Az november 16-i szeminárium témája
Rövid összefoglaló

Először megbeszéltük a következő házifeladat megoldását:

Tekintsünk egy szabályos pénzdarab 10 000 egymás utáni (független) feldobásából származó fej-írás sorozatot. Adjunk becslést a Chebishev egyenlőtlenség segítségével annak valószínűségére, hogy a fej-dobások számának eltérése a várt 5000 számtól legalább 100-zal, illetve legalább 200-zal eltér! Milyen becslést ad ezekre a valószínűségekre a centrális határeloszlástétel?

Megoldás: Vezessük be a következő ξ_j, $\le j \le 10\,000$ valószínűségi változókat, melyekre $\xi_j = 1$, ha a j-ik dobás eredménye fej, $\xi_j = 0$, ha a j-ik dobás eredménye írás. Ekkor ξ_j, $1 \le j \le 10\,000$ független valószínűségi változók, $E\xi_j = \frac{1}{2}$, $\operatorname{Var}\xi_j = \frac{1}{4}$, és a $P\left(\left|\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)\right| > 100\right)$ és $P\left(\left|\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)\right| > 200\right)$ valószínűségekre kell becslét adnunk. A Chebishev egyenlőtlenség az első valószínűségre a $\dfrac{10000\cdot\frac{1}{4}}{100^2} = \dfrac{1}{4}$, a második valószínűségre pedig a $\dfrac{10000\cdot\frac{1}{4}}{200^2} = \dfrac{1}{16}$ felső becslést adja.

A centrális határeloszlástétel szerint $P\left(\dfrac{\sum\limits_{j=1}^{10000}(\xi_j - E\xi_j)}{\sqrt{10000\cdot\frac{1}{4}}} > u\right) \sim 1 - \Phi(u)$. Innen kapjuk, hogy az első valószínűség kiszámításához az $u = \pm\frac{100}{\frac{1}{2}\cdot 100} = \pm 2$ értékeket kell tekinteni, és a vizsgált valószínűség közelítőleg $(1-\Phi(2)) + \Phi(-2) = 2(1-\Phi(2)) \sim 2(1-0.97720) = 0.0456$. A második valószínűség hasonlóan körülbelül $2(1-\Phi(4)) \sim 0$, (az első 4 tizedesjegy 0).

A centrális határeloszlástétel tárgyalása.

Meg lehet adnia az elégséges sőt alkalmas megfogalmazásban szükséges és elégséges feltételét annak, hogy a centrális határeloszlástétel teljesüljön. Ez a következő eredmény:

Centrális határeloszlástétel. *Legyen ξ_k, $k = 1, 2, \dots$, független valószínűségi változók sorozata, melyre $E\xi_k = 0$, $\sigma_k^2 = E\xi_k^2 < \infty$, $k = 1, 2, \dots$, és az $s_n^2 = \sum\limits_{k=1}^{n}\sigma_k^2$, $n = 1, 2, \dots$, sorozat teljesíti a $\lim\limits_{n\to\infty} s_n^2 = \infty$ feltételt. Ha ezenkívül a ξ_n, $n = 1, 2, \dots$, sorozat teljesíti a következő úgynevezett Lindeberg feltételt, mely azt követeli meg, hogy minden $\varepsilon > 0$ számra teljesüljön a*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{n} E\xi_k^2 I(\{|\xi_k| > \varepsilon s_n\}) = 0.$$

reláció, akkor az $\dfrac{S_n}{s_n}$ *sorozat, ahol* $S_n = \sum\limits_{k=1}^{n} \xi_k$ *eloszlásban konvergál a standard normális eloszlásfüggvényhez* $n \to \infty$ *esetén.*

Igaz a fenti állítás megfordítása is: Ha a fenti jelöléseket használva a ξ_k, $E\xi_k = 0$, $E\xi_k^2 = \sigma_k^2$, $k = 1, 2, \dots$, független valószínűségi változókból álló sorozatra $s_n^2 \to \infty$, $\sup\limits_{1\le k\le n} \dfrac{\sigma_k^2}{s_n^2} \to 0$, ha $n \to \infty$, és az $\dfrac{S_n}{s_n}$ sorozat eloszlásban konvergál a standard normális eloszlásfüggvényhez $n \to \infty$ esetén, akkor ez a sorozat teljesíti a Lindeberg feltételt.

Jegyezzük meg, hogy ebben a megfordításban nem csak azt követeltük meg, hogy a határeloszlástétel normális legyen, hanem azt is, hogy a "helyes szórással" rendelkezzen. A $\sup\limits_{1\le k\le n} \dfrac{\sigma_k^2}{s_n^2} \to 0$, ha $n \to \infty$, feltétel megkövetelése természetes. Egyrészt, ez azt fejezi ki szemléletesen, hogy az egyes összeadandók hatása elhanyagolható, azaz a centrális határeloszlástétel sok kis eredmény összhatásaként áll elő, másrészt e feltételnek teljesülnie kell, mert következménye a Lindeberg feltételnek. Valóban, tetszőleges $\varepsilon > 0$ és $1 \le k \le n$ számokra

$$\begin{aligned}\frac{\sigma_k^2}{s_n^2} = \frac{E\xi_k^2}{s_n^2} &= \frac{E(\xi_k^2 I(|\xi_k| \le \varepsilon s_n)}{s_n^2} + \frac{E(\xi_k^2 I(|\xi_k > \varepsilon s_n)}{s_n^2} \\ &\le \varepsilon^2 + \frac{1}{s_n^2}\sum_{k=1}^{n} E\xi_k^2 I(\{|\xi_k| > \varepsilon s_n\}) \le 2\varepsilon^2,\end{aligned}$$

ha $n \ge n_0(\varepsilon)$.

Két természetes kérdés:

1.) Tudunk-e példát adni arra, hogy a centrális határeloszlástétel nem teljesül, és mi van az ilyen példák hátterében?

2.) Milyen egyszerűen ellenőrizhető, hasznos feltételeket tudunk adni arra, hogy a Lindeberg feltétel, ezért a centrális határeloszlástétel teljesüljön?

1.) Legyenek ξ_k, $k = 1, 2, \dots$, független valószínűségi változók, $P(\xi_k = k) = P(\xi_k = -k) = \dfrac{1}{2k^2}$, $P(\xi_k = 0) = 1 - \dfrac{1}{k^2}$. Ekkor $E\xi_k = 0$, $E\xi_k^2 = 1$, $s_n^2 = n$. Viszont a Borel–Cantelli lemma alapján, mivel $\sum\limits_{k=1}^{\infty} P(\xi_k \ne 0) < \infty$ egy valószínűséggel csak véges sok k indexre teljesül az, hogy $\xi_k(\omega) \ne 0$, ezért $\sup\limits_{1\le n<\infty} |S_n(\omega)| < \infty$, és $\dfrac{S_n(\omega)}{s_n} \to 0$ egy valószínűséggel. Lehet rafináltabb példákat is konstruálni. (Például tekinthetjük független η_k, $k = 1, 2, \dots$, valószínűségi változók sorozatát, melyekre $\eta_{2k} = \xi_k$, és az η_{2k-1}, $k = 1, 2, \dots$ valószínűségi változók minden egyéb valószínűségi változótól független standard normális eloszlású valószínűségi változók. Ekkor a $\dfrac{1}{\sqrt{n}} \sum\limits_{k=1}^{n} \xi_k$ véletlen összegek eloszlásban tartanak egy nulla várható értékú, és $\frac{1}{2}$, azaz a vártnál kisebb szórásnégyzetű normális valószínűségi változóhoz.) De minden ellenpélda hátterében az van, hogy bizonyos valószínűségi változók kis valószínűséggel rendkívül nagy értékeket vesznek fel. Ezek a rendkívüli értékek csak nagyon kis mértékben befolyásolják a normalizált összeg eloszlását, de nagyon befolyásolják az összeg szórásnégyzetét.

2.) A következő célunk az, hogy a Lindeberg feltételre, tehát a centrális határeloszlástételre jól ellenőrízhető és elég általános feltételt adjunk. Ennek érdekében a következő állítást illetve annak bizonyítását fogjuk tárgyalni.

Legyen ξ_n, $n = 1, 2, \dots$, független valószínűségi változók sorozata, melyre $E\xi_n = 0$, $\sigma_n^2 = E\xi_n^2 < \infty$, $n = 1, 2, \dots$, $\lim\limits_{n\to\infty} s_n^2 = \infty$, ahol $s_n^2 = \sum\limits_{k=1}^n \sigma_k^2$. Ez a sorozat teljesíti a Lindeberg feltételt, ha a következő tulajdonságok egyike teljesül.

a.) $E|\xi_k|^{2+\alpha} < \infty$, minden $k = 1, 2, \dots$ számra valamilyen $\alpha > 0$ konstanssal, és $\lim\limits_{n\to\infty} \dfrac{\left(\sum\limits_{k=1}^n E|\xi_k|^{2+\alpha}\right)^{2/(2+\alpha)}}{s_n^2} = 0$. Speciálisan, ez a feltétel teljesül akkor, ha $E\xi_k^2 \ge K$ valamilyen $K > 0$ számmal minden $k = 1, 2, \dots$ indexre, és ezenkívül érvényes az $E \lim\limits_{k\to\infty} k^{-\alpha/2} E|\xi_k|^{2+\alpha} = 0$ reláció.

b.) A független ξ_n, $n = 1, 2, \dots$, valószínűségi változók egyforma eloszlásúak. (Tehát független egyforma eloszlású valószűűségi változók sorozata is teljesíti a Lindeberg feltételt, ezért a rájuk vonatkozó centrális határeloszlástétel speciális esete a centrális határeloszlástétel Lindeberg feltételből következő alakjának.)

Bizonyítás: A Hölder egyenlőtlenség alapján

$$\sum_{k=1}^n E\xi_k^2 I(|\xi_k| > \varepsilon s_n) \le \left(\sum_{k=1}^n E|\xi_k|^{(2+\alpha)}\right)^{(2/\alpha+2)} \left(\sum_{k=1}^n P(|\xi_k| > \varepsilon s_n)\right)^{\alpha/(2+\alpha)}.$$

Másrészt, a Chebishev egyenlőtlenség alapján $\sum\limits_{k=1}^n P(|\xi_k| > \varepsilon s_n) \le \sum\limits_{k=1}^n \dfrac{E\xi_k^2}{\varepsilon^2 s_n^2} = \dfrac{1}{\varepsilon^2}$. Ezért ha az a) eset feltételei teljesülnek, akkor

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^n E\xi_k^2 I(|\xi_k| > \varepsilon s_n) = 0,$$

azaz teljesül a Lindeberg feltétel is.

Az a.) rész végén tekintett esetben $s_n \ge \text{const.}\, n$, és $\sum\limits_{k=1}^E |\xi_k|^{2+\alpha} = o\left(n^{(\alpha+2)/2}\right)$, ha $n \to \infty$, ezért ekkor teljesül a a.) részben megfogalmazott feltétel.

Ha $\xi_1, \xi_2, \dots$, független egyforma eloszlású valószínűségi változók sorozata, $E\xi_1 = 0$, $0 < E\xi_1^2 < \infty$, akkor $\dfrac{1}{s_n^2} \sum\limits_{k=1}^n E\xi_k^2 I(|\xi_k| > \varepsilon s_n) = \dfrac{1}{E\xi_1^2} E\xi_1^2 I\left(|\xi_1| > \varepsilon\sqrt{nE\xi_1^2}\right) \to 0$, ha $n \to \infty$. Tehát ebben az esetben is teljesül a Lindeberg feltétel.

A több dimenziós normális eloszlás

Az első megtárgyalandó kérdés az, hogy miért fontos számunkra ez a fogalom. A háttérben a következő kérdés áll. Láttuk, hogy független (valós, azaz egydimenziós) valószínűségi változók normalizált részletösszegei nagyon általános feltételek mellett normális eloszláshoz konvergálnak. Felmerül a kérdés: Tekintsünk független k-dimenziós $\xi_j = \left(\xi_j^{(1)},\dots,\xi_j^{(k)}\right)$, $j=1,2,\dots$, véletlen valószinűségi vektorokat, mely ξ_j vektorok függetlenek. Az egyszerűség kedvéért korlátozzuk figyelmünket arra az esetre, amikor a ξ_j vektorok egyforma eloszlásúak. Tegyük fel továbbá, hogy $E\xi_j^{(l)}=0$, $1\le l\le k$. Viszont az egyes $\xi_j = \left(\xi_j^{(1)},\dots,\xi_j^{(k)}\right)$ vektorok koordinátáinak kapcsolatáról semmilyen korlátozó feltételt nem teszünk. Tekintsük a $\dfrac{1}{\sqrt n}\sum\limits_{j=1}^n \xi_j = \dfrac{1}{\sqrt n}\left(\sum\limits_{j=1}^n \xi_j^{(1)},\dots,\sum\limits_{j=1}^n \xi_j^{(k)}\right)$ normalizált összegeket. Arra vagyunk kiváncsiak, hogy igaz-e az egy-dimenziós centrális határeloszlástétel több-dimenziós változata, azaz igaz-e az, hogy a fenti véletlen vektorokból készített normalizált részletösszegek nagyon általános feltételek mellett eloszlásban konvergálnak. Továbbá, milyen eloszlások jelennek meg mint határeloszlások? A válasz az első kérdésre igenlő. A lehetséges határeloszlások jól jellemezhetőek, és ezeket nevezik több-dimenziós normális eloszlásoknak.

Természetesen a több-dimenziós határeloszlástételeket az egy-dimenziós problémákhoz hasonlóan tárgyalhatjuk. Sőt, a több-dimenziós probléma egyszerűen visszavezethető az egydimenziós esetre a következő észrevétel segítségével.

Lemma: *Legyen $U^{(n)} = \left(U_1^{(n)},\dots,U_k^{(n)}\right)$, $n=1,2,\dots$, k-dimenziós valószínűségi változók sorozata. Ez a sorozat akkor és csak akkor konvergál eloszlásban egy $U=(U_1,\dots,U_k)$ k-dimenziós véletlen vektor eloszlásához, ha tetszőleges $(a_1,\dots,a_k)$ együtthatókra a $\sum\limits_{l=1}^k a_l U_l^{(n)}$ lineáris kombinációk eloszlásban konvergálnak a $\sum\limits_{l=1}^k a_l U_l$ lineáris kombinációhoz, ha $n\to\infty$.*

A lemma bizonyítása azon alapul, hogy a karakterisztikus függvényekre kimondott eredmények érvényesek több-dimenziós eloszlások karakterisztikus függvényeire is. Nevezetesen, egy $U=(U_1,\dots,U_k)$ valószínűségi változó eloszlását meghatározza annak $\varphi(t_1,\dots,t_k) = Ee^{i(t_1U_1+\cdots+t_kU_k)}$, $(t_1,\dots,t_k)\in R^k$, karakterisztikus függvénye, továbbá ha valószínűségi vektorok egy $U^{(n)} = \left(U_1^{(n)},\dots,U_k^{(n)}\right)$ sorozata akkor és csak akkor konvergál egy $U=(U_1,\dots,U_k)$ véletlen vektorhoz, ha e változók $\varphi_n(t_1,\dots,t_k) = Ee^{i(t_1U_1^{(n)}+\cdots+t_kU_k^{(n)})}$ karakterisztikus függvényei minden $(t_1,\dots,t_k)\in R^k$ pontban konvergál az $U=(U_1,\dots,U_k)$ véletlen vektor $\varphi(t_1,\dots,t_k) = Ee^{i(t_1U_1+\cdots+t_kU_k)}$ karakterisztikus függvényéhez. Továbbá, ha $\varphi_{a_1,\dots,a_k}^{(n)}(t) = E\left\{t\sum\limits_{l=1}^k a_lU_l^{(n)}\right\}$ jelöli a $\sum\limits_{l=1}^k a_lU_l^{(n)}$ valószínűségi változók karakterisztikus függvényét, akkor

$$\varphi_{a_1,\dots,a_k}^{(n)}(t) = \varphi^{(n)}(a_1t,\dots,a_kt),$$

és az analog eredmény érvényes az index nélküli $(U_1,\dots,U_k)$ vektorra és az e vektor koordinátáiból képzett lineáris kombinációkra. Ezért átfogalmazva a Lemma állítását a karakterisztikus függvények nyelvére megkapjuk a Lemma bizonyítását.

(Megjegyezzük, hogy az egy-dimenziós esetben megfogalmazott az itt felidézettnél tartalmasabb állítás megfelelője a több-dimenziós esetben is érvényes. Ha tudjuk, hogy az $U^{(n)} = \left(U_1^{(n)},\dots,U_k^{(n)}\right)$ véletlen vektorok $\varphi_n(t_1,\dots,t_k) = Ee^{i(t_1U_1^{(n)}+\cdots+t_kU_k^{(n)})}$ karakterisztikus függvényei egy az origóban folytonos függvényhez konvergálnak, akkor a limeszfüggvény egy eloszlás karakterisztikus föggvénye, és a tekintett véletlen vektorok eloszlásban konvergálnak ehhez az eloszláshoz.)

Térjünk vissza a kiiduló problémához! Tekintsünk független k-dimenziós $\xi_j = \left(\xi_j^{(1)},\dots,\xi_j^{(k)}\right)$, $j=1,2,\dots$, véletlen valószinűségi vektorokat, melyek függetlenek, egyforma eloszlásúak és nulla várható értékűek. Tegyük fel továbbá, hogy $E\left(\xi_1^{(p)}\right)^2<\infty$, $1\le p\le k$. Vezessük be a $D(p,q)=E\xi_1^{(p)}\xi_1^{(q)}<\infty$, $1\le p,q\le k$. Ekkor a $\frac{1}{\sqrt{n}}\sum\limits_{j=1}^n\sum\limits_{p=1}^k a_p\xi_j^{(p)}$ valószínűségi változók az egy-dimenziós centrális határeloszlástétel szerint konvergálnak a nulla várható értékű és $\sum\limits_{p=1}^k\sum\limits_{q=1}^l D(p,q)a_pa_q$ szórásnégyzettel, azaz $\varphi(t)=\exp\left\{-t^2\sum\limits_{p=1}^k\sum\limits_{q=1}^l D(p,q)a_pa_q \Big/ 2\right\}$ karakterisztikus függvénnyel. Innen következik $t=1$ választással, hogy a limesz eloszlás karakterisztikus függvénye az $(a_1,\dots,a_k)$ pontban a $\varphi(a_1,\dots,a_k)=\exp\left\{-\sum\limits_{p=1}^k\sum\limits_{q=1}^l D(p,q)a_pa_q \Big/ 2\right\}$ függvény. E formula segítségével megkaphatjuk a több-dimenziós centrális határeloszlástételt, illetve a határeloszlások jellemzését, azaz a több-dimenziós normális eloszlás fogalmát. Mielőtt a részleteket kezdenénk el tárgyalni idézzünk fel néhány lineáris algebrai fogalmat és eredményt.

Lineáris algebrai ismétlés

Tekintsünk egy X (véges dimenziós) Euklides-i teret, azaz X lineáris tér, melyen létezik skalárszorzat. Az, hogy X lineáris tér azt jelenti, hogy ha $x\in X$, $y\in X$ X-beli elemek (melyeket az irodalomban vektornak neveznek) és α, β valós számok, akkor $\alpha x+\beta y\in X$ is teljesülnek, továbbá teljesülnek bizonyos algebrai azonosságok: $(x+y=y+x$, $x+(y+z)=(x+y)+z$, létezik $0\in X$, (null elem), melyre $x+0=x$, minden x pontnak létezik $-x$ inverze, melyre $x+(-x)=0$, $(\alpha+\beta)x=\alpha x+\beta x$, $(\alpha(\beta x)=(\alpha\beta)x$, $0x=0$, azaz, ha egy vektort beszorzunk a nulla számmal akkor a nulla vektort kapjuk.) Valamely $x_1,\dots,x_k\in X$ vektorokat lineárisan függetlennek nevezzük, ha a $\sum\limits_{j=1}^k\alpha_jx_j=0$ csak triviális módon lehetséges, azaz csak akkor, ha minden $\alpha_j\in R$ együtthatóra $\alpha_j=0$. Azt mondjuk, hogy az $x_1,\dots,x_k$ vektorok generátorrendszert alkotnak, ha minden $y \in X$ előállítható $y = \sum_{j=1}^{k} \alpha_k x_k$ alakban. Ha $x_1, \dots, x_k$ vektorok egyszerre alkotnak generátorrendszert és független rendszert, akkor ezt bázisnak hívják. Az általános eredmények szerint egyrészt minden vektor egyértelműen kifejezhetőek, mint egy bázis elemeinek lineáris lineáris kombinációja. Egy lineáris térnek különböző bázisai adhatók meg, de ezek elemszáma minden bázisra ugyanaz. Ezt a számot nevezik a lineáris tér dimenziójának.

A lineáris algebra egyik fontos fogalma a lineáris transzformáció fogalma. Egy $A\colon X \to X$ leképezést a lineáris transzformációnak nevezünk, ha miden $x \in X$, $y \in X$ vektorra α és β (komplex) számokra $(\alpha x + \beta y)A = \alpha(xA) + \beta(yA)$. (Az irodalomban nem egységes a jelölésrendszer, van ahol xA-t és van ahol Ax-et írnak.) Két lineáris transzformáció szorzatán a transzformációk egymás utáni alkalmazását értjük. Ha rögzítünk egy bázist egy k-dimenziós lineáris térben, akkor minden $x \in X$ vektort természetes módon azonosíthatunk egy szám-k-assal, ha a vektort felírjuk, mint a bázisbeli elemek (egyértelmű) lineáris kombinációját, és a vektort azonosítjuk az ebben a reprezentációban szereplő együtthatókkal. Ugyancsak természetes, (tanult) módon azonosíthatunk egy lineáris transzformációt egy mátrix-szal.

De valójában számunkra bizonyos extra-tulajdonságokkal rendelkező lineáris terek lesznek az érdekesek. Úgynevezett Euklidesi terekről szóló fogalmakra és eredményekre lesz szükségünk, azaz olyan lineáris terekkel foglalkozunk, melyekben be van vezetve egy skalárszorzat, azaz minden $x \in X$ és $y \in X$ vektorokra létezik (x, y) skalárszorzat, mely egy valós (illetve általánosabb esetekben komplex) szám, és teljesíti a következő feltételeket: $(x, x) \ge 0$, sőt $(x, x) > 0$, ha $x \ne 0$, $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$, $(x, y) = \overline{(y, x)}$, ahol $\overline{z}$ a z komplex szám konjugáltját jelöli. Ez a skalászorzat azért olyan fontos a számunkra, mert ez teszi lehetővé, hogy beszélhessünk egy $x \in X$ vektor $|x|$ hosszáról, melyet az $|x|^2 = (x, x)$ formula definiál, valamint két x és y vektor szögéről, (speciálisan merőlegességéről) melyet a $\cos(x \text{ és } y \text{ által bezárt szög}) = \frac{(x, y)}{|x||y|}$ formula definiál. Továbbá definiálhatjuk az úgynevezett $A(x, y)$ bilineáris formákat az $X \times X$ téren. $A(x, y)$ bilineáris forma, ha $A(\alpha_1 x + \alpha_2 x_2, y) = \alpha_1 A(x_1, y) + \alpha_2 A(x_2, y)$, és $A(x, \alpha_1 y_1 + \alpha_2 y_2) = \bar{\alpha}_1 A(x, y_1) + \bar{\alpha}_2 A(x_2, y_2)$. Be lehet látni, hogy egy Euklides-i térben egy természetes kölcsönösen egyértelmű megfeleltetés létezik a lineáris transzformációk és a bilineáris formák között. Nevezetesen, minden A lineáris transzformáció segítségével definiálható egy (xA, y) bilineáris forma, és tetszőleges bilineáris forma egyértelműen ilyen formában írható. Ez a következőképp lehetséges: Igaz az, hogy egy Euklides-i térben létezik $e_1, \dots, e_k$ ortonormált bázis. (Azaz létezik olyan vektorokból álló bázis, mely egymásra merőleges egy hosszúságú vektorokból áll. Ilyen bázis több is létezik. Bár nem lenne kötelező, de Euklidesi terekben gyakorlatilag mindig ortogonális bázis segítségével számolnak.) Egy ilyen bázisban az $A(x, y)$ bilineáris formát megadó (xA, y) kifejezésben szereplő A lineáris transzformáció mátrixa az a mátrix, melynek p-ik sorának q-ik helyén álló elem az $a(p, q) = A(e_p, e_q)$ szám. Egy ilyen koordinátarendszerben, ha $x = (x_1, \dots, x_k) \in X$, $y = (y_1, \dots, y_k) \in X$, akkor az $A(x, y)$ bilineáris formát a $A(x, y) = xAy^* = \sum_{p=1}^{k} \sum_{q=1}^{k} a(p, q) x_p \bar{y}_q$ képlet adja meg, ahol y^* az y

vektorból készített oszlopvektor.

Ha A lineáris transzformáció egy Euklidesi térben, akkor definiálhatjuk, annak transzponáltját az $(xA, y) = (x, yA^*)$ formulával. Ha a lineáris transzformáció mátrixát egy ortonormált bázisban írjuk fel, akkor annak transzponáltját átlóra való tükrözéssel és a mátrix (komplex értékű) elemeinek a konjugálásával kapjuk meg. Egy A lineáris transzformációt önadjungáltnak nevezünk, ha $A = A^*$, unitérnek, ha $UU^* = I$, ahol I az identitás mátrix. Egy unitér mátrixra az is érvényes, hogy $U^*U = I$. Egy unitér mátrix szög és távolságtartó, ezért ortonrmált bázist ortonormált bázisba visz. Egy A önadjungált mátrixra (xA, x) valós szám minden $x \in X$ számra, mert $(xA, x) = (x, xA) = \overline{(xA, x)}$. Azt mondjuk, hogy egy önadjungált mátrix pozitiv szemidefinit, ha $(xA, x) \geq 0$ minden $x \in X$ számra, pozitív definit, ha $(xA, x) > 0$ minden $x \neq 0$ vektorra. Nem nehéz belátni, hogy ha A önadjungált mátrix, akkor ezt a mátrixot illetve az általa definiált (xA, y) bilineáris függvényt meghatározza ennek megszorítása az $x = y$ pontokra, azaz az (xA, x) függvény, (kvadratikus alak).

Annak érdekében, hogy lineáris operátorok (a nekik megfelelő) mátrixát minél jobban át tudjuk tekinteni, érdemes olyan ortonormált rendszert találni, melyben az operátor mátrixa a lehető legegyszerűbb szerkezetű. Ezért fontos eredmény a következő állítás. Ha A önadjungált operátor, egy k-dimenziós Euklidesi térben, akkor létezik az A mátrixnak k darab $e_1, \dots, e_k$ ortonormált sajátvektorból álló bázisa, azaz olyan e_p, $1 \leq p \leq k$ vektorokból álló bázis, mely vektorokra $e_pA = \lambda_p e_p$ Továbbá λ_p valós szám, $p = 1, \dots, k$. Ebben a bázisban az A mátrix egy diagonális Λ mátrix. E mátrix diagonális elemei a λ_p sajátértékek. Az A önadjungált operátor akkor és csak akkor pozitív szemidefinit, ha minden sajátértéke nem-negatív. Megjegyezzük, hogy az A mátrix felírható, mint $A = U\Lambda U^*$ alakban, ahol U unitér operátor. Innen kiolvasható a következő számunkra fontos állítás: Ha A pozitív szemidefinit operátor (mátrix), akkor A felírható $A = BB^*$ alakban. (Sőt, választhatjuk a B mátrixot mint önadjungált mátrixot.) Ugyanis $B = U\sqrt{\Lambda}U^*$, ahol $\sqrt{\Lambda}$ az a diagonális mátrix, melynek elemei $\sqrt{\lambda_p}$, $1 \leq p \leq k$ teljesíti a feltételeket. Ekkor ugyanis $BB^* = \mathbf{U}\sqrt{\Lambda}U^*U\sqrt{\Lambda}U^* = U\Lambda U^* = A$.