KÁLMÁN-FÉLE SZŰRŐK

1. A probléma megfogalmazása.

E jegyzet témája az úgynevezett Kálmán-féle szűrők vizsgálata. A feladat a következő. Adott egy $x(0), x(1), \dots,$ több változós (együttesen) normális, más néven Gauss eloszlású véletlen vektorokból álló sorozat, amelyet valamely rendszer produkál. Úgy képzelhetjük el, hogy az n, $n = 0, 1, \dots,$ időpontban a rendszer generál egy $x(n)$ véletlen vektort. Ezeket az $x(n)$ véletlen vektorokat közvetlenül nem tudjuk megfigyelni, csak ezeknek valamilyen szűrőn keresztül előállított $y(0), y(1), \dots$ traszformáltjait. Az a feladat, hogy egy adott n időpontban adjunk minél jobb becslést az m idő múlva megjelenő $x(n + m)$, $m \geq 1$, vektor értékére az $y(0), y(1), \dots, y(n)$ véletlen vektorok értékeinek az ismeretében.

Pontosabban, részletesebben, megfogalmazva a feladat a következő: Tekintsük az alábbi rekurzív módon definiált sorozatokat:

$$x(n) = \Phi(n-1)x(n-1) + \varepsilon(n-1), \qquad n = 1, 2, \dots, \tag{1.1}$$

$$y(n) = H(n)x(n) + \eta(n), \qquad n = 0, 1, 2, \dots, \tag{1.2}$$

ahol $x(n)$ és $\varepsilon(n)$ p-dimenziós (oszlop) vektorok, $y(n)$ és $\eta(n)$ q-dimenziós véletlen (oszlop) vektorok, $\Phi(\cdot)$ és $H(\cdot)$ $p \times p$ illetve $q \times p$ méretű mártixok, $\varepsilon(1), \varepsilon(2), \dots$ és $\eta(0), \eta(1), \dots$ egymástól is független, független, nulla várható értékű normális eloszlású véletlen vektorokból álló sorozatok, amelyeknek a kovariancia mátrixai

$$E\varepsilon(n)\varepsilon(n)^* = Q(n), \quad n = 0, 1, \dots, \tag{1.3}$$

$$E\eta(n)\eta(n)^* = R(n), \quad n = 0, 1, \dots, \tag{1.4}$$

valamely az n indextől függő $p \times p$ és $q \times q$ méretű (pozitív szemidefinit) mátrixok. (Emlékeztetek arra, hogy $\varepsilon(n)$ és $\eta(n)$ oszlopvektorok. A $*$ jel egy vektor vagy mátrix konjugáltját fogja jelenteni a továbbiakban.) Ezenkívül adva van egy p-dimenziós, nulla várható értékű $x(0)$ normális eloszlású (kiinduló) véletlen vektor ismert $\Sigma(0)$ kovariancia mátrix-szal, amely független az $\varepsilon(0), \varepsilon(1), \dots$ és $\eta(0), \eta(1), \dots$ véletlen vektoroktól.

Ilyen módon az (1.1) képlet definiálja az $x(n)$, $n = 0, 1, \dots,$ sorozatot, az (1.2) formula pedig az $y(n)$, $n = 0, 1, \dots,$ sorozatot. Tegyük fel, hogy ismerjük az (1.1) és (1.2) formulában szereplő $\Phi(n)$ és $H(n)$ mátrixokat valamint az (1.3) és (1.4) képletekben szereplő $Q(n)$ és $R(n)$ kovariancia mátrixokat minden olyan n indexre, amelyekre ezek definiálva vannak. Ezenkívül ismerjük a kiinduló $x(0)$ vektor $Ex(0)x(0)^*$ kovariancia mátrixát is. Minket az $x(\cdot)$ véletlen vektorok értéke érdekel. Egy n időpontban meg akarjuk becsülni az $x(n + m)$ véletlen vektor értékét valamely $n + m$ időpontban, $m \geq 1$. Az n időpontig megfigyeléseket teszünk, de nem tudjuk magukat az $x(\cdot)$ véletlen vektorokat megfigyelni, csak ezek 'hibával terhelt' transzformáltjait, az $y(0), \dots, y(n)$ véletlen vektorokat. Feladatunk az $x(n + m)$ véletlen vektor optimális becslése az $y(0), \dots, y(n)$ véletlen vektorok ismeretében.

Az általános elmélet szerint ez a kérdés ekvivalens az $E(x(n + m)|y(0), \dots, y(n))$ feltételes várható érték kiszámításával. Mivel (együttesen) normális eloszlású vektorokkal dolgozunk ez a valószínűségszámítási probléma átfogalmazható a következő módon. Vezessük be a következő jelöléseket. Legyen $\mathcal{H}_n$ az $L_2(\Omega, \mathcal{A}, P)$ Hilbert térnek az $x(0)$, $\varepsilon(0), \dots, \varepsilon(n)$, $\eta(0), \dots, \eta(n)$ vektorok koordinátáinak lineáris kombinációiból álló altere, ahol $(\Omega, \mathcal{A}, P)$ az a valószínűségi mező, ahol az általunk tekintett valószínűségi változók 'élnek', a skalár szorzatot pedig a szokásos módon, a $(\xi, \eta) = E\xi\eta$ képlettel definiáljuk. Legyen $\mathcal{M}_n$ a $\mathcal{H}_n$ Hilbert térnek az $y(0), \dots, y(n)$ vektorok koordinátái által kifeszített altere. (A $\mathcal{H}_n$ altér tartalmazza az $x(0), \dots, x(n+1)$ és $y(0), \dots, y(n)$ vektorok koordinátáit, de általában nem tartalmazza az $y(n+1)$ vektor koordinátáit.) Ekkor az

$$E(x(n+m)|y(0), \dots, y(n))$$

feltételes várható érték megegyezik az $x(n+m)$ vektornak (pontosabban az $x(n+m)$ vektor minden koordinátájának) a vetületével az $\mathcal{M}_n$ altérre. Ez azt jelenti, hogy a minket érdeklő feladat ezen vetületek kiszámolása. Ennek a vetületnek a kiszámítására ismertetni fogunk egy rekurzív eljárást.

Ha összehasonlítjuk az e jegyzetben vizsgált feladatot e témakör egy másik előadásában tárgyalt predikció elméleti problémával, akkor láthatunk némi hasonlóságot, de különbséget is. Jelen esetben vektor értékű sorozat elemeinek a becslését vizsgáljuk, míg a predikció elméleti problémában skalár értékű valószínűségi változókét. Ennek a különbségnek azonban nem volt elvi oka. Az analóg probléma vizsgálata a predikció elméletben is fontos és sokat vizsgált kérdés. Ennek tárgyalása azonban több időt és energiát követelt volna.

Jelen problémában nem stacionárius sorozatot vizsgáltunk, és a a megfigyeléseket nem a $-\infty$ hanem a nulla időponttól kiindulva tekintettük. E különbség mögött már elvi okok is vannak. A predikció elméletben felhasználtuk azt, hogy a feladatot meg tudjuk fogalmazni a Fourier transzformáltak terében is, és a feladat ilyen átfogalmazása hasznosnak bizonyult. A jelen probléma vizsgálatában az eredeti térben maradtunk, ahol a statcionaritásnak nincs jelentősége, ezért érdemes volt a problémát általánosabban, az n paramétertől függő $\Phi(n)$, $H(n)$, $Q(n)$ és $R(n)$ mátrixok segítségével megfogalmazni. Ezt a különbséget a kétfajta vizsgálat között úgy is meg szokták fogalmazni az irodalomban, hogy e jegyzet vizsgálataiban a vizsgált Gauss folyamat állapotteres, míg a korábban tárgyalt vizsgálatokban fázisteres leírását adtuk.

Megjegyzem, hogy egyéb természetes kérdések is felmerülnek az e jegyzetben vizsgált problémával kapcsolatban, amelyeket itt nem tárgyalunk. Az egyik ilyen kérdés az, hogy amennyiben a feladatban tekintett Φ, H, Q és R mátrixok nem függnek az n paramétertől, akkor milyen feltételek mellett létezik az (1.1) és (1.2) feltételt kielégítő stacionárius Gauss sorozat, és hogyan kell megadni egy stacionárius sorozat kezdő eloszlását.

2. Az optimális becslés megadása.

Ebben a fejezetben ismertetjük azt az eredményt, amely megadja, hogy hogyan lehet az 1. fejezet problémáját megoldani, az ott definiált projekciót kiszámolni. Az eredmény megfogalmazása előtt bevezetek néhány jelölést.

Jelölje $\hat{x}(n+m|n)$ az (1.1) képletben definiált a $\mathcal{H}(n+m)$ Hilbert térben levő $x(n+m)$ véletlen vektor (pontosabban koordinátáinak) vetületét a $\mathcal{H}(n+m)$ Hilbert térnek szintén az első fejezetben bevezetett $\mathcal{M}_n$ alterébe. Használni fogjuk a következő, kissé pongyola, de az irodalomban elterjedt jelölést. Azt mondjuk, hogy egy k-dimenziós z véletlen vektor, $z = (z_1, \dots, z_k)^*$, benne van a $\mathcal{M}_n$ térben, jelölésben $z \in \mathcal{M}_n$, ha a z vektor minden z_j koordinátájára $z_j \in \mathcal{M}_n$, $1 \le j \le k$. Hasonlóan azt mondjuk, hogy z merőleges a $\mathcal{M}_n$ altérre, jelölésben $z \perp \mathcal{M}_n$ ha $z_j \perp \mathcal{M}_n$ a z vektor minden z_j koordinátájára, $1 \le j \le k$. Alább megfogalmazom e jegyzet fő eredményét.

2.1. Tétel. *Legyen $\hat{x}(n+m|n)$ az $x(n+m)$ vektor vetülete az $\mathcal{M}_n$ altérre. Ekkor*

$$\hat{x}(n+m|n) = \Phi(n+m-1)\hat{x}(n+m-1|n), \quad \text{ha } m > 1, \tag{2.1}$$

$$\hat{x}(n+1|n) = \Psi(n)\hat{x}(n|n-1) + K(n)y(n), \tag{2.2}$$

ahol $\hat{x}(0|-1) = 0$ az $n = 0$ esetben,

$$\Psi(n) = \Phi(n) - K(n)H(n), \tag{2.3}$$

és

$$K(n) = \Phi(n)\Sigma(n)H(n)^*\{H(n)\Sigma(n)H(n)^* + R(n)\}^{-1}. \tag{2.4}$$

A (2.4) formulában szereplő $\Sigma(n)$ mátrixot a következő rekurzív formulával számíthatjuk ki.

$$\Sigma(n+1) = \Phi(n)\Sigma(n)\Phi(n)^* + Q(n) - K(n)\{H(n)\Sigma(n)H(n)^* + R(n)\}K(n)^*, \tag{2.5}$$

és $\Sigma(0) = Ex(0)x(0)^*$.

1. Megjegyzés. A 2.1 tételben (a $K(n)$ mátrix (2.4) formulában megadott definiciójában) feltételeztük, hogy a $H(n)\Sigma(n)H(n)^* + R(n)$ mátrix invertálható. A 4. fejezetben elmagyarázom, hogy hogyan tudunk ettől a megszorítástól megszabadulni.

2. Megjegyzés. A 2.1 tétel megfogalmazásában szereplő $\Sigma(n)$ mátrixot a (3.7) formulában fogjuk definiálni. Az $x(n)$ véletlen vektor becslése az $n-1$ időpontban az $\hat{x}(n|n-1)$, és a becslés hibája az $x(n) - \hat{x}(n|n-1)$ vektor. Ez a hiba egy nulla várható értékű normális eloszlású vektor, és $\Sigma(n)$ a kovariancia mátrixa. A 2.1 tételben az $x(n+1)$ vektor $\hat{x}(n+1|n)$ becslését együtt számoljuk ki e becslés hibájának $\Sigma(n+1)$ kovariancia mátrixával. Az $(\hat{x}(n+1|n), \Sigma(n+1))$ pár kiszámításához elég ismernünk az előző lépésben kiszámolt $(\hat{x}(n|n-1), \Sigma(n))$ párt és az utolsó $y(n)$ megfigyelést. Sőt, a $\Sigma(n)$ mátrixokat a megfigyelés megkezdése előtt, az $y(n)$, $n = 0, 1, 2, \dots$ mintaelemek ismerete nélkül is kiszámolhatjuk. Ismerjük a rekurzió kiinduló $(\hat{x}(0|-1), \Sigma(0)) = (0, Ex(0)x(0)^*)$ párját.

A 2.1. tétel megválaszolja azt a kérdést, hogy hogyan adjuk meg az $x(n+m)$, $m \geq 1$, véletlen vektor optimális becslését az $y(1), \dots, y(n)$ véletlen vektorok ismeretében (ismerve a $\Phi(n)$, $H(n)$, $Q(n)$ és $R(n)$ mátrixokat is). Ez az $\hat{x}(n+m|n)$ vektor kiszámítását jelenti. Ha $m \geq 2$, akkor a (2.1) képlet segítségével az $\hat{x}(n+m|n)$ vektor kiszámítását vissza tudjuk vezetni az $\hat{x}(n+1|n)$ vektor kiszámítására. A $\hat{x}(n+1|n)$ vektort a (2.2), (2.3) és (2.4) formulák segítségével n szerinti rekurzióval számolhatjuk ki, feltéve, hogy ki tudjuk számolni a $\Sigma(n)$ mátrixokat. (Annak érdekében, hogy ezt a rekurziót alkalmazni tudjuk tudnunk kell azt is, hogy $\hat{x}(0|-1) = 0$, és ismernünk kell a $\Sigma(0) = Ex(0)x(0)^*$ kovariancia mátrixot.) Viszont a $\Sigma(n)$ mennyiséget ki tudjuk számítani a (2.5) formula segítségével szintén az n változó szerinti rekurzióval.

3. A 2.1. Tétel bizonyítása.

Bontsuk fel az $\mathcal{M}_n$ Hilbert teret a $\mathcal{M}_n = \mathcal{M}_{n-1} \oplus \mathcal{V}_n$ direkt összegre, ahol $\mathcal{V}_n$ az $\mathcal{M}_{n-1}$ altér ortogonális kiegészítője az $\mathcal{M}_n$ Hilbert térben. Ki fogjuk számolni az $y(n) \in \mathcal{M}_n$ és $\hat{x}(n+1|n) \in \mathcal{M}_n$ vektorok felbontását (pontosabban szólva e vektorok minden koordinátájának a felbontását) e vektoroknak az $\mathcal{M}_{n-1}$ és $\mathcal{V}_n$ (ortogonális) alterekre vett vetületeinek az összegére. Az így kapott eredményeket a 3.1. lemmában fogjuk megfogalmazni.

Annak érdekében, hogy kiszámoljuk az $y(n) = H(n)x(n) + \eta(n)$ vektor vetületeit az $\mathcal{M}_{n-1}$ és $\mathcal{V}_n$ alterekre vegyük észre, hogy $\eta(n) \perp \mathcal{M}_{n-1}$. Ez azért igaz, mert $\eta(n)$ független az $y(0), \dots, y(n-1)$ valószínűségi változóktól, ezért minden $v \in \mathcal{M}_{n-1}$ vektortól, és $E\eta(n) = 0$. Innen következik, hogy az $y(n)$ vektor vetülete az $\mathcal{M}_{n-1}$ altérre megegyezik a $H(n)x(n)$ vektor vetületével erre az altérre, azaz egyenlő a $H(n)\hat{x}(n|n-1)$ vektorral. Ezért

$$y(n) = H(n)\hat{x}(n|n-1) + z(n), \quad H(n)\hat{x}(n|n-1) \in \mathcal{M}_{n-1} \quad \text{és } z(n) \in \mathcal{V}_n, \tag{3.1}$$

ahol

$$z(n) = y(n) - H(n)\hat{x}(n|n-1) = \eta(n) + H(n)(x(n) - \hat{x}(n|n-1)). \tag{3.2}$$

Mivel $x(n+1) = \Phi(n)x(n) + \varepsilon(n)$, és $\varepsilon(n) \perp \mathcal{M}_{n-1}$, ezért az $\hat{x}(n+1|n)$ vektor felírható

$$\begin{aligned}\hat{x}(n+1|n) = \hat{x}(n+1|n-1) + u(n) = \Phi(n)\hat{x}(n|n-1) + u(n),\\ \Phi(n)\hat{x}(n|n-1) \in \mathcal{M}_{n-1} \quad \text{és } u(n) \in \mathcal{V}_n,\end{aligned} \tag{3.3}$$

ahol $u(n)$ az $\hat{x}(n+1|n)$ vektor vetülete a $\mathcal{V}_n$ altérre.

Annak érdekében, hogy az $u(n) = (u(n,1), \dots, u(n,p))^*$ vektort kiszámoljuk vegyük észre, hogy a $z(n) = (z(n,1), \dots, z(n,q))^*$ vektor koordinátái generátor rendszert alkotnak a $\mathcal{V}_n$ altérben. Ez következik az $y(n)$ és $z(n)$ vektorok közötti a (3.2) formulában megfogalmazott kapcsolatból, a $z(n) \in \mathcal{V}_n$, $H(n)\hat{x}(n|n-1) \in \mathcal{M}_{n-1}$ relációkból, és abból a tényből, hogy az $y(0), \dots, y(n)$ vektorok az $\mathcal{M}_n$ az $y(0), \dots, y(n-1)$ vektorok pedig az $\mathcal{M}_{n-1}$ alteret generálják. Ezért az $u(n)$ vektor mindegyik $u(n,j)$, $1 \leq j \leq p$, koordinátája kifejezhető $u(n,j) = \sum\limits_{k=1}^{q} c(j,k)z(n,k)$ lineáris kombináció alakban alkalmas $c(j,k)$ együtthatókkal. (A $c(j,k)$ együtthatók valójában függnek az n paramétertől is. De az n paramétert rögzítettük, ezért ezt a függést nem jelezzük.) Számoljuk ki a $c(j,k)$ együtthatókat.

Rögzítsünk egy j, $1 \le j \le p$, paramétert. A $c(j,k)$ együtthatók kiszámolása érdekében vegyük észre, hogy $x(n+1) - \hat{x}(n+1|n) \perp \mathcal{V}_n$. (Valójában az erősebb $x(n+1) - \hat{x}(n+1|n) \perp \mathcal{M}_n$ reláció is érvényes.) Ezért felírva az $\hat{x}(n+1|n)$ és $x(n+1)$ vektor $\hat{x}(n+1|n) = (\hat{x}(n+1|n,1), \dots, \hat{x}(n+1|n,p))^*$ és $x(n+1) = x(n+1,1), \dots, x(n+1,p))^*$ koordinátáit azt kapjuk, felhasználva az $\hat{x}(n+1|n-1) \perp \mathcal{V}_n$ és $z(n) \in \mathcal{V}_n$ relációkat és a (3.3) formula első azonosságát, hogy $Eu(n,j)z(n,l) = E\hat{x}(n+1|n,j)z(n,l) = Ex(n+1,j)z(n,l)$ minden $1 \le l \le q$ indexre. Ezért az $u(n,j) = \sum\limits_{k=1}^{q} c(j,k)z(n,k)$ azonosságot beszorozva $z(n,l)$-lel, és várható értéket véve, azt kapjuk minden $1 \le l \le q$ paraméterre, hogy

$$Ex(n+1,j)z(n,l) = Eu(n,j)z(n,l) = \sum_{k=1}^{q} c(j,k)Ez(n,k)z(n,l), \qquad 1 \le l \le q,$$

A keresett $c(j,k)$ együtthatókat ezen egyenletrendszer megoldásának a segítségével fogjuk megkapni.

Vezessük be a következő jelöléseket. Legyen $C(j) = (c(j,1), \dots, c(j,q))^*$, és $P(j) = (Ex(n+1,j)z(n,1), \dots, Ex(n+1,j)z(n,q))^*$. Ezen vektorok segítségével az előző egyenletet így írhatjuk fel:

$$P(j) = E[z(n)z(n)^*]C(j).$$

Tegyük fel, hogy az $E[z(n)z(n)^*]$ mátrix invertálható. Egyelőre csak ezzel az esettel foglalkozunk. Mivel $(E[z(n)z(n)^*])^{-1}$ szimmetrikus mátrix, ezért az előző egyenlet alapján

$$C(j)^* = P(j)^*(E[z(n)z(n)^*])^{-1},$$

és mivel $u(n,j) = C(j)^*z(n)$, innen azt kapjuk, hogy

$$u(n,j) = C(j)^*z(n) = P(j)^* \left(E[z(n)z(n)^*]\right)^{-1} z(n) \quad \text{minden } 1 \le j \le p \text{ indexre.}$$

Azt, hogy ez az azonosság az $u(n)$ vektor minden $1 \le j \le p$ koordinátájára igaz a következő módon írhatjuk zárt alakban, felhasználva a $P(j)$ vektorok alakját.

$$u(n) = E[x(n+1)z(n)^*] \left(E[z(n)z(n)^*]\right)^{-1} z(n). \tag{3.4}$$

A (3.2) formulában szereplő $z(n) = y(n) - H(n)\hat{x}(n|n-1)$ vektort ki tudjuk számolni az $y(0), \dots, y(n)$ vektorok ismeretében. Meg fogjuk mutatatni, hogy azt a mátrixot, amellyel a $z(n)$ vektort megszoroztuk a (3.4) formula jobboldalán ki tudjuk számolni a $\Phi(l)$, $H(l)$, $Q(l)$, $R(l)$, $1 \le l \le n$, és $\Sigma(0)$ mátrixok ismeretében. Innen következik, hogy az $u(n)$ és (a (3.3) formula alapján) az $\hat{x}(n+1|n)$ vektort is ki tudjuk számolni.

A fenti állítások bizonyításának érdekében kiszámoljuk a (3.4) formulában szereplő $E[x(n+1)z(n)^*]$ és $E[z(n)z(n)^*]$ mátrixokat. Bebizonyítjuk, hogy

$$E[x(n+1)z(n)^*] = \Phi(n)\Sigma(n)H(n)^*, \tag{3.5}$$

és
$$V(n) = E[z(n)z(n)^*] = H(n)\Sigma(n)H(n)^* + R(n), \tag{3.6}$$
ahol
$$\Sigma(n) = E[(x(n) - \hat{x}(n|n-1))(x(n) - \hat{x}(n|n-1))^*]. \tag{3.7}$$
Később a (3.7) formulában szereplő $\Sigma(n)$ kiszámolásáról bebizonyítjuk azt a számunkra hasznos formulát, amelyet a (2.5) képletben fogalmaztunk meg.

A (3.5) formula bizonyításában felhasználjuk, hogy $x(n+1) = \Phi(n)x(n) + \varepsilon(n)$, $\varepsilon(n) \perp \mathcal{M}_n$, (mert $\varepsilon(n)$ független minden az $\mathcal{M}_n$ térben levő valószínűségi változótól), és $z(n) \in \mathcal{M}_n$, ahonnan $E\varepsilon(n)z(n)^* = 0$. Ezért a (3.2) formula felhasználásával azt kapjuk, hogy
$$\begin{aligned} Ex(n+1)z(n)^* &= E(\Phi(n)x(n) + \varepsilon(n))z(n)^* = \Phi(n)Ex(n)z(n)^* \\ &= \Phi(n)E(x(n)[(x(n) - \hat{x}(n|n-1))^*H(n)^* + \eta(n)^*]. \end{aligned}$$
Továbbá, mivel $Ex(n)\eta(n)^* = 0$, az $x(n)$ és $\eta(n)$ valószínűségi változók függetlensége miatt, és $E\hat{x}(n|n-1)[x(n) - \hat{x}(n|n-1)]^* = 0$, ezért
$$\begin{aligned} &Ex(n+1)z(n)^* = \Phi(n)E(x(n)[(x(n) - \hat{x}(n|n-1))^*H(n)^*] \\ &\quad = \Phi(n)E([x(n) - \hat{x}(n|n-1)][(x(n) - \hat{x}(n|n-1))^*H(n)^*] = \Phi(n)\Sigma(n)H(n)^*. \end{aligned}$$
A (3.5) formulát beláttuk.

A (3.6) formula bizonyítása érdekében vegyük észre, hogy
$$E\eta(n)[x(n) - \hat{x}(n|n-1)]^* = E[x(n) - \hat{x}(n|n-1)]\eta(n)^* = 0,$$
mert $E\eta(n)x(n)^* = Ex(n)\eta(n)^* = 0$ az $\eta(n)$ és $x(n)$ változók függetlensége miatt, és
$$E\eta(n)\hat{x}(n|n-1)^* = E\hat{x}(n|n-1)\eta(n)^* = 0,$$
mivel $\eta(n) \perp \mathcal{M}_{n-1}$, és $\hat{x}(n|n-1) \in \mathcal{M}_{n-1}$. Ezért a (3.2) formula alapján
$$\begin{aligned} Ez(n)z(n)^* &= E\left[(\eta(n) + H(n)(x(n) - \hat{x}(n|n-1))(\eta(n) + H(n)(x(n) - \hat{x}(n|n-1))^*\right] \\ &= E(\eta(n)\eta(n)^* + H(n)E\left[(x(n) - \hat{x}(n|n-1))(x(n) - \hat{x}(n|n-1))^*\right]H(n)^* \\ &= R(n) + H(n)\Sigma(n)H(n)^*. \end{aligned}$$
A (3.4), (3.5) és (3.6) formulákból adódik, hogy
$$u(n) = \Phi(n)\Sigma(n)H(n)^*\{H(n)\Sigma(n)H(n)^* + R(n)\}^{-1}z(n) = K(n)z(n), \tag{3.8}$$
a (2.4) formulában definiált $K(n)$ mátrix-szal, feltéve, hogy $H(n)\Sigma(n)H(n)^* + R(n) = E[z(n)z(n)^*]$ invertálható mátrix. (Az $Ez(n)z(n)^*$ mátrix akkor és csak akkor invertálható, ha a $z(n)$ vektor koordinátái nemcsak generátor rendszert, hanem bázist is alkotnak a $\mathcal{V}_n$ térben.)

A most kapott eredményeket foglaljuk össze a következő 3.1. lemmában.

3.1. Lemma. *Tekintsük az (1.1) és (1.2) képlettel definiált Kálmán-féle szűrőt, és benne a 2. fejezetben definiált $\mathcal{M}_n$ Hilbert tereket, és $\hat{x}(n+m|n)$ véletlen vektorokat. Tekintsük az $\mathcal{M}_n$ Hilbert tér $\mathcal{M}_n = \mathcal{M}_{n-1} \oplus \mathcal{V}_n$ ortogonális felbontását, két ortogonális altér direkt összegére, és az $y(n) \in \mathcal{M}_n$ és $\hat{x}(n+1|n) \in \mathcal{M}_n$ felbontását e vektorok $\mathcal{M}_{n-1}$-beli és $\mathcal{V}_n$-beli vetületeinek az összegére. Ezeket a felbontásokat adtuk meg a (3.1) illetve (3.3) formulában, ahol a képletekben szereplő $z(n)$ illetve $u(n)$ $\mathcal{V}_n$ térbeli komponenseket a (3.2) illetve (3.4) vagy (3,8) formulában adtunk meg. A (3.4) és (3.8) formula akkor érvényes, ha az $E[z(n)z(n)^*]$, vagy ami ezzel ekvivalens a $H(n)\Sigma(n)H(n)^*+R(n)$ mátrix invertálható. E formulák bizonyítása során igazoltuk a (3.5) és (3.6) azonosságokat is, amelyek szintén hasznosak lesznek későbbi érvelésünkben.*

Rátérek a 2.1. tétel bizonyíására.

A 2.1. tétel bizonyítása. A (2.1) formula igazolása viszonylag egyszerű. Mivel $x(n+m) = \Phi(n+m-1)x(n+m-1) + \varepsilon(n+m-1)$, és az $\varepsilon(n+m-1)$ véletlen vektor merőleges az $\mathcal{M}_n$ térre, (mert független az $\mathcal{M}_n$ térbeli valószínűségi változóktól), ezért véve ezen azonosság tagjainak vetületét az $\mathcal{M}_n$ altérre a bizonyítandó $\hat{x}(n+m|n) = \Phi(n+m-1)\hat{x}(n+m-1|n)$ azonosságot kapjuk.

A (2.2) formula bizonyításának érdekében felírjuk a (3.3) és (3.8) formulák majd a (3.2) formula segítségével az

$$\begin{aligned}\hat{x}(n+1|n) &= \Phi(n)\hat{x}(n|n-1) + K(n)z(n) \\ &= \Phi(n)\hat{x}(n|n-1) + K(n)(y(n) - H(n)\hat{x}(n|n-1)) \\ &= (\Phi(n) - K(n)H(n))\hat{x}(n|n-1) + K(n)y(n)\end{aligned}$$

azonosságot. Mivel $\Phi(n) - K(n)H(n) = \Psi(n)$ a (2.3) formula szerint, ezekből a relációkból következik a (2.2) reláció (a (2.3) és (2.4) formulában bevezetett formulákkal együtt).

A 2.1. tétel bizonyításának befejezéséhez igazolni kell még a (2.5) azonosságot. Az e formula baloldalán szereplő és a (3.7) formulában definiált $\Sigma(n+1)$ kiszámítása érdekében vizsgáljuk először az $x(n+1) - \hat{x}(n+1|n)$ kifejezést. Felírhatjuk, hogy

$$x(n+1) - \hat{x}(n+1|n) = [x(n+1) - \hat{x}(n+1|n-1)] - [\hat{x}(n+1|n) - \hat{x}(n+1|n-1)].$$

Továbbá

$$\hat{x}(n+1|n) - \hat{x}(n+1|n-1) = u(n) = K(n)z(n)$$

a (3.3) formula első azonossága, és a (3.8) azonosság alapján, és

$$x(n+1) - \hat{x}(n+1|n-1) = \Phi(n)[x(n) - \hat{x}(n|n-1)] + \varepsilon(n) \tag{3.9}$$

az (1.1) formulából, ha azt az $n+1$ paraméterre alkalmazzuk az n paraméter helyett, illetve az $\hat{x}(n+1|n-1) = \Phi(n)\hat{x}(n|n-1)$ azonosságból, amit szintén az $n+1$ paraméterre felírt (1.1) azonosságból kapunk, ha vesszük az azonosság két oldalának a vetítését az $\mathcal{M}_{n-1}$ altérre, és felhasználjuk azt, hogy $\varepsilon(n) \perp \mathcal{M}_{n-1}$. Innen

$$x(n+1) - \hat{x}(n+1|n) = \Phi(n)[x(n) - \hat{x}(n|n-1)] + \varepsilon(n) - K(n)z(n),$$

és a (3.7) formula alapján

$$\Sigma(n+1) = E[\Phi(n)[x(n) - \hat{x}(n|n-1)] + \varepsilon(n) - K(n)z(n)] \\ [\Phi(n)[x(n) - \hat{x}(n|n-1)] + \varepsilon(n) - K(n)z(n)]^*. \tag{3.10}$$

Elvégezve a beszorzásokat a (3.10) formulában azt kapjuk, hogy

$$\begin{aligned}\Sigma(n+1) = \Phi(n)\Sigma(n)\Phi(n)^* + Q(n) + EK(n)z(n)z(n)^*K(n)^* \\ - E[[x(n+1) - \hat{x}(n+1|n-1)](K(n)z(n))^*] \\ - E[(K(n)z(n))[x(n+1) - \hat{x}(n+1|n-1)]^*].\end{aligned} \tag{3.11}$$

A vegyes szorzatok számolásánál a (3.11) formulában kihasználtuk a (3.9) formulát, és azt hogy $E[\Phi(n)[x(n) - \hat{x}(n|n-1)]\varepsilon(n)^* = E\varepsilon(n)[\Phi(n)[x(n) - \hat{x}(n|n-1)]^* = 0$. Ez utóbbi állítás azért igaz, mert az $\varepsilon(n)$ véletlen vektor független, mind az $x(n)$ mind a $\hat{x}(n|n-1) \in \mathcal{M}_{n-1}$ vektortól.

Azt állítom, hogy

$$E[[x(n+1) - \hat{x}(n+1|n-1)](K(n)z(n))^*] = K(n)E[z(n)z(n)^*]K(n)^*, \tag{3.12}$$

ahonnan természetesen a

$$E[(K(n)z(n))[x(n+1) - \hat{x}(n+1|n-1)]^*] = K(n)E[z(n)z(n)^*]K(n)^*$$

azonosság is következik.

Valóban $E\hat{x}(n+1|n-1)z(n)^* = 0$, mert $z(n) \in \mathcal{V}_n$ (definició szerint), és $\hat{x}(n+1|n-1) \perp \mathcal{V}_n$, mivel $\hat{x}(n+1|n-1) \in \mathcal{M}_{n-1}$. Hasonlóan, $E[x(n+1) - \hat{x}(n+1|n)]z(n)^* = 0$, mert $x(n+1) - \hat{x}(n+1|n) \perp \mathcal{M}_n$. Innen

$$\begin{aligned}E[[x(n+1) - \hat{x}(n+1|n-1)](K(n)z(n))^*] &= Ex(n+1)z(n)^*K(n)^* \\ &= E\hat{x}(n+1|n)z(n)^*K(n)^* = Eu(n)z(n)^*K(n)^* = EK(n)z(n)z(n)^*K(n)^*.\end{aligned}$$

A fenti számolás utolsó két azonossága a (3.3) és (3.8) formulákból következik, amelyek szerint $u(n) = K(n)z(n)$, és $u(n)$ az $\hat{x}(n+1|n)$ vektor vetülete a $z(n)$ vektort tartalmazó $\mathcal{V}_n$ altérre.

Ezzel a (3.12) formulát beláttuk. A (3.11) és (3.12) formulákból következik, hogy

$$\Sigma(n+1) = \Phi(n)\Sigma(n)\Phi(n)^* + Q(n) - K(n)Ez[(n)z(n)^*]K(n)^*,$$

ahonnan a (3.6) formula segítségével megkapjuk a bizonyítani kívánt (2.5) formulát.

4. Az eredmények áttekintése. Az általános eset vizsgálata.

E jegyzetben ismertettük azt, hogy ismerve egy az (1.1) és (1.2) formulákban definiált Kálmán-féle szűrő által egy n időpontig előállított $x(1),\dots,x(n)$ változók, $y(1),\dots,$ $y(n)$ 'filtráltjainak' az értékeit, hogyan tudjuk megadni a jövőben megjelenő $x(n+m)$ változók, $m\ge 1$, értékeinek az optimális becslését. Ennek a becslésnek van néhány olyan tulajdonsága, amelyeket érdemes külön tárgyalni.

A vizsgált feladatban az $x(n+m)$, $m\ge 1$, véletlen vektorok $x(n+m|n)$ vetületét akarjuk kiszámolni az $y(1),\dots,y(n)$ véletlen vektorok koordinátái által generált, és a jegyzetben pontosan definiált $\mathcal{M}_n$ Hilbert térbe. A (2.1) képlet alapján az $\hat{x}(n+1|n)$ becslések ismeretében ki tudjuk számolni az $\hat{x}(n+m|n)$ becsléseket tetszőleges $m\ge 1$ paraméterre.

Érdemes megérteni, hogy az $\hat{x}(n+1|n)$ becsléseket az n paraméter szerinti rekurzióval tudjuk kiszámolni. Ez jelen esetben a következőt jelenti. Jelölje $\Sigma(n)=E[x(n)-\hat{x}(n|n-1)][x(n)-\hat{x}(n|n-1)]^*$ az $x(n)$ valószínűségi változóra adott $\hat{x}(n|n-1)$ becslés $x(n)-\hat{x}(n|n-1)$ hibájának a kovariancia mátrixát. Az n-ik időpontban kiszámoljuk az $(\hat{x}(n|n-1),\Sigma(n))$ párt. Ezután, az $n+1$-ik időpontban, az $y(n)$ véletlen vektor megismerése után, ki akarjuk számolni az $(\hat{x}(n+1|n),\Sigma(n+1))$ párt is. Ezt a (2.2), (2.3), (2.4) és (2.5) formulák segítségével tesszük meg. E formulák alkalmazásához elegendő az $(\hat{x}(n|n-1),\Sigma(n))$ párt és az $y(n)$ véletlen vektort ismerni. (Feltesszük, hogy az (1.2), (1.3) formulákban szereplő $\Phi(n)$ és $H(n)$ mátrixokat, illetve az $\varepsilon(n)$ és $\eta(n)$ 'hiba vektorok' $Q(n)$ és $R(n)$ kovariancia mátrixait szintén ismerjük.) Természetesen, ahhoz, hogy ezt a rekurziós eljárást alkalmazni tudjuk, ismernünk kell a kiinduló, nulla időpontbeli $(\hat{x}(0|-1),\Sigma(0))$ párt is. Ezt az $\hat{x}(0|-1)\equiv 0$, és $\Sigma(0)=Ex(0)x(0)^*$ képletek határozzák meg.

A fenti eredmények a következőt jelentik. Az $x(n+1)$ véletlen vektor $\hat{x}(n+1|n)$ optimális becslésének a kiszámításában a teljes $y(0),\dots,y(n)$ megfigyelés sorozatot felhasználhatjuk. De az $y(1),\dots,y(n-1)$ megfigyelések helyett elegendő az általuk meghatározott $\hat{x}(n|n-1)$ és $\Sigma(n)$ mennyiségek ismerete. Ezek tartalmazzák az összes számunkra fontos ismeretet az $(y(1),\dots,y(n-1))$ véletlen vektorról. Sőt, ezek az ismeretek nemcsak az $\hat{x}(n+1|n)$, hanem a $\Sigma(n+1)$ mennyiség kiszámolásához is elegendőek. Ez lehetővé teszi, hogy az $\hat{x}(n+1|n)$ vektort a $\Sigma(n+1)$ mátrix-szal együtt rekurzíve kiszámolhassuk.

A fenti számolások akkor végezhetőek el, ha az $E[z(n)z(n)^*]=H(n)\Sigma(n)H(n)^*+R(n)$ mátrix invertálható. Felmerül a kérdés, mit tehetünk abban az esetben, ha ez a feltétel nem teljesül. Az alábbiakben ezzel a kérdéssel foglalkozom.

A $z(n)$ véletlen vektor úgy jelent meg, mint az $y(n)$ vektor koordinátáinak vetülete a $\mathcal{M}_n$ Hilbert tér $\mathcal{M}_{n-1}$ alterének $\mathcal{V}_n$ ortogonális kiegészítőjére. A $z(n)$ vektor koordinátái generátor rendszert alkotnak az $\mathcal{V}_n$ Hilbert térben, de lehet, hogy nem alkotnak bázist. Ha e koordináták nem alkotnak bázist, akkor alkalmas módon véges sokat el lehet hagyni belőlük úgy, hogy a megtartott koordináták bázist alkossanak. Meg fogjuk mutatni, hogy a 'fölösleges' koordináták elhagyásával meg tudjuk oldani a vizsgált feladatot az eredeti eljárás természetes módosításával. Ekkor ugyanis invertálható mátrixokkal kell dolgoznunk. Az előbb vázolt eljárás jogosságának igazolásához szükségünk lesz az alábbi feladatban megfogalmazott eredményre.

Feladat. Legyen $Z = (Z_1, \dots, Z_q)^*$ egy q dimenziós normális eloszlású véletlen vektor Σ kovariancia mátrix-szal és nulla várható értékkel. Tekintsük ennek $Z(j_1, \dots, j_k) = (Z_{j_1}, \dots, Z_{j_k})$ részvektorát tetszőleges $\{j_1, \dots, j_k\} \subset \{1, \dots, q\}$ indexhalmazra, és jelölje $\Sigma(j_1, \dots, j_k)$ a $Z(j_1, \dots, j_k)$ vektor kovariancia mátrixát. A $\Sigma(j_1, \dots, j_k)$ mátrix megegyezik a Σ mátrix megszorításával a j_1-ik, … és j_k-ik sorokra és oszlopokra. A $Z(j_1, \dots, j_k)$ vektor $Z_{j_1}, \dots, Z_{j_k}$ koordinátái akkor és csak akkor lineárisan függetlenek, azaz e valószínűségi változóknak akkor és csak akkor nincs olyan nem triviális lineáris kombinációja, amelyik egy valószínűséggel nullával egyenlő, ha a $\Sigma(j_1, \dots, j_k)$ mátrix invertálható. Speciálisan, a maximális méretű lineárisan független koordinátákból álló $Z(j_1, \dots, j_k)$ vektorok indexhalmazai megegyeznek azon maximális méretű $\Sigma(j_1, \dots, j_k)$ mátrixok indexhalmazaival, amelyek invertálhatóak.

Segítség. A $Z(j_1, \dots, j_k)$ vektor normális eloszlású minden $j_1, \dots, j_k$ indexhalmazra, ezért felírható $Z(j_1, \dots, j_k) = \eta A$ alakban, ahol $\eta = (\eta_{j_1}, \dots, \eta_{j_k})^*$ független, standard normális η_{j_s} valószínűségi változókból álló vektor, A pedig $k \times k$ méretű mátrix. (Ebben a jelölésben elhagytuk az A mátrix és η vektor függését a $j_1, \dots, j_k$ indexektől.) Írjuk fel az A mátrixot $A = P\Lambda Q$ alakban, ahol P és Q unitér, Λ pedig diagonális mátrix, nem negatív elemekkel a mátrix átlójában. A $Z(j_1, \dots, j_k)$ vektor koordinátái akkor és csak akkor lineárisan függetlenek, ha az A mátrix invertálható. Ez akkor és csak akkor teljesül, ha a Λ mátrix diagonális elemei mind nullától különböznek. A $\Sigma(j_1, \dots, j_k) = P\Lambda^2 P$ azonosság is teljesül, ahonnan következik, hogy $\Sigma(j_1, \dots, j_k)$ akkor és csak akkor invertálható, ha a Λ mátrix diagonális elemei mind nullától különböznek.

Annak érdekében, hogy a Kálmán-féle szűrő becslésének a feladatát megoldjuk az általános esetben, (akkor is, ha az $E[z(n)z(n)^*]$ mátrix nem invertálható), vegyük észre, hogy a 2.1. tétel akkor is érvényes marad, ha az az $y(n)$ véletlen vektor (1.2) formulában megadott definiciójában megengedjük, hogy az ott szereplő $H(n)$ mátrix és $\eta(n)$ vektor, $q(n) \times p$, illetve $q(n)$ méretű legyen, azaz az $y(n)$ vektor dimenziója függhet az n paramétertől.

Legyen a $V(n) = E[z(n)z(n)^*] = H(n)\Sigma(n)H(n)^* + R(n)$ kovariancia mátrixra a $\{j_1, j_2, \cdots, j_k\} \subset \{1, \dots, q\}$ egy olyan maximális, $k = k(n)$ méretű halmaz, amelyre a $\Sigma = V(n)$ mátrixhoz tartozó, a 'Feladat'-ban bevezetett $\Sigma(j_1, \dots, j_k)$ mátrix invertálható. Vegyük észre, hogy a 'Feladat' eredményéből következik, hogy a $z(n) = (z(n,1), \dots, z(n,q))^*$ Gauss eloszlású véletlen vektor j_1-ik, … , j_k-ik koordinátái egy maximális független rendszert alkotnak. Mivel a $z(n,j)$, $1 \le j \le q$, valószínűségi változók generátort alkotnak a $\mathcal{V}_n$ térben innen az is következik, hogy $z(n, j_1)$, … , $z(n, j_k)$ bázis a $\mathcal{V}_n$ térben. Jelölje $\bar{H}(n) = \bar{H}(n, j_1, \dots, j_k)$ azt a $k \times p$ méretű mátrixot, amelyet úgy kapunk, hogy az (1.2) formulának csak a j_1-ik, … , és j_k-ik sorait tartjuk meg, legyen $\bar{\eta} = \bar{\eta}(j_1, \dots, j_k)$ a szintén (1.2) formulában szereplő η vektor j_1-ik, … , és j_k-ik koordinátáiból álló vektor, és tekintsük a

$$\bar{y}(n) = \bar{H}(n)x(n) + \bar{\eta}(n) \tag{4.1}$$

egyenletet, azaz vegyük az (1.2) egyenlet megszorítását a j_1-ik, … , és j_k-ik sorokra.

Ekkor a $\bar{\eta}(n)$ vektor $\bar{R}(n) = E\bar{\eta}(n)\bar{\eta}(n)^*$ kovarianciamátrixa megegyezik az (1.4) formulában definiált $R(n)$ kovariancia mátrix megszorításával a j_1-ik, … , és j_k-ik sorokra és oszlopokra. Továbbá a $\bar{y}(n)$ vektor $z(n, j_s)$, $1 \le s \le k$ ortogonális vetületei az $\mathcal{M}_n$ térre ugyanazt a $\mathcal{V}_n$ Hilbert teret generálják, mint az $y(n)$ tér $z(n)$ vetülete, és ezenkívül a $\mathcal{V}_n$ tér bázisát is alkotják. Ezért a Kálmán féle feladat ekvivalens változatát kapjuk, ha az (1.2) formulát a (4.1) formulával helyettesítjük.

Az új módosított feladatban (ahol az (1.2) formulát a (4.1) formulával helyettesítettük) a (3.6) formulában definiált $V(n) = E[z(n)z(n)^*]$ mátrix helyett a $V(n)$ mátrix j_1-ik, ..., és j_k-ik soraira és oszlopaira vett $V(n|j_1, \dots, j_k)$ megszorításával kell számolnunk, ami egy invertálható mátrix. Ezért ebben az esetben alkalmazható a 2.1. tétel. Az egyetlen különbség az, hogy a $K(n)$ mátrix (2.4) formulában megadott definiciójában a $H(n)$ mátrixot a $\bar{H}(n) = \bar{H}(n|j_1, \dots, j_k)$, az $R(n)$ mátrixot pedig az $\bar{R}(n) = \bar{R}(n|j_1, \dots, j_k)$ mátrix-szal kell helyettesíteni.