{ függvényében}}$$

$$cov\{a, a|\boldsymbol{z}\} = \boldsymbol{\Sigma}_{\boldsymbol{aa}|\boldsymbol{z}} = [\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U} + \boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1}$$

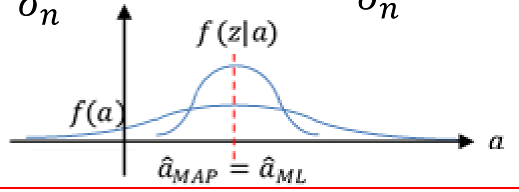
1. Példa: DC szint zajban:

$$\hat{a}_{MS} = \frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} z_k}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}, \quad var\{\tilde{a}\} = \frac{\sigma_a^2}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}, \quad b(a) = \frac{\mu_a - a}{1 + N\frac{\sigma_a^2}{\sigma_n^2}},$$

2. Példa: ismert jel ismeretlen amplitúdója:

$$\hat{a}_{MAP} = \frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k z_k}{1 + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}, \quad var\{\tilde{a}\} = \frac{\sigma_a^2}{1 + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}$$

3.2. Maximum Likelihood (ML) becslő



$$\left.\frac{\partial f(\boldsymbol{z}|\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{a=\hat{a}_{ML}} = 0, \qquad \left.\frac{\partial \ln f(\boldsymbol{z}|\boldsymbol{a})}{\partial \boldsymbol{a}}\right|_{a=\hat{a}_{ML}} = 0$$

3.2.1. Gauss-Markov (GM) becslő:

Speciális ML becslő: a megfigyelési zaj **Gauss** eloszlású, a megfigyelési egyenlet **lineáris**.

$$f(\boldsymbol{z}|\boldsymbol{a}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{nn}}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})}$$

$$\hat{\boldsymbol{a}}_{GM} = [\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{z}$$

$$cov(\hat{\boldsymbol{a}}_{GM}, \hat{\boldsymbol{a}}_{GM}) = [\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U}]^{-1}$$

Példa: DC szint zajban:

$$f(\boldsymbol{z}|a) = \frac{1}{(2\pi\sigma_n^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1}(z_k - a)^2}, \qquad \hat{a}_{ML} = \hat{a}_{GM} = \frac{1}{N}\sum_{k=0}^{N-1} z_k,$$

Lineáris megfigyelési egyenlet és Gauss eloszlású csatorna zaj esetén a legjobb (Gauss-Markov) becslő **az egyszerű átlagolás!**

3.3. Becslők determinisztikus modellel jellemzett paraméterek esetén

3.3.1. Becslések minősítése

Torzítatlan becslő: Az a becslő, amelyik várható értékben a helyes értéket adja: $E(a - \hat{a}) = 0$

$$var(\hat{A}_1) = var[\tfrac{1}{N}\sum_{k=0}^{N-1} z_k] = E\left\{\left(\left(\tfrac{1}{N}\sum_{k=0}^{N-1} z_k\right) - E\left\{\tfrac{1}{N}\sum_{k=0}^{N-1} z_k\right\}\right)^2\right\} = E\left\{\left(\tfrac{1}{N}\sum_{k=0}^{N-1}(z_k - A)\right)^2\right\} = \frac{1}{N^2}\sum_{k=0}^{N-1} var(z_k) = \frac{1}{N^2}N\sigma^2 = \frac{\sigma^2}{N},$$

$$var(\hat{A}_2) = var(z_0) = \sigma^2 > var(\hat{A}_1)$$

3.3.2. A minimális variancia kritériuma

$$mse(\hat{a}) = E[(\hat{a} - a)^2]$$

Torzítás

Zavarás

$$mse(\hat{a}) = E\{[\hat{a} - E[\hat{a}] + E[\hat{a}] - a]^2\} = E\{[\hat{a} - E[\hat{a}] + b(a)]^2\} = var(\hat{a}) + b^2(a).$$

Ismétlés

3.3.3. Minimális varianciájú, torzítatlan becslők (Minimum Variance Unbiased Estimator, MVU Estimator)

Nem biztos, hogy MVU becslő létezik! Lehet, hogy **nincsen** torzítatlan becslő. Lehet, hogy **egyik** becslő **se minimális varianciájú**.

Érdekes módon van módszer arra, hogy meghatározzuk, hogy mekkora az elvileg elérhető legkisebb variancia!

Cramer-Rao alsó korlát (Cramer-Rao Lower Bound: CRLB, skalár paraméter esete) $var(\hat{a}) \geq CRLB(a)$

Ha $E\left[\frac{\partial \ln f(z;a)}{\partial a}\right] = 0$ akkor $var(\hat{a}) \geq \left[-E\left(\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right)\right]^{-1}$ $\frac{\partial \ln f(z;a)}{\partial a} = I(a)(g(z) - a)$ $\hat{a} = g(z)$ a variancia minimum: $= \frac{1}{I(a)}$.

1. Példa: $z_0 = A + w_0,$ $\frac{\partial \ln f(z_0;A)}{\partial A} = \frac{1}{\sigma^2}[z_0 - A],$ $-\frac{\partial^2 \ln f(z_0;A)}{\partial A^2} = \frac{1}{\sigma^2}.$ $\hat{A} = g(z_0) = z_0,$ $var(\hat{A}) \geq \sigma^2$

2. Példa: $z_k = A + w_k,$ $\frac{\partial \ln f(z;A)}{\partial A} = \frac{1}{\sigma^2}\sum_{k=0}^{N-1}(z_k - A),$ $-\frac{\partial^2 \ln f(z;A)}{\partial A^2} = \frac{N}{\sigma^2}.$ $\hat{A} = \frac{1}{N}\sum_{k=0}^{N-1} z_k,$ $var(\hat{A}) \geq \frac{\sigma^2}{N}.$

3. Példa: $z_k = A + w_k,$ $\frac{\partial \ln f(z;\sigma^2)}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{k=0}^{N-1}(z_k - A)^2$ $= \frac{N}{2\sigma^4}\left[\frac{1}{N}\sum_{k=0}^{N-1}(z_k - A)^2 - \sigma^2\right]$ $\widehat{\sigma^2} = \frac{1}{N}\sum_{k=0}^{N-1}(z_k - A)^2$

Zaj variancia varianciája:

$$-E\left[\frac{\partial^2 \ln f(z;\sigma^2)}{\partial(\sigma^2)^2}\right] = -\frac{N}{2\sigma^4} + \frac{N\sigma^2}{\sigma^6} = \frac{N}{2\sigma^4}$$

$var(\widehat{\sigma^2}) \geq \frac{2\sigma^4}{N}.$

4. Példa: $x_k = s_k(a) + w_k,$

$$var(\hat{a}) \geq \frac{\sigma_w^2}{\sum_{k=0}^{N-1}\left[\frac{\partial s_k(a)}{\partial a}\right]^2}.$$

5. Példa: $\alpha = g(a)$ paraméter varianciájának alsó korlátja:

$$var(\hat{\alpha}) \geq \frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left\{\left[\frac{\partial \ln f(z;a)}{\partial a}\right]^2\right\}} = -\frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right]}.$$

6. Példa: DC szint (A) teljesítményének (A^2) „mérése” a DC szint mérésére alapozva.

$$var(\hat{\alpha}) = var(\hat{A}^2) \geq \frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left\{\left[\frac{\partial \ln f(z;a)}{\partial a}\right]^2\right\}} = -\frac{\left[\frac{\partial g(a)}{\partial a}\right]^2}{E\left[\frac{\partial^2 \ln f(z;a)}{\partial a^2}\right]} = \frac{(2A)^2}{\frac{N}{\sigma^2}} = \frac{4A^2\sigma^2}{N}$$

Mivel $E\{\hat{A}^2\} = E^2(\hat{A}) + var(\hat{A}) = A^2 + \frac{\sigma^2}{N} \neq A^2$ A becslés nem torzítatlan.

Mivel $var(\hat{A}^2) = \frac{4A^2\sigma^2}{N} + \frac{2\sigma^4}{N^2} > CRLB$ a javasolt becslő nem lesz hatásos.

Cramer-Rao alsó korlát (Cramer-Rao Lower Bound: CRLB, vektor-paraméter esete)

Tegyük fel, hogy a mérési (vektor)adatok valószínűség sűrűségfüggvényére teljesül a regularitási feltétel minden $\boldsymbol{a}$ esetében:

$$E\left[\frac{\partial lnf(z;\boldsymbol{a})}{\partial \boldsymbol{a}}\right]=0.$$

Ekkor bármely torzítatlan becslőre igaz, hogy a becslő kovariancia mátrixának és az ún. **Fisher információs mátrix** inverzének a különbsége pozitív szemidefinit: $\boldsymbol{C}_{\hat{\boldsymbol{a}}}-\boldsymbol{I}^{-1}(\boldsymbol{a})\geq 0$

A Fisher információs mátrix elemei:

$$\boldsymbol{I}_{ij}(\boldsymbol{a})=-E\left(\frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial a_i\partial a_j}\right)$$

Egy olyan torzítatlan becslő, amelyik eléri a Cramer-Rao alsó korlátot akkor és csak akkor található, ha:

$$\frac{\partial lnf(z;\boldsymbol{a})}{\partial \boldsymbol{a}}=\boldsymbol{I}(\boldsymbol{a})(\boldsymbol{g}(z)-\boldsymbol{a})$$

Ilyenkor a becslő $\hat{\boldsymbol{a}}=\boldsymbol{g}(z)$, és a kovariancia minimum $\boldsymbol{I}^{-1}(\boldsymbol{a})$.

7. Példa: Additív, Gauss eloszlású fehér zajjal terhelt **DC szint és szórásának** mérése mérési sorozatra alapozva:

$z_k=A+w_k$, $k=0,1,\dots,N-1$, ahol a w_k korrelálatlan minden mintával, és valószínűségsűrűség függvénye: $\mathcal{N}(0,\sigma^2)$.

Most $\boldsymbol{a}=[A\quad \sigma^2]^T$. A sűrűség függvény logaritmusa: $lnf(z;\boldsymbol{a})=-\frac{N}{2}ln2\pi-\frac{N}{2}ln\sigma^2-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}(z_k-A)^2$.

$$\boldsymbol{I}(\boldsymbol{a})=-E\begin{bmatrix}\frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A^2} & \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A\partial\sigma^2}\\ \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial A\partial\sigma^2} & \frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial(\sigma^2)^2}\end{bmatrix}=\begin{bmatrix}\frac{N}{\sigma^2} & 0\\ 0 & \frac{N}{2\sigma^4}\end{bmatrix}.$$

Ez a mátrix diagonális, így könnyen invertálható, amiből:

$$\boldsymbol{C}_{\hat{\boldsymbol{a}}}=\begin{bmatrix}var(\hat{A}) & 0\\ 0 & var(\widehat{\sigma^2})\end{bmatrix}\geq\begin{bmatrix}\frac{\sigma^2}{N} & 0\\ 0 & \frac{2\sigma^4}{N}\end{bmatrix}$$

8. Példa: Ha $\alpha=g(\boldsymbol{a})$ a becsülendő mennyiség, és az $\hat{\boldsymbol{a}}$ kovarianciája $\boldsymbol{I}^{-1}(\boldsymbol{a})$, akkor

$$\boldsymbol{C}_{\hat{\alpha}}\geq\left(\frac{\partial g(\boldsymbol{a})}{\partial\boldsymbol{a}}\right)\left[-E\left[\frac{\partial^2 lnf(z;\boldsymbol{a})}{\partial\boldsymbol{a}^2}\right]\right]^{-1}\left(\frac{\partial g(\boldsymbol{a})}{\partial\boldsymbol{a}}\right)^T,$$

ahol a függvénykapcsolat paraméter szerinti deriváltját sorvektorként értelmezzük.

Ha az előző példához kapcsolódóan az $\alpha=g(\boldsymbol{a})=A^2/\sigma^2$ jel/zaj viszony becslése varianciájának alsó határát keressük, akkor

$$\frac{\partial g(\boldsymbol{a})}{\partial\boldsymbol{a}}=\begin{bmatrix}\frac{\partial g(\boldsymbol{a})}{\partial A} & \frac{\partial g(\boldsymbol{a})}{\partial\sigma^2}\end{bmatrix}=\begin{bmatrix}\frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4}\end{bmatrix},$$

Ezt felhasználva a keresett alsó határ:

$$var(\hat{\alpha})\geq\begin{bmatrix}\frac{2A}{\sigma^2} & -\frac{A^2}{\sigma^4}\end{bmatrix}\begin{bmatrix}\frac{\sigma^2}{N} & 0\\ 0 & \frac{2\sigma^4}{N}\end{bmatrix}\begin{bmatrix}\frac{2A}{\sigma^2}\\ -\frac{A^2}{\sigma^4}\end{bmatrix}=\frac{4A^2}{N\sigma^2}+\frac{2A^4}{N\sigma^4}=\frac{4\alpha+2\alpha^2}{N}$$

Megjegyzés: Az előzőekben bemutatott példák közös jellemzője, hogy **determinisztikus paraméter**, és **ismert** valószínűségsűrűség függvényű **csatorna-karakterisztika** feltételezésével a becslés varianciájának **elvi minimumát** kapjuk meg.

Az, hogy ezt **milyen mérési eljárás** alkalmazása esetén érjük el, ill. **létezik-e, ismert-e** ilyen eljárás, sok esetben **nyitott kérdés.**

A továbbiakban korlátozzuk magunkat arra az esetre, amikor a **megfigyelési modell lineáris**.

Látni fogjuk, hogy ilyenkor hatásos becslőkhöz jutunk, azaz minden esetben elérjük a becslés varianciájára meghatározható alsó korlátot (CRLB).

Additív, Gauss eloszlású fehér zajjal terhelt lineáris modellek esete: $\boxed{\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}}$

z	N*1	dimenziós megfigyelési vektor
U	N*M	dimenziós, ismert megfigyelési mátrix
a	M*1	dimenziós, becsülendő paramétervektor
w	N*1	dimenziós zaj vektor, amelynek valószínűség sűrűségfüggvénye N(0,$\sigma^2\boldsymbol{I}$).

a megfigyelési vektor minden eleme az ismeretlen paraméterek lineáris kombinációjaként áll elő, amit additív zaj terhel.

Ahhoz, hogy M paramétert becsülni tudjunk legalább ugyanennyi megfigyelést/mérést kell elvégeznünk, azaz $N \geq M$, ráadásul a zaj hatásának csökkentése érdekében tipikusan $N > M$ vagy $N \gg M$.

Ebben az esetben a CRLB és az ezt elérő MVU becslő kiszámítható:

1. lépés: $\ln f(\boldsymbol{z};\boldsymbol{a})$ számítása;

2. lépés: Az $\boldsymbol{I}(\boldsymbol{a}) = -E\left[\frac{\partial^2 \ln f(\boldsymbol{z};\boldsymbol{a})}{\partial \boldsymbol{a}^2}\right]$ Fisher információs mátrix és abból az $\hat{\boldsymbol{a}}$ kovariancia mátrixának kiszámítása: $\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \boldsymbol{I}^{-1}(\boldsymbol{a})$.

3. lépés: A $g(\boldsymbol{z})$ MVU becslő meghatározása a következő szorzat alakból: $\boxed{\frac{\partial \ln f(\boldsymbol{z};\boldsymbol{a})}{\partial \boldsymbol{a}} = \boldsymbol{I}(\boldsymbol{a})[g(\boldsymbol{z}) - \boldsymbol{a}].}$

1. lépés: $\ln f(\boldsymbol{z};\boldsymbol{a}) = -\ln\left(\sqrt{2\pi\sigma^2}\right)^N - \frac{1}{2\sigma^2}(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})^T(\boldsymbol{z}-\boldsymbol{U}\boldsymbol{a})$

2. lépés: $\frac{\partial \ln f(\boldsymbol{z},\boldsymbol{a})}{\partial \boldsymbol{a}} = -\frac{1}{2\sigma^2}\frac{\partial}{\partial \boldsymbol{a}}[\boldsymbol{z}^T\boldsymbol{z} - 2z^T\boldsymbol{U}\boldsymbol{a} + \boldsymbol{a}^T\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{a}] = \frac{1}{\sigma^2}[\boldsymbol{U}^T\boldsymbol{z} - \boldsymbol{U}^T\boldsymbol{U}\boldsymbol{a}]$,

$$\boldsymbol{I}(\boldsymbol{a}) = -E\left[\frac{\partial^2 \ln f(\boldsymbol{z};\boldsymbol{a})}{\partial \boldsymbol{a}^2}\right] = \frac{1}{\sigma^2}\boldsymbol{U}^T\boldsymbol{U},$$

$$\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \boldsymbol{I}^{-1}(\boldsymbol{a}) = \sigma^2[\boldsymbol{U}^T\boldsymbol{U}]^{-1}$$

3. lépés: $\frac{\partial \ln f(\boldsymbol{z};\boldsymbol{a})}{\partial \boldsymbol{a}} = \boldsymbol{I}(\boldsymbol{a})[g(\boldsymbol{z}) - \boldsymbol{a}] = \frac{\boldsymbol{U}^T\boldsymbol{U}}{\sigma^2}[(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z} - \boldsymbol{a}]$.

$$\boxed{\hat{\boldsymbol{a}} = g(\boldsymbol{z}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z},} \quad \boxed{\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \boldsymbol{I}^{-1}(\boldsymbol{a}) = \sigma^2(\boldsymbol{U}^T\boldsymbol{U})^{-1}}$$

Vagyis **additív**, **Gauss** eloszlású **fehér** zajjal terhelt **lineáris** modellek esetében:

- Az MVU becslő: $\hat{\boldsymbol{a}} = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z}$. A becslő **hatásos** és **eléri a CLRB korlátot**;
- A becslő **torzítatlan**, mert: $E(\hat{\boldsymbol{a}}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T E(\boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}) = \boldsymbol{a}$;
- A becslő statisztikai viselkedése **teljes mértékben specifikált**, mert $\hat{\boldsymbol{a}}$ a Gauss eloszlású $\boldsymbol{z}$ vektornak **lineáris transzformáltja**:

$$\boxed{\hat{\boldsymbol{a}} \sim \mathcal{N}\left(\boldsymbol{a}, \sigma^2(\boldsymbol{U}^T\boldsymbol{U})^{-1}\right).}$$

Példák ennek a modellcsaládnak az alkalmazására, ahol: $\boxed{\hat{\boldsymbol{a}} = g(\boldsymbol{z}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z},}$ $\boxed{\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \boldsymbol{I}^{-1}(\boldsymbol{a}) = \sigma^2(\boldsymbol{U}^T\boldsymbol{U})^{-1}}$

1. Példa: az illesztendő modell a diszkrét n időindex polinomja: $x_n = a_0 + a_1 n + a_2 n^2 + \cdots + a_{M-1}n^{M-1} + w_n$

$$\boxed{\boldsymbol{z} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^{M-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & N-1 & (N-1)^2 & \dots & (N-1)^{M-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w},}$$

$$\boxed{\hat{\boldsymbol{a}} = g(\boldsymbol{z}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z},}$$

$$\boxed{\hat{\boldsymbol{a}} \sim \mathcal{N}\left(\boldsymbol{a}, \sigma^2(\boldsymbol{U}^T\boldsymbol{U})^{-1}\right)}$$

2. Példa: Az illesztendő modell az időben diszkrét Fourier sorfejtés: $x_n = \sum_{k=1}^{M} a_k \cos\frac{2\pi kn}{N} + \sum_{k=1}^{M} b_k \sin\frac{2\pi kn}{N} + w_n$
Most DC komponenst nem illesztünk.

$$\boxed{\boldsymbol{z} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} \cdots & 1 & \cdots & \cdots & 0 & \cdots \\ \cdots & \cos\frac{2\pi k}{N} & \cdots & \cdots & \sin\frac{2\pi k}{N} & \cdots \\ \cdots & \cos\frac{4\pi k}{N} & \cdots & \cdots & \sin\frac{4\pi k}{N} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cos\frac{2\pi}{N}(N-1) & \cdots & \cdots & \sin\frac{2\pi}{N}(N-1) & \cdots \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \\ b_1 \\ b_2 \\ \vdots \\ b_M \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}}$$

$$\boxed{\hat{\boldsymbol{a}} = g(\boldsymbol{z}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z},}$$

$$\boxed{\hat{\boldsymbol{a}} \sim \mathcal{N}\left(\boldsymbol{a}, \sigma^2(\boldsymbol{U}^T\boldsymbol{U})^{-1}\right)}$$

$$(\boldsymbol{U}^T\boldsymbol{U})^{-1} = \frac{2}{N}\boldsymbol{I},$$

$$\hat{a}_k = \frac{2}{N}\sum_{n=0}^{N-1} x_n \cos\frac{2\pi}{N}kn,$$

$$\hat{b}_k = \frac{2}{N}\sum_{n=0}^{N-1} x_n \sin\frac{2\pi}{N}kn.$$

A becslők egymástól függetlenek. $\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \sigma^2[\boldsymbol{U}^T\boldsymbol{U}]^{-1} = \frac{2}{N}\sigma^2\boldsymbol{I}$

3. Példa: A mozgó átlagoló (FIR szűrő), amikor a gerjesztés belépő függvény, tehát $x_n = 0$, ha $n < 0$: $\boxed{y_n = \sum_{k=0}^{M-1} a_k\, x_{n-k} + w_k}$

$$\boxed{\boldsymbol{z} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} x_0 & 0 & \cdots & \cdots & 0 \\ x_1 & x_0 & \cdots & \cdots & 0 \\ x_2 & x_1 & x_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & 0 \\ x_{N-1} & x_{N-2} & x_{N-3} & \cdots & x_{N-M} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{M-1} \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}}$$

$$\boxed{\hat{\boldsymbol{a}} = g(\boldsymbol{z}) = (\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{z},}$$

$$\boxed{\hat{\boldsymbol{a}} \sim \mathcal{N}\left(\boldsymbol{a}, \sigma^2 (\boldsymbol{U}^T\boldsymbol{U})^{-1}\right)}$$

4. Példa: Mozgó-átlag (Moving Average: MA) paraméterek becslése. Most nem belépő függvényre: $\boxed{y_n = \sum_{k=0}^{M-1} a_k\, x_{n-k} + w_k}$

$$\boldsymbol{U}^T\boldsymbol{U} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \\ x_{-1} & x_0 & \cdots & x_{N-2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1-M} & x_{2-M} & \cdots & x_{N-M} \end{bmatrix} \begin{bmatrix} x_0 & x_{-1} & \cdots & x_{1-M} \\ x_1 & x_0 & \cdots & x_{2-M} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-M} \end{bmatrix} = \sum_{n=0}^{N-1} \boldsymbol{X}_n \boldsymbol{X}_n^T = \underbrace{\sum_{n=0}^{N-1} \begin{bmatrix} x_n^2 & x_n x_{n-1} & \cdots & x_n x_{n-M+1} \\ x_{n-1}x_n & x_{n-1}^2 & \cdots & x_{n-1}x_{n-M+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n-M+1}x_n & x_{n-M+1}x_{n-1} & \cdots & x_{n-M+1}^2 \end{bmatrix}}$$

ahol $\boldsymbol{X}_n^T = [x_n \quad x_{n-1} \quad \cdots \quad x_{n-M+1}]$

$\widehat{\boldsymbol{R}}_{xx}$ $\boxed{\text{„autokorreláció” becslés.}}$ Normálás!

$$\boldsymbol{U}^T\boldsymbol{z} = \sum_{n=0}^{N-1} \begin{bmatrix} x_n z_n \\ x_{n-1} z_n \\ \vdots \\ x_{n-M+1} z_n \end{bmatrix}$$

$\widehat{\boldsymbol{R}}_{xz}(k)$ $\boxed{\text{„keresztkorreláció” becslés.}}$ Normálás!

Ezzel: $\boxed{\hat{\boldsymbol{a}} = \widehat{\boldsymbol{R}}_{xx}^{-1}\widehat{\boldsymbol{R}}_{xz}}$

5. Példa: Lineáris modell ismert ***s*** jelkomponens esetén: $\boxed{\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{s} + \boldsymbol{w}}$ $\boldsymbol{z}' = \boldsymbol{z} - \boldsymbol{s}$ bevezetésével: $\boxed{\hat{\boldsymbol{a}} = [\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{U}^T(\boldsymbol{z} - \boldsymbol{s})}$

Fehér zaj esetén: $\boldsymbol{C}_{\hat{\boldsymbol{a}}} = \sigma^2[\boldsymbol{U}^T\boldsymbol{U}]^{-1}$.

A egy ismeretlen konstans, r pedig egy ismert konstans: $\boxed{x_n = A + r^n + w_n}$

$$\boxed{\boldsymbol{z} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} A + \begin{bmatrix} 1 \\ r \\ \vdots \\ r^{N-1} \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix},}$$

$$\boxed{\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1}(x_n - r^n)} \qquad \boxed{var\{\hat{A}\} = \frac{\sigma^2}{N}}$$

6. Példa: Lineáris modell színes Gauss zaj esetén: $\boxed{\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{w}, \ \ \boldsymbol{w} \sim \mathcal{N}(0, \boldsymbol{C}),}$

C pozitív definit, ezért létezik olyan invertálható ***D*** mátrix, amellyel $\boldsymbol{C}^{-1} = \boldsymbol{D}^T\boldsymbol{D}$.

$$E\left[(\boldsymbol{D}\boldsymbol{w})\left((\boldsymbol{D}\boldsymbol{w})\right)^T\right] = E[\boldsymbol{D}\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{D}^T] = \boldsymbol{D}\boldsymbol{C}\boldsymbol{D}^T = \boldsymbol{D}\boldsymbol{D}^{-1}(\boldsymbol{D}^T)^{-1}\boldsymbol{D}^T = \boldsymbol{I}$$

felhasználásával bevezethetünk egy transzformációt:

$$\boxed{\boldsymbol{z}' = \boldsymbol{D}\boldsymbol{z} = \boldsymbol{D}\boldsymbol{U}\boldsymbol{a} + \boldsymbol{D}\boldsymbol{w} = \boldsymbol{U}'\boldsymbol{a} + \boldsymbol{w}', \quad \boldsymbol{w}' = \boldsymbol{D}\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}).}$$

$\boxed{\boldsymbol{z}' = \boldsymbol{D}\boldsymbol{z} = \boldsymbol{D}\boldsymbol{U}\boldsymbol{a} + \boldsymbol{D}\boldsymbol{w} = \boldsymbol{U}'\boldsymbol{a} + \boldsymbol{w}', \quad \boldsymbol{w}' = \boldsymbol{D}\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}).}$ Ebből az ismeretlen paramétervektor kifejezhető:

$\boxed{\hat{\boldsymbol{a}} = [\boldsymbol{U}'^T\boldsymbol{U}']^{-1}\boldsymbol{U}'^T\boldsymbol{z}' = [\boldsymbol{U}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{z} = [\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{z},}$ $\boxed{\boldsymbol{C}_{\hat{\boldsymbol{a}}} = [\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U}]^{-1}.}$

Megjegyzés: A házi feladatban színes zajt kell generálni. Ha $\boldsymbol{w}$ fehér zaj, egységnyi szórással, akkor transzformáljuk: $\boldsymbol{B}$-vel $E[(\boldsymbol{B}\boldsymbol{w})((\boldsymbol{B}\boldsymbol{w}))^T] = E[\boldsymbol{B}\boldsymbol{w}\boldsymbol{w}^T\boldsymbol{B}^T] = \boldsymbol{B}\boldsymbol{B}^T = \boldsymbol{C}$ Egy lehetséges megoldás: **Cholesky felbontás**. Vagy: **mvnrnd( )** multivariate normal random numbers

3.3.4. A legjobb lineáris torzítatlan becslő (Best Linear Unbiased Estimator, BLUE):

- Az MVU becslő nem mindig létezik vagy lehetetlenség megtalálni.
- Az adatok valószínűség sűrűségfüggvénye nem ismert.

A **BLUE** becslő egy **szuboptimális** becslő, amelyik:

- az adatok **lineáris** függvénye: $\boxed{\hat{\boldsymbol{a}} = \boldsymbol{H}\boldsymbol{z};}$
- csak **torzítatlan** lehet: $E(\hat{\boldsymbol{a}}) = \boldsymbol{H}E(\boldsymbol{z}) = \boldsymbol{a};$
- **minimalizálja** a becslő **varianciáját**;
- csak az adatok **átlagára** és **varianciájára** van szükség.

A BLUE meghatározása (skalár esetben):

A $z_k,\ k = 0,1,\dots,N-1$ adatokhoz az $\boxed{\hat{a} = \sum_{k=0}^{N-1} h_k z_k = \boldsymbol{h}^T\boldsymbol{z}}$ lineáris becslőt rendeljük, ahol $\boldsymbol{h} = \{h_0, h_1, \dots, h_{N-1}\}^T$.

A becslőtől elvárjuk a torzítatlanságot: $\boxed{E(\hat{a}) = \sum_{k=0}^{N-1} h_k E(z_k) = a}$

Minimalizáljuk a varianciát: $var(\hat{a}) = E\left\{\left(\hat{a} - E(\hat{a})\right)^2\right\} = E\left\{\left(\boldsymbol{h}^T\boldsymbol{z} - \boldsymbol{h}^T E(\boldsymbol{z})\right)^2\right\} =$
$= E\{\boldsymbol{h}^T[\boldsymbol{z} - E(\boldsymbol{z})][\boldsymbol{z} - E(\boldsymbol{z})]^T\boldsymbol{h}\} = \boldsymbol{h}^T\boldsymbol{C}\boldsymbol{h}.$

1. Példa: Ismert jel amplitúdójának detektálása zajban: $\boxed{z_k = a s_k + w_k}$

A lineáris becslő: $\hat{a} = \sum_{k=0}^{N-1} h_k z_k = \boldsymbol{h}^T\boldsymbol{z}$. Ez torzítatlan: $E(\hat{a}) = \boldsymbol{h}^T E(\boldsymbol{z}) = \boldsymbol{h}^T\boldsymbol{s}a = a,$ ahol ehhez $\boldsymbol{h}^T\boldsymbol{s} = 1$ kell legyen, és $\boldsymbol{s} = \{s_0, s_1, \dots, s_{N-1}\}^T$. $\boxed{\text{Minimalizálandó } \boldsymbol{h}^T\boldsymbol{C}\boldsymbol{h} \text{ a } \boldsymbol{h}^T\boldsymbol{s} = 1 \text{ feltétel betartása mellett!}}$

Ehhez a Lagrange multiplikátoros technikát használjuk. Keressük a következő kifejezés szélsőértékét: $\boxed{J = \boldsymbol{h}^T\boldsymbol{C}\boldsymbol{h} + \lambda(\boldsymbol{h}^T\boldsymbol{s} - 1)}$

$\frac{\partial J}{\partial \boldsymbol{h}} = 2\boldsymbol{C}\boldsymbol{h} + \lambda\boldsymbol{s} = \boldsymbol{0}$, ebből $\boldsymbol{h} = -\frac{\lambda}{2}\boldsymbol{C}^{-1}\boldsymbol{s}$, ezzel $\boldsymbol{h}^T\boldsymbol{s} = 1 = -\frac{\lambda}{2}\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}$, $-\frac{\lambda}{2} = \frac{1}{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}}$, amit visszahelyettesítve $\boldsymbol{h}$ kifejezésébe:

$\boldsymbol{h} = \frac{\boldsymbol{C}^{-1}\boldsymbol{s}}{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}}$, és végül $\hat{a} = \boldsymbol{h}^T\boldsymbol{z}$ felhasználásával az optimális megoldás: $\boxed{\hat{a} = \frac{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{z}}{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}},}$ $\boxed{var(\hat{a}) = \frac{1}{\boldsymbol{s}^T\boldsymbol{C}^{-1}\boldsymbol{s}}}$

A BLUE meghatározása (vektoros esetben):

$\boldsymbol{z} = \boldsymbol{Ua} + \boldsymbol{w},$ ahol w nulla várható értékű zaj, melynek kovariancia mátrixa $\boldsymbol{C},$ sűrűségfüggvénye tetszőleges,

akkor az $\boldsymbol{a}$ paraméter **BLUE** becslője és a becslés kovariancia mátrixa: $\boxed{\hat{\boldsymbol{a}} = (\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{z},}$ $\boxed{\boldsymbol{C}_{\hat{\boldsymbol{a}}} = (\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}}$

Megjegyzés: Ha a zaj Gauss eloszlású, akkor a BLUE egy MVU becslő.

2. Példa: DC szint zajban: $\boxed{z_k = A + w_k,}$ w_k sűrűségfüggvényét nem ismerjük, csak a varianciáját: $var(w_k) = \sigma_k^2.$

$U = [1,1,\ldots,1]^T,$
$a = A.$
A kovariancia mátrix: $\boldsymbol{C} = \begin{bmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & \sigma_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{N-1}^2 \end{bmatrix} \Rightarrow \boldsymbol{C}^{-1} = \begin{bmatrix} \frac{1}{\sigma_0^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_1^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_{N-1}^2} \end{bmatrix}$

$$\boxed{\hat{a} = (\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{z} = \left(\sum_{k=0}^{N-1}\frac{1}{\sigma_k^2}\right)^{-1}\sum_{k=0}^{N-1}\frac{z_k}{\sigma_k^2}}$$

$$\boxed{\boldsymbol{C}_{\hat{a}} = (\boldsymbol{U}^T\boldsymbol{C}^{-1}\boldsymbol{U})^{-1} = \left(\sum_{k=0}^{N-1}\frac{1}{\sigma_k^2}\right)^{-1}}$$

3.3.5. Maximum Likelihood (ML) becslők

Probléma, hogy **MVU becslők** esetlegesen léteznek vagy nem találhatók meg. A BLUE becslő **lineáris modellekre** szorítkozik.

A **Maximum Likelihood (ML) becslők** determinisztikus modellel leírt paraméterek esetén is használhatók. Tulajdonságaik:
- mindig használhatók, ha a valószínűségsűrűség függvény (**csatornakarakterisztika**) ismert;
- nagy adatméret esetén **optimálisak** (aszimptotikusan hatásosak, elérik a CRLB értéket);
- számítástechnikailag komplexek, numerikus módszerek használatát igénylik.

Likelihood függvény: az $f(\boldsymbol{z};\boldsymbol{a})$ sűrűségfüggvény, ahol $\boldsymbol{a}$ változó (nem paraméter)
ML becslő: $\boldsymbol{a}$ azon értéke, amelyhez a likelihood függvény maximuma tartozik.
Eljárása: képezzük a $ln f(\boldsymbol{z};\boldsymbol{a})$ log-likelihood függvényt, és deriváljuk $\boldsymbol{a}$ szerint, majd ezt nullával egyenlővé téve megoldjuk $\boldsymbol{a}$-ra.

$\boxed{\text{Eredményül az } \hat{\boldsymbol{a}}_{ML} \text{ becslőt kapjuk.}}$

1. Példa: Additív, Gauss eloszlású fehér zajjal terhelt DC szint mérése ismeretlen zaj variancia mellett: $z_n = A + w_n.$

Tegyük fel, hogy $A > 0$ és $\sigma^2 = A$. A valószínűség sűrűségfüggvény:

CRLB?

Létezik MVU becslő?

$$f(z;A) = \frac{1}{(2\pi A)^{\frac{N}{2}}} exp\left[-\frac{1}{2A}\sum_{n=0}^{N-1}(z_n - A)^2\right].$$

$$\frac{\partial \ln f(z;A)}{\partial A} = -\frac{N}{2A} + \frac{1}{A}\sum_{n=0}^{N-1}(z_n - A) + \frac{1}{2A^2}\sum_{n=0}^{N-1}(z_n - A)^2$$

Az ML becslőt a derivált egyenlő nulla feltételből kapjuk: $\hat{A}_{ML} = -\frac{1}{2} + \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} z_n^2 + \frac{1}{4}}$

3.3.6. Legkisebb négyzetes hibájú (LS) becslők

Nincs előzetes ismeretünk sem a mérendő paraméterről, sem a csatorna karakterisztikáról (a zajról).

Minden megfigyelést a következőképpen képzelünk el: $z_k = s_k(\boldsymbol{a}) + e_k,$ ahol e_k a mérési és a modellezési hiba.

A cél az LS költség minimalizálása: $J(\boldsymbol{a}) = \sum_{k=0}^{N-1}\left(z_k - s_k(\boldsymbol{a})\right)^2.$ Nincs valószínűségi feltételezés csak egy determinisztikus jelmodell.

Ha a megfigyelési egyenlet lineáris: $\boldsymbol{z} = \boldsymbol{U}\boldsymbol{a} + \boldsymbol{e}$, akkor explicit megoldás adható.

Feltételezzük, hogy az $\boldsymbol{a}$ paraméter $\hat{\boldsymbol{a}}$ értéket vesz fel, és felállítjuk a megfigyelés modelljét: $\boldsymbol{U}\hat{\boldsymbol{a}}$.

$$J(\boldsymbol{a},\hat{\boldsymbol{a}}) = (\boldsymbol{z} - \boldsymbol{U}\hat{\boldsymbol{a}})^T(\boldsymbol{z} - \boldsymbol{U}\hat{\boldsymbol{a}}) = \boldsymbol{z}^T\boldsymbol{z} - \boldsymbol{z}^T\boldsymbol{U}\hat{\boldsymbol{a}} - \hat{\boldsymbol{a}}^T\boldsymbol{U}^T\boldsymbol{z} + \hat{\boldsymbol{a}}^T\boldsymbol{U}^T\boldsymbol{U}\hat{\boldsymbol{a}} = \boldsymbol{z}^T\boldsymbol{z} - 2\hat{\boldsymbol{a}}^T\boldsymbol{U}^T\boldsymbol{z} + \hat{\boldsymbol{a}}^T\boldsymbol{U}^T\boldsymbol{U}\hat{\boldsymbol{a}}$$

melynek szélsőértékét (minimumát) keressük:

$$\left.\frac{\partial J(\boldsymbol{a},\hat{\boldsymbol{a}})}{\partial \hat{\boldsymbol{a}}}\right|_{\hat{\boldsymbol{a}}=\hat{\boldsymbol{a}}_{LS}} = -2\boldsymbol{U}^T\boldsymbol{z} + 2\boldsymbol{U}^T\boldsymbol{U}\hat{\boldsymbol{a}} = 0 \quad \rightarrow \quad \hat{\boldsymbol{a}}_{LS} = [\boldsymbol{U}^T\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{z}$$

A minimális LS költség: $J(\boldsymbol{a},\hat{\boldsymbol{a}}_{LS}) = \boldsymbol{z}^T(\boldsymbol{z} - \boldsymbol{U}\hat{\boldsymbol{a}}_{LS})$

Megjegyzések: Súlyozott négyzetes kritériumot is használhatunk, ha bevezetünk egy $\boldsymbol{Q}$ négyzetes súlyozó mátrixot, amivel:

$$J(\boldsymbol{a},\hat{\boldsymbol{a}}) = (\boldsymbol{z} - \boldsymbol{U}\hat{\boldsymbol{a}})^T\boldsymbol{Q}(\boldsymbol{z} - \boldsymbol{U}\hat{\boldsymbol{a}}), \qquad \hat{\boldsymbol{a}}_{LS} = [\boldsymbol{U}^T\boldsymbol{Q}\boldsymbol{U}]^{-1}\boldsymbol{U}^T\boldsymbol{Q}\boldsymbol{z}.$$

Ha $\boldsymbol{Q} = \boldsymbol{\Sigma}_{nn}^{-1}$, akkor formálisan a Gauss-Markov becslőt kapjuk, tehát a GM becslő egy súlyozott legkisebb négyzetes hibájú becslő, ahol a súlyozást a megfigyelési zaj kovariancia mátrixának inverze adja.