Méréselmélet

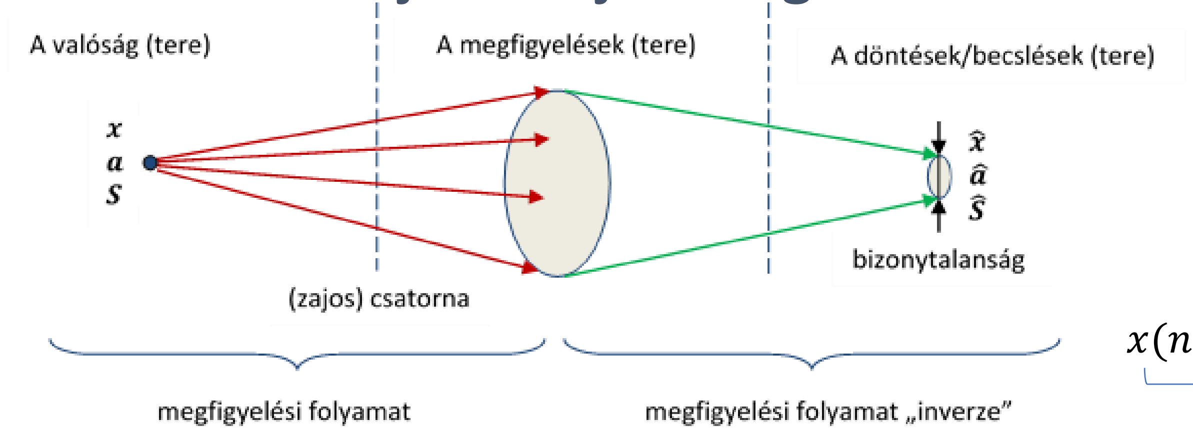
2. fejezet: A döntéselmélet alapjai

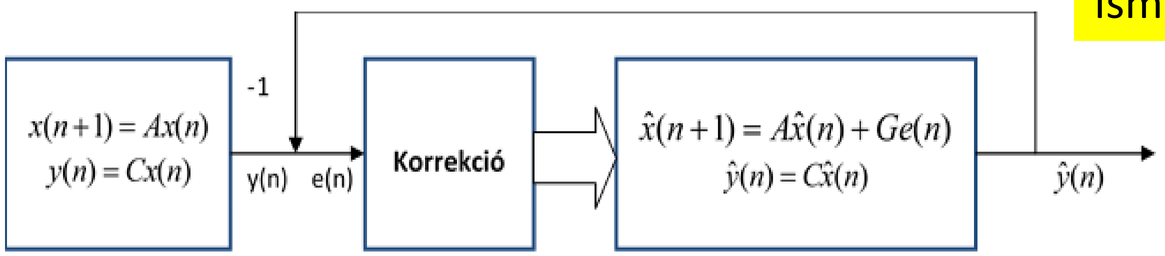
3. fejezet: A becsléselmélet alapjai

2021. február 17.

1.3. A mérési eljárás sajátosságai

Ismétlés





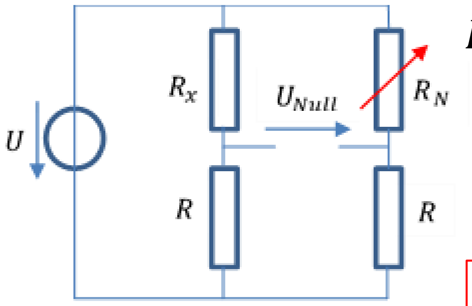
Célunk az $x(n)$ állapotvektor becslése.

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

$$\varepsilon(n+1) = F\varepsilon(n)$$

1. $F = A - GC = 0.$ Ebben az esetben $G = AC^{-1}$.
2. $F^N = (A - GC)^N = 0.$ $x(N) - \hat{x}(N) = (A - GC)^N(x(0) - \hat{x}(0)) = 0$ **N lépéses konvergencia, $\forall$ sajátérték nulla.**
3. Ha $F^N = (A - GC)^N \neq 0$, **a hiba exponenciális jelleggel csökken.**

„Majdnem minden rendszer megfigyelő.”



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

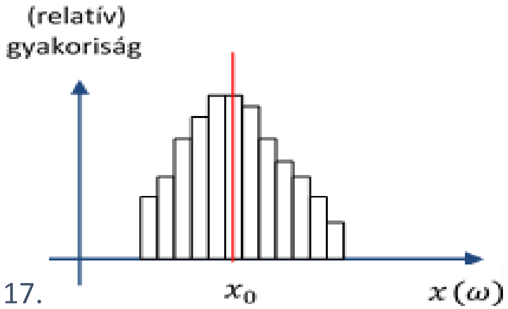
valós idejű **kiszámíthatóság!**

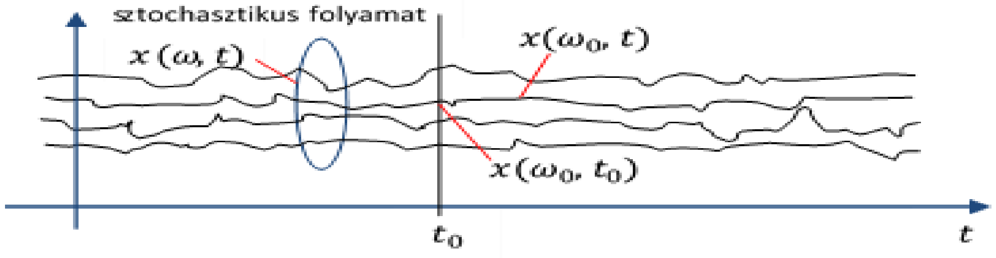
Megfigyelés zajos csatorna esetén: $y(n) = Cx(n)$+**zaj** $\rightarrow$ véletlen $\hat{x}(n+1)$, $\varepsilon(n+1)$.

$$E\{\textstyle\sum_{k=0}^{N-1} \varepsilon_k^2(n)\} = E[\boldsymbol{\varepsilon}^T(n)\boldsymbol{\varepsilon}(n)] = trace\, E[\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^T(n)] = trace P(n) \xrightarrow[n\to\infty]{} min.$$

$\varepsilon(n+1) = F\varepsilon(n)$ helyett $E[\varepsilon(n+1)\varepsilon^T(n+1)] = FE[\varepsilon(n)\varepsilon^T(n)]F^T = FP(n)F^T$

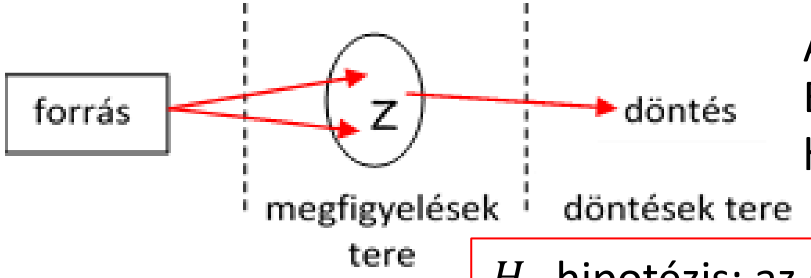
1.4. A zajos csatorna modellezése





2. A döntéselmélet alapjai

Példa: detektálás radarral. Bináris vagy kéthipotézises döntés: eldöntendő a jelenlét vagy a jelen nem lét kérdése.



A csatorna zajos, ugyanarról a jelenségről rendre eltérő értékű megfigyeléseket kapunk. El kell döntenünk, hogy – egy vagy több megfigyelésre alapozva – a két lehetséges hipotézisből melyiket fogadjuk el:

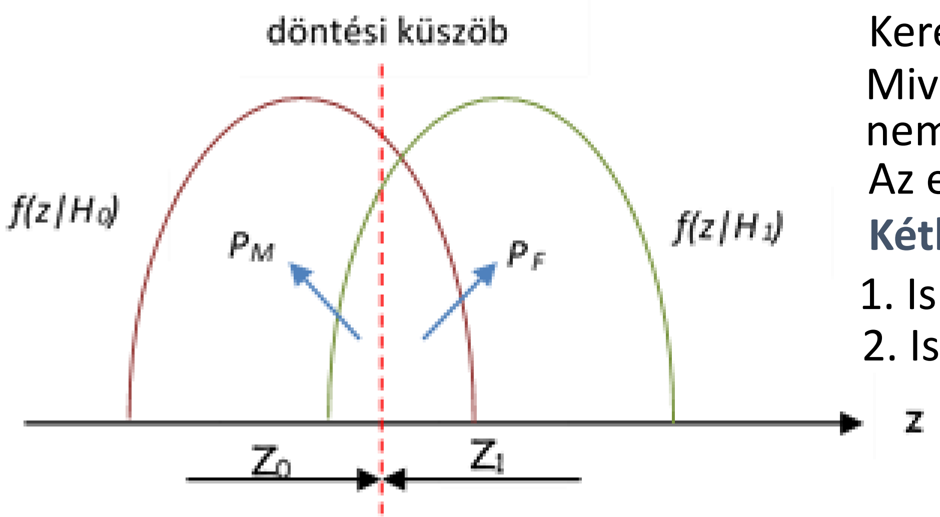
H_0 hipotézis: az (ellenséges) objektum nincs jelen.

H_1 hipotézis: az (ellenséges) objektum jelen van.

Lehetséges hibák: Elfogadjuk H_0-t, holott H_1 igaz. Ennek valószínűsége P_M (miss probability). Tévesztés valószínűsége

Elfogadjuk H_1-t, holott H_0 igaz. Ennek valószínűsége P_F (false alarm probability). Téves riasztás valószínűsége

A döntéshez előzetesen felvesszük a megfigyelések hisztogramját, és abból közelítőleg előállítjuk az $f(z|H_0)$ és $f(z|H_1)$ **feltételes sűrűségfüggvényeket**. Ezeket **csatorna- karakterisztikáknak** is nevezzük.



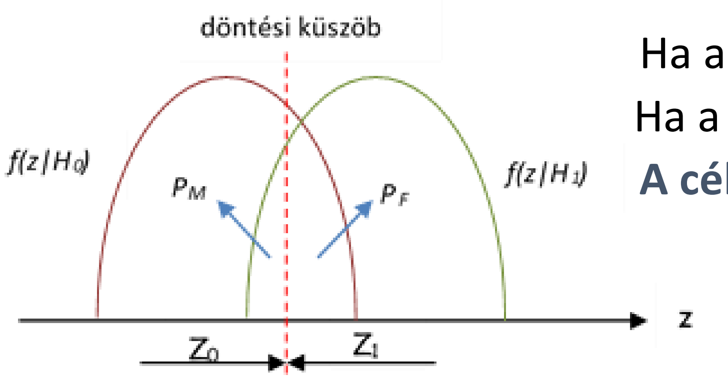
Keressük a döntési küszöböt valamilyen **optimum kritérium** szerint.
Mivel a sűrűségfüggvények átlapolódnak, ezért a döntési küszöb meghatározása nem triviális.
Az ehhez alkalmazható konkrét stratégia a rendelkezésre álló **információ** függvénye.

Kéthipotézises Bayes döntés: Alkalmazási feltételek:

1. Ismerjük az ún. a priori valószínűségeket: $H_0 \rightarrow P_0$ és $H_1 \rightarrow P_1$.
2. Ismerjük a csatornakarakterisztikákat: $f(z|H_0)$ és $f(z|H_1)$.

Definiáljuk a költségeket:
C_{ij} annak a költsége, hogy az i-edik hipotézist fogadtuk el, holott a j-edik igaz.

Értelmezzük a bekövetkezési valószínűségeket:
$P(H_i|H_j)$, ahol az i index a feltételezett hipotézis, a j index pedig a bekövetkezett kimenetel azonosítója. (i és j most 0 vagy 1.)



Ha a küszöb alatti (Z_0) tartományba esik a mért érték, akkor H_0-t fogadjuk el.
Ha a küszöb feletti (Z_1) tartományba esik a mért érték, akkor H_1-t fogadjuk el.
A cél: a küszöb olyan beállítása, amely az átlagos **kockázatot** (risk)/**költséget** (cost) minimalizálja:

$$R = \underbrace{C_{00}P_0P(H_0|H_0)}_{\text{Annak költsége, hogy } H_0 \text{ következett be, és azt is fogadtuk el.}} + \underbrace{C_{10}P_0P(H_1|H_0)}_{\text{Annak költsége, hogy } H_0 \text{ következett be, és } H_1\text{-t fogadtuk el.}} + \underbrace{C_{01}P_1P(H_0|H_1)}_{\text{Annak költsége, hogy } H_1 \text{ következett be, és } H_0\text{-t fogadtuk el.}} + \underbrace{C_{11}P_1P(H_1|H_1)}_{\text{Annak költsége, hogy } H_1 \text{ következett be, és azt is fogadtuk el.}}$$

A kifejezés minimumát a döntési küszöb értékének alkalmas megválasztásával érjük el.

A bekövetkezési valószínűségek számítása:

$$P(H_i|H_j) = \int_{Z_i} f(z|H_j)dz,$$

$$R = C_{00}P_0\int_{Z_0} f(z|H_0)dz + C_{10}P_0\int_{Z_1} f(z|H_0)dz + C_{01}P_1\int_{Z_0} f(z|H_1)dz + C_{11}P_1\int_{Z_1} f(z|H_1)dz.$$

Mivel $\int_{Z_1} f(z|H_0)dz = 1 - \int_{Z_0} f(z|H_0)$, és $\int_{Z_1} f(z|H_1)dz = 1 - \int_{Z_0} f(z|H_1)dz.$

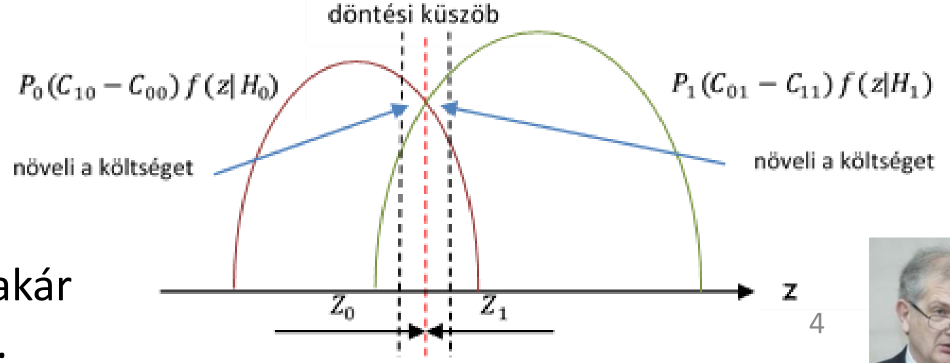
$$R = C_{10}P_0 + C_{11}P_1 + \underbrace{P_0(C_{00} - C_{10})\int_{Z_0} f(z|H_0)dz}_{<0} + \underbrace{P_1(C_{01} - C_{11})\int_{Z_0} f(z|H_1)dz}_{>0}.$$

Tegyük fel, hogy $C_{10} > C_{00}$, és $C_{01} > C_{11}$

Tekintsük a döntési küszöb helyét az összefüggésbeli egyváltozós integrál függvény független változójának!

Az átlagos kockázat minimuma a z változó azon értékénél (a keresett döntési küszöbértéknél) van, amelyre

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



A kiadódó döntési küszöbértéktől akár jobbra, akár balra eltérve az átlagos kockázat növekedni fog.

valószerűség

Ismétlés

$$\boxed{P_0(C_{10}-C_{00})f(z|H_0)=P_1(C_{01}-C_{11})f(z|H_1)} \Rightarrow \Lambda(z)=\frac{f(z|H_1)}{f(z|H_0)}\bigg|_{z_{küszöb}}=\frac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})}=\eta$$

ún. „likelihood” arány függvény

Ha az aktuálisan megfigyelt értéket behelyettesítjük a „likelihood” arány függvénybe, és ha $\Lambda(z)>\eta$, akkor a döntés H_1,

ha $\Lambda(z)<\eta$, akkor a döntés H_0.

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

Ez az ún. Bayes döntési szabály vagy likelihood arány teszt.

Megjegyzések:

1. Ha a költségeket úgy választjuk meg, hogy $\eta=\frac{P_0}{P_1}$ legyen, akkor a szabály $P_1 f(z|H_1)\underset{H_0}{\overset{H_1}{\gtrless}}P_0 f(z|H_0)$ alakban írható.

utólagos

Ez megegyezik a $P(H_1|z)\underset{H_0}{\overset{H_1}{\gtrless}}P(H_0|z)$ feltétellel, vagyis ilyenkor a döntést az *a posteriori* valószínűségek alapján hozhatjuk meg.

$P(H_0|z)$ annak a valószínűsége, hogy z-t mérve H_0, $P(H_1|z)$ pedig annak a valószínűsége, hogy z-t mérve H_1 következett be.

Ezt a speciális esetet maximum a posteriori (MAP) döntésnek nevezzük.

valószerűség

2. Szokás a $\lambda(z)=ln\,\Lambda\,(z)\underset{H_0}{\overset{H_1}{\gtrless}}ln\,\eta=\gamma$, ún. log-likelihood arányt is használni.

1. Példa:

Konstans jel detektálása zajos csatornán keresztül: jelen van-e a jel a megfigyelésekben vagy nincsen?

Tegyük fel, hogy a megfigyelések független, Gauss eloszlású valószínűségi változók nulla várható értékkel, és σ_n^2 varianciával.

A $\boldsymbol{H_0}$ **hipotézis** az, hogy a megfigyelés $z_k=n_k$, azaz **a jel nincsen jelen**, csak a zaj aktuális mintáját kapjuk.

A $\boldsymbol{H_1}$ **hipotézis** az, hogy a megfigyelés $z_k=a+n_k$, azaz **a jel jelen van**, a jel és a zaj aktuális mintájának összegét figyeltük meg.

$k=0,1,\dots,N-1$, azaz N mintát veszünk, és ezek együttes figyelembevételével döntünk. Ekkor z egy N dimenziós vektor!

A döntés során az együttes feltételes sűrűségfüggvényeket fogjuk használni, hiszen egyszerre veszünk figyelembe N mintát!

Minden mintáról tudjuk, hogy Gauss eloszlású. Mivel a minták függetlenek, az együttes feltételes sűrűség függvény az egyes feltételes sűrűségfüggvények szorzata lesz!

Egyetlen megfigyelés esetén a feltételes sűrűségfüggvények:

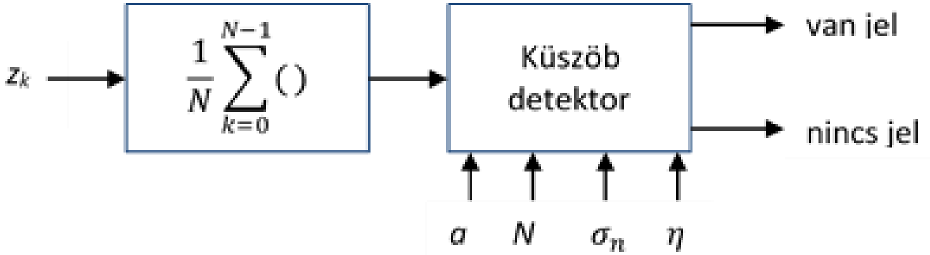
$$f(z|H_0) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{z_k^2}{2\sigma_n^2}},\ f(z|H_1) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(z_k-a)^2}{2\sigma_n^2}}.$$

$$\boxed{\Lambda(z) = \prod_{k=0}^{N-1} \frac{f(z_k|H_1)}{f(z_k|H_0)} = \prod_{k=0}^{N-1} \frac{e^{-\frac{(z_k-a)^2}{2\sigma_n^2}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta,}$$

$$\ln\Lambda(z) = \lambda(z) = -\sum_{k=0}^{N-1} \frac{(z_k - a)^2}{2\sigma_n^2} + \sum_{k=0}^{N-1} \frac{z_k^2}{2\sigma_n^2} = \frac{1}{2\sigma_n^2} \sum_{k=0}^{N-1} -(z_k^2 - 2az_k + a^2 - z_k^2) = \frac{a}{\sigma_n^2} \sum_{k=0}^{N-1} z_k - \frac{Na^2}{2\sigma_n^2} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta = \gamma$$

Átrendezve:

$$\boxed{\underbrace{\frac{1}{N} \sum_{k=0}^{N-1} z_k}_{\text{előfeldolgozás}} \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{\sigma_n^2}{Na} \ln\eta + \frac{a}{2}}_{\text{küszöbérték}}}$$



A döntőkészülék blokkvázlata

Ha $\eta = 1$, akkor $\ln\eta = 0$

ekkor

$$\boxed{\frac{1}{N} \sum_{k=0}^{N-1} z_k \underset{H_0}{\overset{H_1}{\gtrless}} \frac{a}{2}}$$

2. Példa: Változó amplitúdójú jel detektálása zajos csatornán keresztül: jelen van-e a jel a megfigyelésekben vagy nincsen?

Tegyük fel, hogy a megfigyelések független, Gauss eloszlású valószínűségi változók nulla várható értékkel, és σ_n^2 varianciával.

A H_0 **hipotézis** az, hogy a megfigyelés $z_k = n_k$, azaz **a jel nincsen jelen**, csak a zaj aktuális mintáját kapjuk.

A H_1 **hipotézis** az, hogy a megfigyelés $z_k = a_k + n_k$, azaz **a jel jelen van**, a jel és a zaj aktuális mintájának összegét figyeltük meg.

$k = 0, 1, \dots, N-1$, azaz N mintát veszünk, és ezek együttes figyelembevételével döntünk. Ekkor z egy N dimenziós vektor!

A döntés során az együttes feltételes sűrűségfüggvényeket fogjuk használni, hiszen egyszerre veszünk figyelembe N mintát! Minden mintáról tudjuk, hogy Gauss eloszlású.

$$f(z|H_0) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{z_k^2}{2\sigma_n^2}},\quad f(z|H_1) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(z_k-a_k)^2}{2\sigma_n^2}}.$$

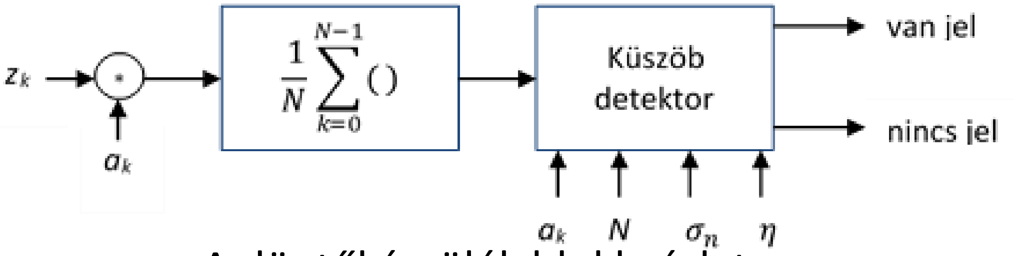
A megfigyelések függetlensége miatt a N megfigyelés együttes sűrűségfüggvénye az egyedi megfigyelések sűrűségfüggvényeinek szorzata.

$$\Lambda(z) = \prod_{k=0}^{N-1} \frac{f(z_k|H_1)}{f(z_k|H_0)} = \prod_{k=0}^{N-1} \frac{e^{-\frac{(z_k-a_k)^2}{2\sigma_n^2}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

$$\ln \Lambda(z) = \lambda(z) = -\sum_{k=0}^{N-1} \frac{(z_k-a_k)^2}{2\sigma_n^2} + \sum_{k=0}^{N-1} \frac{z_k^2}{2\sigma_n^2} = \frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} z_k a_k - \frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1} a_k^2 \underset{H_0}{\overset{H_1}{\gtrless}} \ln\eta = \gamma$$

Átrendezve: $$\boxed{\underbrace{\frac{1}{N}\sum_{k=0}^{N-1} a_k z_k}_{\text{előfeldolgozás}} \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{\sigma_n^2}{N}\ln\eta + \frac{1}{2N}\sum_{k=0}^{N-1} a_k^2}_{\text{küszöbérték}}}$$

„illesztett szűrő”



A döntőkészülék blokkvázlata

Ha a minták egyformák, akkor az előző eredményt kapjuk.

3. példa: Véletlen amplitúdójú jel detektálása zajos csatornán keresztül: jelen van-e a jel a megfigyelésekben vagy nincsen?

Tegyük fel, hogy a jel és a zaj egyaránt időben diszkrét, stacionárius sztochasztikus fehér zaj folyamatok, Gauss eloszlással, nulla várható értékkel, és σ_a^2, ill. σ_n^2 varianciával.

A $\boldsymbol{H_0}$ **hipotézis** az, hogy a megfigyelés $z_k = n_k$, azaz **a jel nincsen jelen**, csak a zaj aktuális mintáját kapjuk.

A $\boldsymbol{H_1}$ **hipotézis** az, hogy a megfigyelés $z_k = a_k + n_k$, azaz **a jel jelen van**, a jel és a zaj aktuális mintájának összegét figyeltük meg.

$k = 0,1,\dots, N-1$, azaz összesen N mintát figyelünk meg egyszerre.

A döntés során az együttes feltételes sűrűségfüggvényeket fogjuk használni.

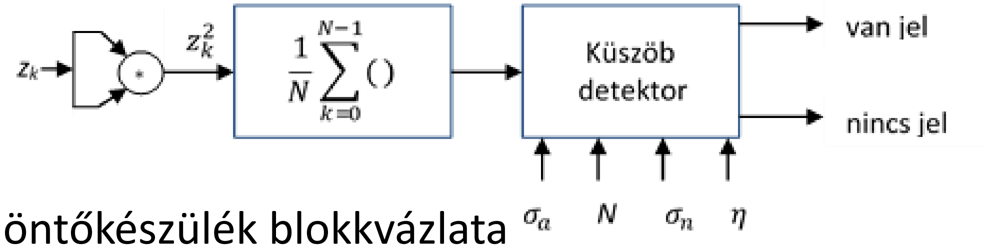
$$f(z|H_0)=\frac{1}{\sqrt{2\pi}\sigma_n}e^{-\frac{z_k^2}{2\sigma_n^2}},\quad f(z|H_1)=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_a^2+\sigma_n^2}}e^{-\frac{z_k^2}{2(\sigma_a^2+\sigma_n^2)}}.$$

$$\boxed{\Lambda(z)=\prod_{k=0}^{N-1}\frac{f(z_k|H_1)}{f(z_k|H_0)}=\prod_{k=0}^{N-1}\left[\frac{\dfrac{\sigma_n}{\sqrt{\sigma_a^2+\sigma_n^2}}e^{-\frac{z_k^2}{2(\sigma_a^2+\sigma_n^2)}}}{e^{-\frac{z_k^2}{2\sigma_n^2}}}\right]\underset{H_0}{\overset{H_1}{\gtrless}}\eta.}$$

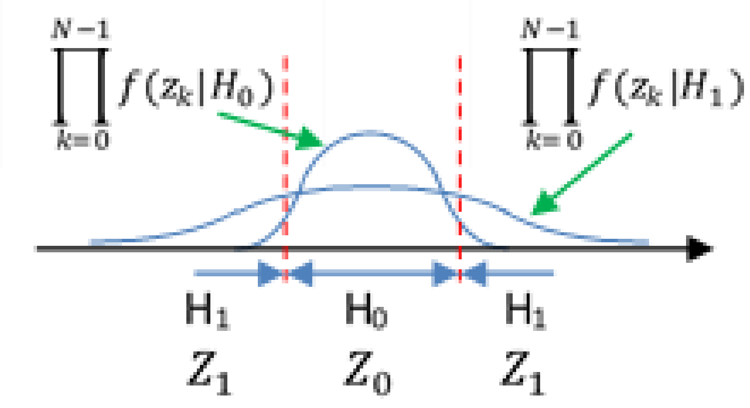
$$\ln\Lambda(z)=\frac{N}{2}\ln\frac{\sigma_n^2}{\sigma_a^2+\sigma_n^2}-\sum_{k=0}^{N-1}\frac{z_k^2}{2(\sigma_a^2+\sigma_n^2)}+\sum_{k=0}^{N-1}\frac{z_k^2}{2\sigma_n^2}=\frac{\sigma_a^2}{2\sigma_n^2(\sigma_a^2+\sigma_n^2)}\sum_{k=0}^{N-1}z_k^2+\frac{N}{2}\ln\frac{\sigma_n^2}{\sigma_a^2+\sigma_n^2}\underset{H_0}{\overset{H_1}{\gtrless}}\ln\eta=\gamma$$

Átrendezve:

$$\boxed{\underbrace{\frac{1}{N}\sum_{k=0}^{N-1}z_k^2}_{\text{előfeldolgozás}}\underset{H_0}{\overset{H_1}{\gtrless}}\underbrace{\frac{2\sigma_n^2(\sigma_a^2+\sigma_n^2)}{\sigma_a^2}\left[\frac{1}{2}\ln\frac{\sigma_a^2+\sigma_n^2}{\sigma_n^2}+\frac{1}{N}\ln\eta\right]}_{\text{küszöbérték}}}$$

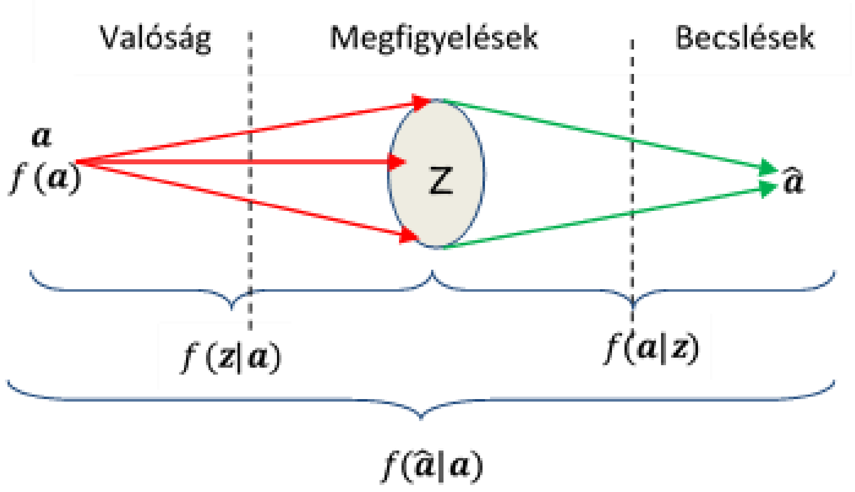


A döntőkészülék blokkvázlata



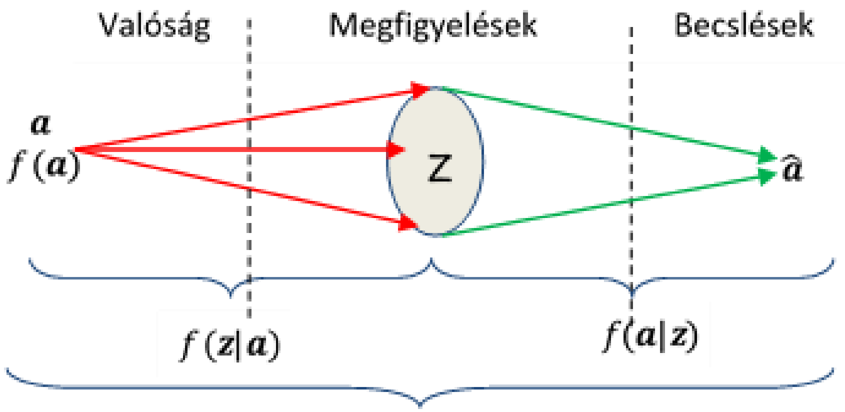
Döntési tartományok

3. A becsléselmélet alapjai



A cél az $\boldsymbol{a}$ paraméter(vektor) $\hat{\boldsymbol{a}}$ becslőjének meghatározása az ábrán látható előzetesen ismert (és az azokból származtatott) sűrűségfüggvények ismeretében.

3. A becsléselmélet alapjai



A cél az $\boldsymbol{a}$ paraméter(vektor) $\hat{\boldsymbol{a}}$ becslőjének meghatározása az ábrán látható előzetesen ismert (és az azokból származtatott) sűrűségfüggvények ismeretében.

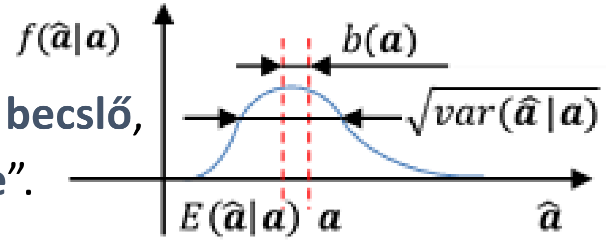
$\boldsymbol{a}$ véletlen (bizonytalan) érték $\rightarrow$ hisztogram $\rightarrow f(\boldsymbol{a})$

$\boldsymbol{a}$ fix érték $\rightarrow$ $\boldsymbol{z}$ véletlen (bizonytalan) érték $\rightarrow$ hisztogram $\rightarrow f(\boldsymbol{z}|\boldsymbol{a})$ csatorna karakterisztika

Mi a megfigyelt értéket/érték-halmazt $(\boldsymbol{z})$ ismerjük!

Vajon milyen értékű és gyakoriságú $\boldsymbol{a}$ eredményezhetett ilyen $\boldsymbol{z}$ értéket? $\rightarrow$ hisztogram $\rightarrow f(\boldsymbol{a}|\boldsymbol{z})$ csatorna karakterisztika „inverze”

Amit mi keresünk: $\hat{\boldsymbol{a}}$ **becslő**, és annak „**minősítése**”.



1. **Feltételes várhatóérték**: $E\{\hat{\boldsymbol{a}}|\boldsymbol{a}\} = \int_{-\infty}^{\infty} \hat{\boldsymbol{a}}\, f(\hat{\boldsymbol{a}}|\boldsymbol{a}) d\hat{\boldsymbol{a}}$
2. **Feltételes kovariancia mátrix**: $cov\{\hat{\boldsymbol{a}},\hat{\boldsymbol{a}}|\boldsymbol{a}\} = E\left\{\left(\hat{\boldsymbol{a}} - E(\hat{\boldsymbol{a}}|\boldsymbol{a})\right)\left(\hat{\boldsymbol{a}} - E(\hat{\boldsymbol{a}}|\boldsymbol{a})\right)^T |\boldsymbol{a}\right\}$
3. **Feltételes torzítás** (bias): $b(\boldsymbol{a}) = E\{\hat{\boldsymbol{a}}|a\} - \boldsymbol{a}$
4. **Átlagos négyzetes hiba** (Mean Square Error (MSE)) **mátrix**: $E\{(\hat{\boldsymbol{a}} - \boldsymbol{a})(\hat{\boldsymbol{a}} - \boldsymbol{a})^T|\boldsymbol{a}\} = cov\{\hat{\boldsymbol{a}},\hat{\boldsymbol{a}}|\boldsymbol{a}\} + b(\boldsymbol{a})b^T(\boldsymbol{a})$

Ha ismert az $f(\boldsymbol{a})$ valószínűség-sűrűségfüggvény, akkor definiálható:

5. **Feltétel nélküli várható érték:** $E(\hat{\boldsymbol{a}}) = E\{E\{\hat{\boldsymbol{a}}|\boldsymbol{a}\}\}$
6. **Feltétel nélküli kovariancia mátrix:** $cov\{\hat{\boldsymbol{a}},\hat{\boldsymbol{a}}\} = E\{cov\{\hat{\boldsymbol{a}},\hat{\boldsymbol{a}}|\boldsymbol{a}\}\}$

3.1. Bayes becslők

Az ismeretlen (mérendő) paramétert **valószínűségi változóként** modellezzük. Az egyszerűség kedvéért a paraméter legyen skalár.

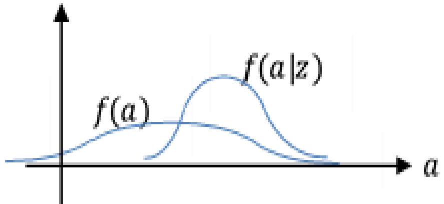
Feltételezzük, hogy ismert a megfigyelt paraméter **sűrűségfüggvénye:** $f(a)$, és **a csatorna karakterisztikája**: $f(z|a)$.

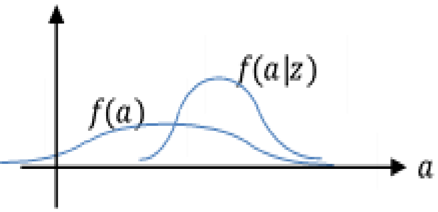
A **Bayes szabály** felhasználásával előállítható az $f(a|z)$ ún. *a posteriori* (utólagos) sűrűségfüggvény:

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

ahol

$$f(z) = \int_{-\infty}^{\infty} f(a)f(z|a)da$$





Ha a megfigyelések tartalmaznak a paraméterre vonatkozó információt, az *a posteriori* sűrűségfüggvény a konkrét paraméter-értékek egy szűkebb környezetére terjed ki.

Az ***a posteriori* sűrűségfüggvény** segítségével határozzuk meg az ismeretlen paraméter (vektor) **„legjobb" becslőjét.**
Ehhez értelmezünk kell a **„legjobb"** fogalmát alkalmas **költségfüggvényeken** keresztül:

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

A **Bayes költség** ennek a minimuma:

$$R_B = \min_{\hat{a}} R(\hat{a},a) = \min_{\hat{a}} E\{C(\hat{a},a)\}.$$

Itt $C(\hat{a},a)$ az ún. **költségfüggvény,** vagy **kockázatfüggvény**, vagy **kritériumfüggvény**, vagy **hibafüggvény,** vagy **illesztési függvény**.

I. **Négyzetes kritérium** (vektor paraméter megengedett):

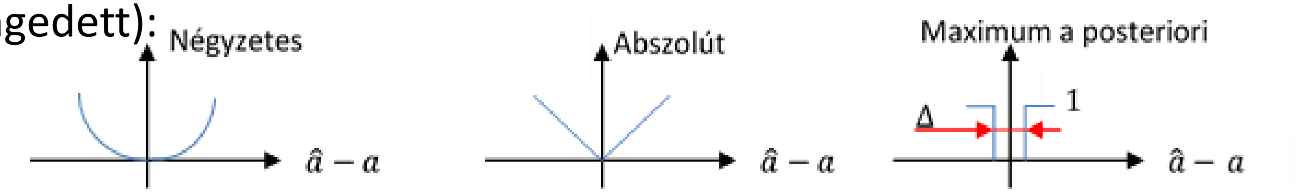
$$C(\hat{a},a) = \sum_{i=1}^{m} (\hat{a}_i - a_i)^2 = (\hat{a}-a)^T(\hat{a}-a)$$

II. **Abszolút kritérium** (vektor paraméter megengedett):

$$C(\hat{a},a) = \sum_{i=1}^{m} |\hat{a}_i - a_i|$$

III. **Maximum a posteriori kritérium** (vektor paraméter megengedett):

$$C(\hat{a},a) = \begin{cases} 0 & ha \quad |\hat{a}_i - a_i| \le \frac{\Delta}{2} \ \forall i - re \\ 1 & egyébként \end{cases}$$



3.1.1. Minimális átlagos négyzetes hibájú becslés

$$R(\hat{a},a) = E\{(\hat{a}-a)^T(\hat{a}-a)\} = \int_{-\infty}^{\infty}\left[\underbrace{\int_{-\infty}^{\infty}(\hat{a}-a)^T(\hat{a}-a) f(a|z)\,da}\right] f(z)\,dz = \int_{-\infty}^{\infty} g(\hat{a}) f(z)\,dz$$

felhasználjuk, hogy $f(a,z) = f(a|z)f(z)$.

$$g(\hat{a}) = \int_{-\infty}^{\infty} (\hat{a}-a)^T(\hat{a}-a)\, f(a|z)\,da$$

Csak ez függ $\hat{a}$-tól, ezért elég ennek a minimumát keresni!

$$g(\hat{a}) = \int_{-\infty}^{\infty} (\hat{a}-a)^T(\hat{a}-a)\, f(a|z)da$$ minimumának feltétele $$\boxed{\frac{\partial}{\partial \hat{a}} \int_{-\infty}^{\infty} (\hat{a}-a)^T (\hat{a}-a) f(a|z)da \Big|_{\hat{a} = \hat{a}_{MS}} = 0}$$

Mivel $\frac{\partial}{\partial \hat{a}} (\hat{a}-a)^T (\hat{a}-a) = 2(\hat{a}-a) \rightarrow \int_{-\infty}^{\infty} \hat{a}_{MS} f(a|z)da = \boxed{\hat{a}_{MS} = \int_{-\infty}^{\infty} a f(a|z)da}$

A legjobb becslés az *a posteriori* **várható érték**!

3.1.2. Minimális átlagos abszolút hibájú becslés

Skalár esetre bemutatva:

$$R(\hat{a},a) = E\{|\hat{a}-a|\} = \int_{-\infty}^{\infty} \underbrace{[\int_{-\infty}^{\hat{a}} (\hat{a}-a) f(a|z)da - \int_{\hat{a}}^{\infty} (\hat{a}-a) f(a|z)da]}_{g(\hat{a})} f(z)dz = \int_{-\infty}^{\infty} g(\hat{a}) f(z)dz$$

Itt is elég $g(\hat{a})$ minimumát megkeresni:

$$\frac{\partial}{\partial \hat{a}} \left[\int_{-\infty}^{\hat{a}} (\hat{a}-a) f(a|z)da - \int_{\hat{a}}^{\infty} (\hat{a}-a) f(a|z)da \right] \Big|_{\hat{a} = \hat{a}_{ABS}} = 0 \rightarrow \boxed{\int_{-\infty}^{\hat{a}_{ABS}} f(a|z)da = \int_{\hat{a}_{ABS}}^{\infty} f(a|z)da}$$

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

3.1.3. Maximum a posteriori (MAP) becslés

$$g(\hat{a}) = \int_{-\infty}^{\hat{a}-\frac{\Delta}{2}} f(a|z)da + \int_{\hat{a}+\frac{\Delta}{2}}^{\infty} f(a|z)da = 1 - \int_{\hat{a}-\frac{\Delta}{2}}^{\hat{a}+\frac{\Delta}{2}} f(a|z)da$$

Ha Δ kicsi, de $\Delta \neq 0$, akkor az optimum az *a posteriori* sűrűségfüggvény **maximumhelye**

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

Megjegyzések:

1. A Bayes becslések **mindig** az *a posteriori* sűrűségfüggvények alapján történnek.
2. Vektor paraméterre **általánosíthatók.**
3. Az MS becslés **lineáris** abban az alábbi értelemben:

Ha $b = Aa + c$, akkor $\hat{b}_{MS} = A\hat{a}_{MS} + c$, továbbá $E\{a+b|z\} = E\{a|z\} + E\{b|z\} = \hat{a}_{MS} + \hat{b}_{MS}$

3.1.4. Bayes becslő Gauss eloszlások esetén

Tegyük fel, hogy a keresett *a* paraméter és a megfigyelési zaj **Gauss eloszlásúak**!

A paraméter lehet **vektor**, és lehet, hogy a kiértékeléseknél **több megfigyelést** veszünk **egyidejűleg** figyelembe.

Adott $E\{\boldsymbol{a}\} = \boldsymbol{\mu_a}$, $cov\{\boldsymbol{a},\boldsymbol{a}\} = \boldsymbol{\Sigma_{aa}} \rightarrow$ Gauss eloszlás esetén „mindent" tudunk $\boldsymbol{a}$-ról!

$E\{\boldsymbol{n}\} = 0$, $cov\{\boldsymbol{n},\boldsymbol{n}\} = \boldsymbol{\Sigma_{nn}} \rightarrow$ Gauss eloszlás esetén „mindent" tudunk **a csatornáról**!

Ha "minden" Gauss eloszlású, akkor az a posteriori sűrűségfüggvény momentumai **explicit formában** megadhatók.

Tegyük fel, hogy az **additív zajjal** terhelt megfigyelés az alábbi összefüggéssel írható le: $\boxed{\boldsymbol{z} = \boldsymbol{Ua} + \boldsymbol{n}}$

A megfigyelés modellje lineáris!

Itt $\dim \boldsymbol{a} = p, \dim \boldsymbol{z} = q, \dim \boldsymbol{U} = q * p$. Az $\boldsymbol{U}$ az ún. **megfigyelési mátrix**.

Az *a posteriori* várható érték explicit formája ilyen körülmények között:

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

Az *a posteriori* kovariancia explicit formája ilyen körülmények között:

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

Megjegyzés: $\hat{\boldsymbol{a}}_{MS} = \hat{\boldsymbol{a}}_{ABS} = \hat{\boldsymbol{a}}_{MAP}$, mert az a Gauss *a posteriori* sűrűségfüggvény szimmetrikus.

1. Példa: Mérendő egy ellenállás értékét úgy, hogy ismert áram által ejtett feszültséget mérünk. *N* megfigyelést végzünk.

"$U = IR + zaj$" a modellünk. $\boxed{\text{A megfigyelt értékek: } z_k = a + n_k,\ k = 0, 1, \dots, N-1,}$

a az ismeretlen paraméter (ellenállás), n_k az additív zaj mintája.

Tegyük fel, hogy az **ellenállás értéke** és a **megfigyelési zaj** egyaránt **Gauss eloszlású** valószínűségi változóknak tekinthetők.

A zaj megfigyelési értékei **korrelálatlanok**. Tegyük fel, hogy **ismert** μ_a és $\sigma_a^2 \rightarrow$ az ismeretlen paraméter jellemzése.

a zaj várható értéke $\mu_n = 0$, $cov\{n_k, n_j\} = \sigma_n^2\delta_{kj}$, ahol $\delta_{kj} = \begin{cases} 1 \text{ ha } k = j \\ 0 \text{ ha } k \neq j \end{cases} \rightarrow$ a csatorna jellemzése.

Vektoros alakban: $\boxed{\boldsymbol{z} = \boldsymbol{Ua} + \boldsymbol{n}}$ $\boldsymbol{z}^T = [z_0, z_2, \dots, z_{N-1}]$, $\boldsymbol{n}^T = [n_0, n_2, \dots, n_{N-1}]$, $\boldsymbol{U} = [1 \quad 1 \quad \cdots \quad 1]^T$, $\boldsymbol{\Sigma_{nn}} = \sigma_n^2\boldsymbol{I}$

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

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

$$\hat{a}_{MS} = \mu_{a|z} = \mu_a + \left[\frac{N}{\sigma_n^2} + \frac{1}{\sigma_a^2}\right]^{-1}\left(\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} z_k - \frac{N}{\sigma_n^2}\mu_a\right) = \underbrace{\mu_a}_{\text{előzetes ismeret}} + \underbrace{\frac{N\frac{\sigma_a^2}{\sigma_n^2}}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}(\underbrace{\frac{1}{N}\sum_{k=0}^{N-1} z_k}_{\text{minták átlaga}} - \mu_a)} = \frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} z_k}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}$$

Ha $\sigma_a << \sigma_n$, akkor $\hat{a}_{MS} \cong \mu_a$,

ha $\sigma_a >> \sigma_n$ vagy $N \to \infty$, akkor $\hat{a}_{MS} \cong \frac{1}{N}\sum_{k=0}^{N-1} z_k$.

predikció + korrekció a méréssel szerzett információ alapján.

A becslési hiba **varianciája**: az *a posteriori* kovariancia:

$$cov\{a, a|\boldsymbol{z}\} = \boldsymbol{\Sigma}_{aa|\boldsymbol{z}} = [\boldsymbol{U}^T\boldsymbol{\Sigma}_{\boldsymbol{nn}}^{-1}\boldsymbol{U} + \boldsymbol{\Sigma}_{\boldsymbol{aa}}^{-1}]^{-1} = var\{\tilde{a}\} = \frac{\sigma_a^2}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}$$

$\tilde{a} = \hat{a} - a$

Ha $\sigma_a << \sigma_n$, akkor $var\{\tilde{a}\} \cong \sigma_a^2$,

ha $\sigma_a >> \sigma_n$ vagy $N \to \infty$, akkor $var\{\tilde{a}\} \cong \frac{1}{N}\sigma_n^2$.

A becslés feltételesen torzított: $b(a) = E\{\hat{a}_{MS}|a\} - a =$

$$= E\left\{\frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} z_k}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}\right\} - a = \frac{\mu_a - a}{1 + N\frac{\sigma_a^2}{\sigma_n^2}}.$$

Itt felhasználtuk, hogy $E\{z_k\} = a$.

2. Példa: Mérendő egy ismert jel ismeretlen amplitúdója.

$z_k = a s_k + n_k,\ k = 0, 1, \dots, N-1$. Az ismeretlen a amplitúdó és n_k Gauss eloszlású. $E\{a\} = \mu_a$, $var\{a\} = \sigma_a^2$,

$E\{n_k\} = 0$, $cov\{n_i, n_j\} = \sigma_n^2\delta_{ij}$, $cov\{a, n_i\} = 0$ $\forall i$ -re, $\forall j$ -re.

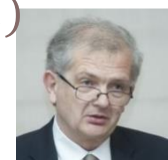
Most használjuk a **maximum a posteriori** (MAP) becslést!

$$\left.\frac{\partial f(a|\boldsymbol{z})}{\partial a}\right|_{a=\hat{a}_{MAP}} = 0 \rightarrow \left.\frac{\partial \ln f(\boldsymbol{z}|a)}{\partial a} + \frac{\partial \ln f(a)}{\partial a}\right|_{a=\hat{a}_{MAP}} = 0$$

Most $f(a) = \frac{1}{\sqrt{2\pi}\sigma_a} e^{-\frac{(a-\mu_a)^2}{2\sigma_a^2}} \rightarrow \frac{\partial \ln f(a)}{\partial a} = -\frac{a - \mu_a}{\sigma_a^2}$

$$f(a|z) = \frac{f(z|a)f(a)}{f(z)} \qquad \frac{\partial \ln f(z)}{\partial a} = 0$$

$$f(\boldsymbol{z}|a) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} e^{-\frac{1}{2\sigma_n^2}\sum_{k=0}^{N-1}(z_k - a s_k)^2} \rightarrow \frac{\partial \ln f(\boldsymbol{z}|a)}{\partial a} = \frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} s_k(z_k - a s_k)$$



$$\boxed{\left.\frac{\partial \ln f(\mathbf{z}|a)}{\partial a} + \frac{\partial \ln f(a)}{\partial a}\right|_{a=\hat{a}_{MAP}} = 0} \qquad \frac{\partial \ln f(a)}{\partial a} = -\frac{a-\mu_a}{\sigma_a^2} \qquad \frac{\partial \ln f(\mathbf{z}|a)}{\partial a} = \frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} s_k\left(z_k - a s_k\right)$$

$$\left.\frac{1}{\sigma_n^2}\sum_{k=0}^{N-1} s_k\left(z_k - a s_k\right) - \frac{a-\mu_a}{\sigma_a^2}\right|_{a=\hat{a}_{MAP}} = 0 \;\rightarrow\; \boxed{\hat{a}_{MAP} = \frac{\mu_a + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k z_k}{1 + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}} \qquad \boxed{var\{\tilde{a}\} = \frac{\sigma_a^2}{1 + \frac{\sigma_a^2}{\sigma_n^2}\sum_{k=0}^{N-1} s_k^2}}$$

Ha s_k=1, $\forall k$, akkor megkapjuk az előző példa eredményét!

Itt is azonosítható a döntéselméleti rész 2. feladatában említett „illesztett” szűrő. Természetesen $\hat{a}_{MAP} = \hat{a}_{MS}$.