Kiegészítés a Valószínűségszámítás I. előadássorozathoz

Áttekintés a felhasznált lineáris algebrai ismeretekről.

A valószínűségszámítás (és a matematika) bizonyos kérdéseiben fontos szerepet játszik a lineáris algebra néhány fogalma és eredménye. Egyrészt meg kell érteni az Euklideszi tereken értelmezett bilineáris függvények fogalmát, ezek kapcsolatát a lineáris transzformációkkal, mátrixokkal, illetve az ezen fogalmakhoz kapcsolódó legfontosabb eredményeket. Ezek az eredmények fontos szerepet játszanak a több-változós valószínűségi változók (különösen a több-dimenziós normális eloszlású valószínűségi változók) vizsgálatában. Egy másik fontos, a valószínűségi vizsgálatokban szintén használt, és a lineáris algebrához kapcsolódó eredmény a több-változós integrálok kiszámítását segítő, az integrálok értékét megőrző integráltranszformációk leírása. Ahhoz, hogy ezt az eredményt megértsük, először a determinánsok, pontosabban a mátrixok determinánsainak szemléletes tartalmát kell megértenünk. Az alábbiakban ezeket a fogalmakat és eredményeket tekintjük át.

Egy tipikus lineáris algebrai modell.

Először tekintsük a következő, a lineáris tereket és az ott megjelenő fogalmakat tartalmazó legjellemzőbb példát. Tekintsük az $(x_1,\dots,x_k)$ k hosszúságú sorozatokból álló teret, és definiáljuk két $(x_1,\dots,x_k)$ és $(y_1,\dots,y_k)$ sorozat összegét, mint $(x_1,\dots,x_k)+(y_1,\dots,y_k)=(x_1+y_1,\dots,x_k+y_k)$, és egy $(x_1,\dots,x_k)$ sorozat szorzatát egy α konstanssal mint $\alpha(x_1,\dots,x_k)=(\alpha x_1,\dots,\alpha x_k)$. A k-hosszúságú sorozatok terét a fenti műveletekkel lineáris térnek nevezzük, a tér elemeit, a k hosszúságú sorozatokat pedig (k-dimenziós) vektoroknak hívjuk. Jelöljük E_k-val a fent definiált lineáris teret. Ha bevezetjük két $x=(x_1,\dots,x_k)$ és $y=(y_1,\dots,y_k)$ E_k térbeli vektor skalárszorzatát is az $(x,y)=\sum\limits_{p=1}^{k} x_p y_p$ képlettel, akkor az E_k lineáris teret ezzel a skalárszorzattal együtt Euklideszi térnek nevezzük. A skalárszorzat bevezetése azért hasznos, mert ez lehetővé teszi, hogy definiáljuk egy vektor hosszát, két vektor által bezárt szöget, egy a k-dimenziós térben tekintett (szép) halmaz térfogatát, azaz, hogy felépítsük az E_k tér geometriáját.

A k-hosszú sorozatokból álló E_k lineáris térben definiálhatjuk a lineáris operátorokat és bilineáris függvényeket a következő módon: Egy $E_k\to E_k$ $z=A(x)$ transzformációt lineáris transzformációnak nevezünk, ha $A(\alpha x+\beta y)=\alpha A(x)+\beta A(y)$ minden $x\in E_k$, $y\in E_k$ vektorra valamint α és β számra. Egy $A(x,y)$, $x\in E_k$, $y\in E_k$ kétváltozós valós (vagy komplex) szám értékű függvény bilineáris függvény, ha teljesülnek az $A(\alpha x_1+\beta x_2,y)=\alpha A(x_1,y)+\beta A(x_2,y)$, és $A(x,\alpha y_1+\beta y_2)=\bar\alpha A(x,y_1)+\bar\beta A(x,y_2)$ azonosságok minden $x_1\in E_k$, $x_2\in E_k$, $y\in E_k$, vektorra, α, és β számra, illetve minden $y_1\in E_k$, $y_2\in E_k$, $x\in E_k$ vektorra és α és β számra, ahol $\bar z$ a z komplex szám konjugáltját jelöli. A következő módon tudunk definiálni lineáris transzformációkat: Legyen A egy $k\times k$ méretű mátrix. Akkor $A(x)=xA$, $x\in E_k$, lineáris transzformáció, ahol xA az x vektor és A mátrix szokásos szorzatát jelöli. Hasonlóan egy A $k\times k$ méretű mátrix segítségével definiálhatunk egy bilineáris függvényt, ha az E_k térben bevezetjük a skalárszorzatot is, azaz E_k-t nemcsak lineáris, hanem Euklideszi térnek is tekintjük. Ezt a következő módon tehetjük: Legyen $A(x,y) = (xA, y)$, ahol xA vektor és mátrix szokásos szorzatát, $(\cdot,\cdot)$ pedig a skalárszorzatot jelöli. Az általános elméletből következik, hogy az E_k téren minden lineáris transzformációt illetve bilineáris függvényt meg lehet adni ilyen módon egy alkalmas A $k \times k$ méretű mátrix segítségével. Megjegyzem, hogy a fenti fogalmakat és eredményeket természetes módon lehet általánosítani arra az esetre is, amikor az E_k térnek egy E_l térbe való lineáris leképezéseit akarjuk tekinteni, illetve olyan $A(x,y)$ bilineáris függvényeket akarunk definiálni, melyekre $x \in E_k$, $y \in E_l$, és a k illetve l indexek különbözőek is lehetnek. Viszont a minket érdeklő problémák vizsgálatában elegendő csak a $k = l$ esettel foglalkozni.

A lineáris algebra néhány fontos fogalma és kérdése.

Megadjuk a lineáris terek, Euklideszi terek és a rajtuk értelmezett lineáris transzformációk, bilineáris függvények általános, ,,absztrakt" definicióját, illetve a velük kapcsolatos legfontosabb eredményeket. Bár ez a definició formálisan általánosabb mint az előbb tekintett példákban leírt eset, valójában be lehet látni, hogy tetszőleges lineáris tér, illetve Euklideszi tér és a rajtuk definiált lineáris transzformációk illetve biliniáris függvények izomorfak a fenti példában tekintett modellek valamelyikével. Mégis érdemes az ,,absztrakt" definiciót bevezetni. Bizonyos problémákat egyszerűbb és természetesebb módon lehet vizsgálni, ha az általános modellt tekintjük, és annak strukturáját jobban megértjük.

Vezessük be először a lineáris tér fogalmát. Egy X halmazt lineáris térnek neveznek, ha minden $x \in X$ és $y \in X$ elemre, (melyeket az irodalomban vektornak neveznek) és α, valós (komplex) számra, definiálva van az $x + y \in X$ összeg és az $\alpha x \in X$ konstanssal képzett szorzat. Továbbá megköveteljük, hogy ezek a műveletek teljesítsék a következő algebrai azonosságokat: $x + y = y + x$, $x + (y + z) = (x + y) + z$, létezik $0 \in X$, (a lineáris tér null eleme), 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 függetleneknek nevezünk, ha a $\sum\limits_{j=1}^{k} \alpha_j x_j = 0$ egyenlet csak triviális módon teljesülhet, azaz csak akkor, ha mindegyik α_j számra (együtthatóra) $\alpha_j = 0$. Azt mondjuk, hogy az $x_1, \dots, x_k$ vektorok generátorrendszert alkotnak, ha minden $y \in X$ előállítható $y = \sum\limits_{j=1}^{k} \alpha_k x_k$ alakban. Ha az $x_1, \dots, x_k$ vektorok egyrészt generátorrendszert alkotnak, másrészt lineárisan függetlenek, akkor ezen vektorok rendszerét bázisnak hívják. A lineáris algebra néhány egyszerű eredménye kifejezi a bázisok fontos tulajdonságait. Ezen eredmények egyike szerint minden vektor *egyértelműen* fejezhető ki, mint egy bázis elemeinek lineáris kombinációja. Egy lineáris térnek különböző bázisai léteznek. Viszont egy lineáris tér minden bázisa ugyanannyi elemből áll. Egy lineáris tér bázisainak elemszámát hívják a tér dimenziójának. Mi csak azzal az esettel fogunk foglalkozni, amikor a lineáris tér dimenziója egy véges szám.

A lineáris algebra egyik fontos fogalma a lineáris transzformáció. Egy X lineáris tér $A\colon X \to X$ önmagába való leképezését lineáris transzformációnak nevezzük, ha minden $x \in X$ és $y \in X$ vektorra α és β (komplex) számra $(\alpha x + \beta y)A = \alpha(xA) + \beta(yA)$. (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, ami szintén lineáris transzformáció.

Sok fontos vizsgálat elvégzése érdekében érdemes bevezetni bizonyos extra-tulajdonságokkal rendelkező lineáris tereket, melyeket Euklideszi tereknek hívnak. Egy X lineáris tér Euklideszi tér, ha értelmezve van benne egy skalárszorzat, azaz, ha minden $x \in X$ és $y \in X$ vektorra létezik egy az x és y skalárszorzatának nevezett (x, y) mennyiség, mely valós (illetve általánosabb esetekben komplex) szám, és teljesíti a következő feltételeket: $(x, x) \ge 0$, sőt $(x, x) > 0$, ha $x \neq 0$, $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$, $(x, y) = \overline{(y, x)}$, ahol $\bar{z}$ a z komplex szám konjugáltját jelöli. A skalászorzat lehetővé, teszi, hogy egy X Euklideszi térben 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 $\cos(x$ és y által bezárt szög$) = \frac{(x,y)}{|x||y|}$ formula definiál. Az $x \in X$ és $y \in X$ vektorokat merőlegeseknek vagy más szóval ortogonálisoknak nevezzük, ha $(x, y) = 0$.

Definiálhatunk úgynevezett $A(x, y)$ bilineáris függvényeket is egy X Euklideszi téren. Az $A(x, y)$, $x \in X$, $y \in X$, valós vagy komplex értékű függvény bilineáris függvény, ha $A(\alpha_1 x + \alpha_2 x_2, y) = \alpha_1 A(x_1, y) + \alpha_2 A(x_2, y)$, és $A(x, \alpha_1 y_1 + \alpha_2 y_2) = \bar{\alpha}_1 A(x, y_1) + \bar{\alpha}_2 A(x_2, y_2)$.

Rögzítsük egy X k-dimenziós X lineáris tér vagy speciálisan Euklideszi tér egy $e_1, \dots, e_k$ bázisát. Ha ismerjük egy A lineáris transzformáció képét mindegyik e_j, $1 \le j \le k$, bázisvektorra, azaz ismerjük az összes $a_{j,p}$, $1 \le j, p \le k$, együtthatót az $e_j A = \sum\limits_{p=1}^{k} a_{j,p} e_p$, $1 \le j \le k$, egyenletekben, akkor tetszőleges x vektor xA képét ki tudjuk számolni. A lineáris algebra természetes és fontos problémái a következő kérdések megválaszolása: Hogyan lehet ezt az xA vektort effektive kiszámolni, milyen számolásokat kell végrehajtani, ha egy $e_1, \dots, e_k$ bázis helyett egy másik $\bar{e}_1, \dots, \bar{e}_k$ bázisra térünk át, és ott akarjuk kiszámolni az A lineáris transzformáció hatását? Mely bázis választással tudjuk a legegyszerűbben látni egy A lineáris transzformáció viselkedését?

Hasonló kérdések fogalmazhatóak meg egy az X Euklideszi térben definiált $A(\cdot, \cdot)$ bilineáris függvény esetében. Be lehet látni, hogy egy k-dimenziós Euklideszi térben nemcsak bázis, hanem speciális, úgynevezett ortonormált bázis is létezik, azaz meg lehet adni, olyan $e_1, \dots, e_k$ bázist, melyre $(e_j, e_p) = 0$, ha $j \neq p$, $(e_j, e_j) = 1$, $1 \le j, p \le k$. Bár nem lenne kötelező, de ha Euklideszi térben bilineáris függvényeket vizsgálnak mindig ortonormált bázisban dolgoznak. Ha rögzítünk egy $e_1, \dots, e_k$ ortonormált bázist, akkor az $A(e_j, e_p)$, $1 \le j, p \le k$, értékek segítségével ki lehet számítani tetszőleges $x \in X$ és $y \in X$ vektorra az $A(x, y)$ bilineáris függvény értékét. Itt is felmerül a kérdés, hogyan lehet ezt a számolást effektive elvégezni, hogyan tudjuk kiszámítani a bilineáris függvényt, ha egy ortonormált bázisról áttérünk egy másikra, és melyik ortonormált bázisban tudjuk a bilineáris függvény viselkedését a legegyszerűbben látni.

Ha rögzítjük egy X k-dimenziós lineáris tér egy $e_1,\dots,e_k$ bázisát, és tekintünk ezen a téren egy A lineáris transzformációt, melyre $e_jA=\sum\limits_{p=1}^k a_{j,p}e_p$, $1\le j\le k$, valamilyen $a_{j,p}$, $1\le j,p\le k$, együtthatókkal, akkor érdemes a következő jelöléseket bevezetni: Definiáljuk azt az $\bar A$ $k\times k$ méretű mátrixot, melynek a j-ik sorában és p-ik oszlopában álló elem a fenti azonosságban szereplő $a_{j,p}$ szám. Ezt az $\bar A$ mátrixot nevezzük nevezzük az A lineáris transzformáció mátrixának (a rögzített $e_1,\dots,e_k$ bázis esetében). Feleltessük meg továbbá az $x=\sum\limits_{j=1} x_je_j$ vektornak az $\bar x=(x_1,\dots,x_k)$ szám k-ast. Be lehet látni, hogy ha az x vektornak az $\bar x$ szám k-as, akkor az xA vektornak az $\bar x\bar A$ szám k-as felel meg a fenti megfeleltetésben. Továbbá az A lineáris transzformáció mátrixa $\bar A$, a B lineáris transzformáció mátrixa $\bar B$, akkor az AB lineáris transzformáció mátrixa az $\bar A\bar B$ mátrix, ahol AB a két lineáris transzformáció szorzatát (azaz egymás után való alkalmazását) $\bar A\bar B$ pedig a két mátrix (szokásos értelemben vett) szorzatát jelöli. A fenti $A\to\bar A$ megfeleltetés kölcsönösen egyértelmű megfeleltetést ad az X k-dimenziós lineáris tér lineáris leképezései és a $k\times k$ méretű mátrixok között (rögzített $e_1,\dots,e_k$ bázis esetében).

Hasonló eredmény érvényes egy Euklideszi térben definiált bilineáris függvényekre. Be lehet látni, hogy egy X Euklideszi térben egy természetes kölcsönösen egyértelmű megfeleltetés létezik a lineáris transzformációk és a bilineáris függvények között. Nevezetesen, egy A lineáris transzformáció és az $(\cdot,\cdot)$ skalárszorzat segítségével definiálható az $A(x,y)=(xA,y)$ kifejezés, és ez bilineáris függvény. Megfordítva, tetszőleges $A(\cdot,\cdot)$ bilineáris függvény egyértelműen felírható ilyen formában egy alkalmas A lineáris transzformáció segítségével. Legegyszerűbb ezt a megfeleltetést úgy megadni, hogy rögzítünk egy $e_1,\dots,e_k$ ortonormált bázist az Euklideszi térben, és definiáljuk az $A(\cdot,\cdot)$ bilineáris függvény valamint az $e_1,\dots,e_k$ bázis és $(\cdot,\cdot)$ skalárszorzat segítségével azt az $\bar A=\bar A(e_1,\dots,e_k)$ mátrixot, melynek j-ik sorában és p-ik oszlopában az $a_{j,p}=A(e_j,e_p)$ szám áll. Ekkor $A(x,y)=\bar x\bar A\bar y^*$ minden $x\in X$, $y\in X$ vektorra, ahol $x=\sum\limits_{j=1}^k x_je_j$ $y=\sum\limits_{j=1}^k y_je_j$, $\bar x=(x_1,\dots,x_k)$ $\bar y=(y_1,\dots,y_k)$, és $\bar y^*$ a $\bar y$ vektor transzponáltját jelöli. Természetesen az előbb definiált $\bar A$ mátrix függ attól, hogy melyik ortonormált bázist választottuk, de be lehet látni, hogy az a lineáris transzformáció, melyet ez a mátrix meghatároz már nem függ az (ortonormált) bázis megválasztásától. A fenti megfeleltetések egyben azt is mutatják, hogy hogyan lehet egy általános lineáris transzformációnak illetve bilineáris függvénynak rögzített (bilineáris függvény esetében az Euklideszi térben ortonormált) bázis esetében egy olyan speciális modellt izomorf módon megfeleltetni, mint amilyet az ismertetés elején megadtam.

Néhány eredmény lináris transzformációkról.

Tekintsük át először röviden azt, hogy hogyan lehet egy A lineáris transzformációt egyszerűen megadni egy alkalmas bázisban. Az ilyen vizsgálatokban rendkívül hasznos fogalom egy lineáris operátor sajátvektorának és a sajátvektor sajátértékének a fogalma. Egy e vektor az A operátor sajátvektora λ sajátértékkel, ha teljesül az $eA = \lambda e$ azonosság. Felmerül a kérdés, hogy van-e minden operátornak olyan bázisa, amelyiknek minden eleme sajátvektor. Ha az A lineáris transzformációnak van ilyen bázisa, akkor érdemes az A transzformáció mátrixát ebben a bázisban felírni. Az A transzformáció mátrixa egy ilyen bázisban diagonális mátrix, azaz a mátrixban a főátlón kívül mindenütt nulla van. Továbbá a mátrix diagonálisában az A transzformáció sajátértékei állnak. De nem minden lineáris operátornak van ilyen bázisa. Az általános esetben egy lineáris leképezés a legegyszerűbb módon az úgynevezett Jordan féle normálalakban adható meg. A Jordan féle normálalak létezése nagyon érdekes eredmény, amelyik sok matematikai problémában, például a (többváltozós lineáris) differenciálegyenletek elméletében nagyon hasznos. Egy A lineáris leképezés Jordan féle normálakjának előállításában fontos lépés az A lineáris operátor invariáns altereinek a megtalálása, azaz az olyan $D \subset E$ lineáris tereké, melyekre $x \in D$ esetében az $xA \in D$. Az A lineáris operátor egydimenziós invariáns alterei, megegyeznek a $D = \{\alpha e : \alpha \text{ tetszőleges szám}\}$ alakú egydimenziós alterekkel, ahol e az A operátor sajátvektora. De mivel a minket érdeklő kérdések vizsgálatában a Jordan féle normálalak nem jelenik meg, ezért ennek a kérdésnek a tárgyalását elhagyom. Hasonló okokból nem tárgyalom, hogyan változik egy lineáris transzformáció mátrixa, ha egy bázisból egy másik bázisra térünk át. Csak röviden megadom az eredményt:

Legyen $e_1, \dots, e_k$ és $\bar{e}_1, \dots \bar{e}_k$ két bázis egy E lineáris térben. Ekkor létezik egy egyértelműen meghatározott, és invertálható B lineáris transzformáció, melyre $e_j B = \bar{e}_j$, $1 \le j \le k$. A B transzformáció $\bar{B}$ mátrixa megegyezik az $e_1, \dots, e_k$ és $\bar{e}_1, \dots \bar{e}_k$ bázisokban, és ugyanez az állítás érvényes a B mátrix B^{-1} inverzének a $\bar{B}^{-1}$ mátrixára is. Be lehet látni, hogy ha egy A lineáris transzformáció mátrixa az $\bar{A}$ mátrix az $e_1, \dots, e_k$ bázisban, akkor az A transzformáció mátrixa az $\bar{e}_1, \dots \bar{e}_k$ bázisban a $\bar{B}\bar{A}\bar{B}^{-1}$ mátrix.

Bilineáris függvények néhány fontos tulajdonsága. Önadjungált és unitér leképezések.

A mi vizsgálatainkban fontosabb szerepet játszik egy Euklideszi térben definiált bilineáris függvények tulajdonságainak a megértése. Mint korábban írtam, egy X Euklideszi térben minden $A(\cdot,\cdot)$ bilineáris függvény felírható egyértelműen $A(x,y) = (xA, y)$ alakban, ahol A egy az X Euklideszi téren definiált lineáris transzformáció, továbbá minden A lineáris transzformáció meghatároz ezen a módon egy bilineáris függvényt. Ezért egy $A(\cdot,\cdot)$ bilineáris függvényt azonosíthatunk az őt meghatározó A lineáris transzformációval. Ez a megjegyzés lehetővé teszi, hogy definiáljuk egy X Euklideszi térben megadott A lineáris transzformáció A^* adjungáltját a következő módon: Az A^* lineáris transzformációt meghatározza az $(xA, y) = (x, yA^*)$ azonosság teljesülése minden $x \in X$ és $y \in X$ vektorra. Ha egy A lineáris transzformáció mátrixát egy ortonormált bázisban írjuk fel, akkor könnyen megadhatjuk az A transzformáció A^* adjungáltjának a mátrixát ugyanabban a bázisban. Nevezetesen a A^* adjungált operátor mátrixát úgy kapjuk meg, hogy az A operátor mátrixának az elemeit az átlóra tükrözzük, és ha komplex értékű akkor konjugáljuk. Azaz ha $A = (a_{j,p})$, $1 \le j, p \le k$, akkor $A^* = (\bar{a}_{p,j})$, $1 \le j, p \le k$. Ez egyben azt is jelenti, hogy beszélhetünk egy A mátrix adjungáltjáról is, azaz noha ugyanaz a mátrix több különböző lineáris transzformáció mátrixaként jelenhet meg, attól függően, hogy mely (ortonormált) bázisban dolgozunk, az A mátrix A^* adjungáltja mindig ugyanaz az A^* mátrix. Felírom az $A \to A^*$ leképezéssel kapcsolatos legfontosabb azonosságokat. Ezek: $(A+B)^* = A^* + B^*$, $(\alpha A)^* = \bar{\alpha} A^*$, $(A^*)^* = A$, $(AB)^* = B^* A^*$, (tehát az utolsó azonosság jobboldalán fel kell cserélni a tényezőket), és $\left(A^{-(1}\right)^* = (A^*)^{-1}$.

Az Euklideszi téren értelmezett lineáris transzformációk között különösen fontos szerepet játszanak az önadjungált és unitér leképezések. Leírom ezek definicióját és a velük kapcsolatos legfontosabb tudnivalókat.

Az X Euklideszi tér egy A lineáris transzformációját önadjungáltnak nevezünk, ha $A = A^*$, egy U transzformációját unitérnek, ha $UU^* = U^*U = I$, ahol I az identitás leképezés. Az, hogy az A lineáris transzformáció önadjungált úgy is megfogalmazható, hogy az általa meghatározott $A(\cdot,\cdot)$ bilineáris függvény teljesíti az $A(x,y) = \overline{A(y,x)}$ azonosságot. Az unitér transzformációk egyik fontos tulajdonsága az, hogy valójában elég az $UU^* = I$ vagy $U^*U = I$ azonosságok egyikét megkövetelni, a másik azonosság ebből következik. Egy k-dimenziós Euklideszi tér valamely transzformációja akkor és csak akkor unitér, ha rögzítve az Euklideszi tér egy ortonormált bázisát és felírva a transzformáció mátrixát ebben a bázisban az így kapott mátrix sorai (vagy oszlopai) ortonormált rendszert alkotnak a szám k-asok terében. Az unitér transzformációknak fontos geometriai jelentésük van. Ezek az X Euklideszi tér távolság és szögtartó leképezései. Speciálisan, az unitér transzformációk ortonormált bázist ortonormált bázisba képeznek.

Egy önadjungált lineáris leképezés mátrixát valamilyen rögzített ortonormált bázisban önadjungált mátrixnak, egy unitér lineáris leképezés mátrixát pedig unitér mátrixnak hívják. Ez az elnevezés jogos, mert az, hogy egy mátrix önadjungált vagy unitér-e eldönthető csak a mátrix ismeretében, és nem függ attól, hogy milyen módon kaptuk ezt a mátrixot egy alkalmas operátor segítségével. Nevezetesen, egy $A = (a_{j,p})$, $1 \le j,p \le k$, $k \times k$ méretű mátrix akkor és csak akkor önadjungált, ha $a_{j,p} = \bar{a}_{p,j}$ minden $1 \le j,p \le k$ indexre. Egy $U = (u_{j,p})$, $1 \le j,p \le k$, $k \times k$ méretű mátrix akkor és csak akkor unitér, ha az U mátrix sorai ortonormáltak, azaz $\sum\limits_{p=1}^{k} u_{j,p}\bar{u}_{j,p} = 1$, minden $1 \le j \le k$ indexre, és $\sum\limits_{p=1}^{k} u_{j,p}\bar{u}_{l,p} = 0$, minden $1 \le j,l \le k$ indexre, ha $j \ne l$.

Egy A önadjungált leképezésre (xA,x) valós szám minden $x \in X$ vektorra, mert $(xA,x) = (x,xA) = \overline{(xA,x)}$. Azt mondjuk, hogy egy önadjungált leképezés pozitív szemidefinit, ha $(xA,x) \ge 0$ minden $x \in X$ vektorra, pozitív definit, ha $(xA,x) > 0$ minden $x \ne 0$ vektorra. Nem nehéz belátni, hogy ha A önadjungált transzformáció, akkor ezt a transzformációt illetve az általa definiált (xA,y) bilineáris függvényt meghatározza e fügvény megszorítása az $x = y$ pontokra, azaz az (xA,x) függvény, melyet kvadratikus alaknak hívnak az irodalomban.

A lineáris algebra egyik fontos eredménye arról szól, hogy hogyan lehet önadjungált operátorokat a legegyszerűbb módon jellemezni egy alkalmas ortonormált bázis segítségével. A következő állítás érvényes. Ha A önadjungált operátor egy X k-dimenziós Euklideszi térben akkor létezik az A transzformációnak k darab $e_1, \dots, e_k$ ortonormált sajátvektorból álló bázisa, azaz olyan egymásra merőleges (egy hosszúságú) e_p, $1 \le p \le k$ vektorokból álló bázis, mely vektorokra $e_p A = \lambda_p e_p$ valamely λ_p sajátértékkel. Továbbá a λ_p sajátértékek valós számok minden $p = 1, \dots, k$ indexre. Ebben a sajátvektorokból álló bázisban az A transzformáció mátrixa egy diagonális Λ mátrix, és a Λ mátrix diagonális elemei a λ_p sajátértékek. Egy A önadjungált operátor akkor és csak akkor pozitív szemidefinit, ha minden sajátértéke nem-negatív, és akkor és csak akkor pozitív definit, ha mindegyik sajátértéke szigorúan pozitív.

A fenti állítás arról, hogy egy X k-dimenziós Euklideszi térben definiált A önadjungált operátornak létezik $e_1, \dots, e_k$ sajátvektorokból álló ortonormált bázisa, a következő jellemzését adja az A operátornak mátrix általános ortonormált bázisban. Legyen $f_1, \dots, f_k$ tetszőleges ortonormált bázis. Ekkor az a sajátértékekből álló ortonormált bázis elemei kifejezhetőek $e_j = \sum\limits_{p=1}^{k} u_{j,p} f_p$, $j = 1, \dots, k$, alakban. Továbbá az $U = (u_{j,p})$, $1 \le j, p \le k$, mátrix unitér. Ugyanis abból, hogy mind az $e_1, \dots, e_k$ mind az $f_1, \dots, f_k$ bázis ortonormált rendszert alkot következik, hogy az U mátrix sorai ortonormált rendszert alkotnak a szám k-asok terében. Be lehet látni, hogy az A önadjungált operátor mátrixa az $f_1, \dots, f_k$ bázisban az $U^* \Lambda U$ mátrix, ahol U az előbb definiált unitér mátrix, Λ pedig az a diagonális mátrix, melynek p-ik sorában (és p-ik oszlopában) az e_p sajátvektor λ_p sajátértéke áll. Rejtett módon egy transzformáció A mátrixának $A = U^* \Lambda U$ alakú előállítása jelenti azt az eredményt, melyet az irodalomban főtengely transzformációnak neveznek. Ez azt mondja ki, hogy egy önadjungált operátor mátrixa alkalmas ortonormált bázisban diagonális. Ahhoz, hogy ezt az előállítást megadjuk az operátor sajátvektorait és sajátértékeit kell ismernünk.

Röviden ismertetem, hogy felhasználva azt az eredményt mely szerint minden A önadjungált operátornak létezik az A operátor ortogonális sajátvektoraiból álló bázisa, hogyan lehet látni azt, hogy az A operátor mátrixa $U^* \Lambda U$ alakban írható. Legyen az e_j sajátvektor előállítása a vizsgált $f_1, \dots, f_k$ ortonormált bázisban $e_j = \sum\limits_{p=1}^{k} u_{j,p} f_p$ alakú, és vezessük be az $u^{(j)} = (u_{j,1} \dots, u_{j,k})$, $j = 1, \dots, k$ jelölést. Ekkor elég belátni, hogy teljesül az $(e_j A, e_l) = u^{(j)} U^* \Lambda U u^{(l)^*}$ azonosság minden $1 \le j, l \le k$ indexre. Ezt az azonosságot viszont nem nehéz ellenőrizni, felhasználva a következő azonosságokat: Egyrészt $(e_j A, e_j) = \lambda_j$, és $(e_j A, e_l) = 0$, ha $j \ne l$, $1 \le j, l \le k$. Másrészt $u^{(j)} U^*$ az a k hosszúságú számsorozat, melynek j-ik koordinátája 1, és az összes többi koordinátája nulla, és az $U u^{(l)^*}$ vektor az a k hosszúságú oszlopvektor, melynek l-ik koordinátája 1, és az összes többi koordinátája nulla.

Jegyezzük meg, hogy érvényes és hasonlóan bizonyítható a következő állítás: Ha egy X k-dimenziós Euklideszi térben definiált (mem feltétlenül önadjungált) bilineáris függvény mátrixa egy B mátrix valamilyen $e_1, \dots, e_k$ ortonormált bázisban, és tekintünk egy másik ortonormált $f_1, \dots, f_k$ bázist az X térben, akkor felírhatjuk az $e_j = \sum\limits_{p=1}^{k} u_{j,p} f_p$ azonosságokat, $1 \le j \le k$, alkalmas $u_{j,p}$ együtthatókkal, és az $U = (u_{j,p})$, $1 \le j, p \le k$ $k \times k$ méretű mátrix unitér. Továbbá a tekintett bilineáris függvény mátrixa az új

$f_1, \dots, f_k$ bázisban az U^*BU mátrix. Jegyezzük meg, hogy az $U^* = U^{-1}$ mátrix az U mátrix inverze. A most említett formula alapján egy önadjungált operátor felírása $U^*\Lambda U$ alakban informálisan a következő módon interpretálható: Ez a képlet azt mutatja meg, hogy hogyan „látjuk" egy önadjungált operátornak a sajátvektorok által meghatározott bázisban megjelenő diagonális alakját egy másik ortonormált bázisból.

Természetesen felmerül az a kérdés, hogy van-e hasonló szép, alkalmas koordinátarendszerben diagonális mátrix formában megadható előállítása az unitér transzformációknak. Ezzel a kérdéssel ekvivalens az a probléma, hogy létezik-e minden unitér operátornak ortogonális sajátvektorokból álló bázisa. A válasz erre a kérdésre igenlő. A különbség mindössze annyi, hogy az önadjungált operátorok sajátértékei valós számok, míg az unitér operátorok sajátértékei egy abszolut értékű komplex számok. Érdemes megjegyezni, hogy a fent említett állítás az unitér operátorokról csak akkor érvényes, ha nem valós hanem úgynevezett komplex Euklideszi terekben dolgozunk, azaz az Euklideszi tér elemeit nemcsak valós, hanem komplex számokkal is megszorozhatjuk. Szintén ismert az, hogy hogyan lehet valós Euklideszi tereken értelmezett unitér transzformációk mátrixát egyszerű módon előállítani alkalmas bázisban. Ennek az eredménynek érdekes geometriai következményei vannak, de ezt a kérdést itt nem tárgyalom. Egyébként ismert az is, hogy melyek azok az operátorok, melyeknek létezik sajátvektorokból álló ortonormált bázisa. Ezek az úgynevezett normális operátorok. Egy operátor normális, ha $AA^* = A^*A$.

A lineáris algebra alkalmazása valószínűségszámítási problémákban.

Tekintsük az ismertetés elején tárgyalt k hosszúságú $(x_1, \dots, x_k)$ sorozatokból álló E_k Euklideszi teret az $(x, y) = \sum\limits_{p=1}^{k} x_p \bar{y}_p$ skalárszorzattal. Legyen továbbá adva k $\xi_1, \dots, \xi_k$ valószínűségi változó egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyekre $E\xi_p = 0$, $E\xi_p^2 < \infty$, $1 \le p \le k$. E valószínűségi változók segítségével definiáljuk a következő bilineáris függvényt az E_k Euklideszi téren. Ha $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k)$ akkor legyen $A(x, y) = E\left(\sum\limits_{j=1}^{k} x_j \xi_j\right)\left(\sum\limits_{l=1}^{k} y_l \xi_l\right)$. Nem nehéz belátni, hogy ez a bilineáris függvény megadható $A(x, y) = xDy^*$ alakban, ahol $D = d_{j,l}$, $1 \le j, l \le k$, $d_{j,l} = E\xi_j \xi_l = \operatorname{Cov}(\xi_j, \xi_l)$, a $(\xi_1, \dots, \xi_k)$ véletlen vektor kovariancia mátrixa. Ennek a bilineáris függvénynek, illetve a neki megfelelő kovarianciamátrixnak a vizsgálata fontos szerepet játszott a több-dimenziós normális eloszlás, illetve a több-dimenziós centrális határeloszlástétel vizsgálatában. Láttuk, hogy a D kovarianicia mátrix önadjungált, és pozitív szemidefinit. Megjegyeztük, hogy a lineáris algebra eredményeiből következik, hogy minden A pozitív szemidefinit mátrix felírható $A = B^*B$ alakban. Az előadásban láttuk, hogy ennek az eredménynek az az egyik következménye, hogy minden szimmetrikus pozitív szemidefinit mátrix esetében létezik olyan több-dimenziós normális eloszlású valószínűségi változó, amelyiknek ez a kovariancia mártixa.

Lássuk, hogy miért igaz ez az állítás. Egy pozitív szemidefinit diagonális Λ mátrix esetében ez az állítás nagyon egyszerűen bizonyítható. Legyenek $\lambda_1, \dots, \lambda_k$ a Λ mátrix átlójában lévő elemek. Az, hogy a Λ mátrix pozitív szemidefinit, azt jelenti, hogy $\lambda_j \ge 0$

minden $1 \leq j \leq k$ indexre. Ezért $\Lambda = \sqrt{\Lambda}\sqrt{\Lambda} = \left(\sqrt{\Lambda}\right)^* \sqrt{\Lambda}$, ahol $\sqrt{\Lambda}$ az a diagonális mátrix, melynek elemei a $\sqrt{\lambda_j}$, $1 \leq j \leq k$ számok. Az általános esetben egy A szimmetrikus mátrix felírható $A = U^* \Lambda U$ alakban, ahol U unitér, Λ pedig diagonális mátrix. Az, hogy az A mátrix pozitív szemidefinit azt jelenti, hogy a Λ mátrix diagonálisában álló elemek (az A mátrix sajátvektoraihoz tartozó sajátértékek) nem negatív számok. Ezért teljesül az $A = B^2 = B^* B$ azonosság $B = U^* \sqrt{\Lambda} U$ választással. Ugyanis ekkor $B^* B = U^* \sqrt{\Lambda} U U^* \sqrt{\Lambda} U = U^* \Lambda U = A$, és $B = B^*$. (Ezenkívül azt is biztosítottuk, hogy a fenti előállításban szereplő B mátrix pozitív szemidefinit.) Az $A = B^* B$ előállítás nem egyértelmű. Valóban, legyen V tetszőleges unitér mátrix, és definiáljuk a $\bar{B} = VB$ mátrixot. Ekkor $\bar{B}^* \bar{B} = B^* V^* V B = B^* B = A$, ha teljesül az $A = B^* B$ azonosság.

Mátrixok determinánsa, Egy transzformáció Jacobianja, Többváltozós integrálok transzformációi.

Tekintsük egy A $k \times k$ méretű $A = (a_{j,l})$, $1 \leq j, l \leq k$, alakú mátrix $\det A$ determinánsát. Ennek a determinánsnak a következő szemléletes jelentése van: Tekintsük az A mátrix $a^{(j)} = (a_{j,1}, \dots, a_{j,k})$ sorait, $1 \leq j \leq k$, mint vektorokat a k-dimenziós térben. Ekkor $\det A$ megegyezik az $a^{(j)}$, $1 \leq j \leq k$, vektorok által kifeszített parallelepipedon előjeles térfogatával. Az A mátrix sorvektorai által kifeszített parallelepipedon a k-dimenziós egységkocka képe, ha az A mátrix által meghatározott lineáris transzformációt alkalmazzuk. Ezért a fent említett eredmény szerint a $\det A$ geometriai tartalma azt fejezi ki, hogy az A mátrix hányszorosára nagyítja egy k-dimenziós vektorok által kifeszített parallelepipedon előjeles térfogatát. Ennek a geometriai ténynek fontos következményei vannak. Ez az eredmény magyarázza meg a $\det(AB) = \det A \det B$ azonosság okát és szemléletes tartalmát. Nevezetesen, $\det(AB)$ azt méri, hogy hányszorosára nagyítja az előjeles térfogatot az A és B lineáris transzformációk egymás utáni alkalmazása, míg az azonosság jobboldala azt méri, hogy az egyes A és B transzformációk külön-külön hányszorosra nagyítják az előjeles térfogatot. Sőt, ennek a geometriai képnek más számunkra fontos következménye is van. Ez magyarázza meg azt, hogy hogyan kell kiszámítani többváltozós függvények integrálját, ha a függvényt transzformáljuk, és miért jelenik meg ebben a formulában a leképezés Jacobian-jának nevezett mennyiség.

Magának a fent említett ténynek a bizonyítását elhagyom, csak röviden utalok az eredmény okára. Ha tekintünk k darab k-dimenziós vektort, akkor mind az általuk kifeszített parallelepipedon előjeles térfogata, mind azon mátrix determinánsának az értéke, mely mátrix j-ik sorában az $a^{(j)}$ vektor áll, az $a^{(j)}$ vektorok multilineáris függvénye. Azaz, bárhogy rögzítjük egy vektor kivételével az összes többi vektor értékét, akkor az így kapott függvény a nem rögzített vektorváltozó lineáris függvénye. Annak belátásához, hogy két multilineáris függvény megegyezik elegendő belátni a két függvény azonosságát bizonyos speciális esetekben. Meg lehet mutatni, hogy jelen esetben elegendő belátni azt, hogy a tekintett előjeles térfogat és determináns értéke megegyezik azokban a speciális esetekben, amikor mindegyik $a^{(j)}$ vektor olyan, hogy az egyik koordinátája 1, az összes többi koordinátája pedig nulla. Ezekben az esetekben viszont nem nehéz ellenőrizni az állítást.

Tekintsünk egy $f(y_1, \dots, y_k)$ függvényt, illetve annak $\int_B f(y_1, \dots, y_k)\, dy_1 \dots dy_k$ integrálját valamely B tartományban. Arra vagyunk kiváncsiak, hogy hogyan lehet kiszámolni ezt az integrált, ha a k-dimenziós tér A halmazának egy $y_j = T_j(x_1, \dots, x_k)$, $1 \le j \le k$, leképezését alkalmazzuk erre a B tartományra. Ezt a leképezést röviden az $y = T(x)$ formában írjuk, ahol az $y = (y_1, \dots, y_k)$, $x = (x_1, \dots, x_k)$, $T = (T_1, \dots, T_k)$ jelölést alkalmazzuk. Tulajdonképpen az

$$\int_a^b f(g(x))g'(x)\, dx = \int_{g(a)}^{g(b)} f(y)\, dy$$

azonosság többdimenziós változatát keressük. Tegyük fel, hogy a további érvelésekben az összes előforduló függvény és halmaz elég szép, azok minden számunkra szükséges tulajdonsággal rendelkeznek. Továbbá, ha másképp nem mondjuk feltételezzük azt is, hogy a k-dimenziós tér T transzformációja az A k-dimenziós tartománynak invertálható leképezése a k-dimenziós egy B tartományára, azaz az $y = T(x)$ egyenletnek egyértelmű megoldása van az x változóban minden $y \in B$ vektorra.

A vizsgálandó $\int_B f(y_1, \dots, u_k)\, dy_1 \dots dy_k$ integrált a következő jellegű integrálközelítő összegek segítségével számolhatjuk ki. Osszuk fel az A tartományt kis átmérőjű, diszjunkt $A_1, \dots, A_n$ halmazok uniójára, és definiáljuk a $B_s = T(A_s)$, $s = 1, \dots, n$ halmazokat, ahol $T(C) = \{T(x) \colon x \in C\}$ minden $C \subset A$ halmazra. Válasszunk ki mindegyik B_s halmazból egy $v^{(s)} = T(u^{(s)}) = \left(v_1^{(s)}, \dots, v_k^{(s)}\right)$, $s = 1, \dots, n$, pontot, ahol $u^{(s)} = \left(u_1^{(s)}, \dots, u_k^{(s)}\right) \in A_s$. Tekintsük ezután az $\int_B f(y_1, \dots, y_k)\, dy_1 \dots dy_k$ integrál $\sum\limits_{s=1}^{n} f\left(v^{(s)}\right) \lambda(B_s)$ alakú integrál közelítő összegeit, ahol $\lambda(C)$ egy C k-dimenziós halmaz térfogatát jelöli. A vizsgálandó integrál ezen közelítő összegek határértéke, ha az $A_1, \dots, A_n$ illetve a T transzformáció (egyenletes) folytonossága miatt a $B_1, \dots, B_n$ halmazrendszerben szereplő halmazok átmérőjének a maximuma nullához tart. Hogyan tudjuk a tekintett integrál közelítését vizsgálni a T leképzés képterében?

Az előbb tekintett integrálközelítő összegekre felírhatjuk a következő azonosságot:

$$\sum_{s=1}^{n} f\left(v^{(s)}\right) \lambda(B_s) = \sum_{s=1}^{n} f\left(T\left(u^{(s)}\right)\right) \frac{\lambda(B_s)}{\lambda(A_s)} \lambda(A_s), \tag{$*$}$$

és a vizsgált integrálra az azonosság jobboldalán szereplő kifejezés kis átmérőjű A_s halmazok választása esetén egyrészt jó közeítést ad a vizsgált integrálra, másrészt be lehet látni, hogy ez a kifejezés jól közelít egy az A tartományon alkalmasan definiált integrált, ha jó becslést tudunk adni a $\dfrac{\lambda(B_s)}{\lambda(A_s)}$ hányadosokra.

Vegyük észre, hogy $\dfrac{\lambda(B_s)}{\lambda(A_s)} = \dfrac{\lambda(T(A_s))}{\lambda(A_s)}$.

Mivel az A_s halmazok átmérője kicsi, $u^{(s)} \in A_s$, ezért a $T = (T_1, \dots, T_k)$ leképezés megszorítása az A_s halmazra jól közelíthető a

$$y_j^{(s)} - v_j^{(s)} \sim \sum_{p=1}^{k} \left. \frac{\partial T_j(x_1, \dots, x_k)}{\partial x_p} \right|_{(x_1, \dots, x_k) = (u_1^{(s)}, \dots, u_k^{(s)})} \left(x_p^{(s)} - u_p^{(s)}\right), \quad j = 1, \dots, k,$$

kifejezéssel, azaz a T transzformációnak az $u^{(s)} = \left(u_1^{(s)}, \dots, u_k^{(s)}\right)$ pont körüli Taylor sorának első tagjával.

A fenti azonosság vektor jelöléssel

$$y^{(s)} - v^{(s)} \sim \left(x^{(s)} - u^{(s)}\right) \left(\frac{\partial T(x)}{\partial x}\right)\bigg|_{x=u^{(s)}}$$

alakban írható, ahol

$$\begin{aligned} x^{(s)} - u^{(s)} &= \left(x_1^{(s)} - u_1^{(s)}, \dots, x_k^{(s)} - u_k^{(s)}\right) \\ y^{(s)} - v^{(s)} &= \left(y_1^{(s)} - v_1^{(s)}, \dots, y_k^{(s)} - v_k^{(s)}\right), \end{aligned}$$

és $\left(\frac{\partial T(x)}{\partial x}\right)\bigg|_{x=u^{(s)}}$ az a $k \times k$ méretű mátrix, melynek j-ik sorában és l-ik oszlopában a $\frac{\partial T_j(x_1, \dots, x_k)}{\partial x_l}\bigg|_{x=u^{(s)}}$ elem áll.

Ez azt jelenti, hogy a $\left(\frac{\partial T_j(x_1, \dots, x_k)}{\partial x_l}\bigg|_{x=u^{(s)}}\right)$ mátrix (elhanyagolhatón kis hibától eltekintve) az A_s halmazt a B_s halmazba viszi. Ez a tény, illetve a mátrix determinánsának megtárgyalt geometriai jelentése miatt

$$\frac{\lambda(B_s)}{\lambda(A_s)} \sim \left|\det\left(\frac{\partial T_j(x_1, \dots, x_k)}{\partial x_l}\bigg|_{x=u^{(s)}}\right)\right|.$$

Jegyezzük meg, hogy itt abszolut értéket kellett venni, mert a több-változós integrálokat olyan közelítő összegek segítségével vizsgáljuk, melyekben a halmazok térfogatát és nem előjeles térfogatát tekintjük. Ez némi eltérést jelent az egy és több-dimenziós integrálok definíciójában és viselkedésében.

A fenti reláció azt sugallja, hogy az A halmaz finom felosztása esetében a $(*)$ formula jobboldalán szereplő összeg jól közelíti az $\int_A f(T(x)) \left|\det \frac{\partial T(x)}{\partial x}\right| dx$ integrált, ezért határátmenetet végrehajtva megkapjuk az

$$\int_B f(y)\, dy = \int_A f(T(x)) \left|\det \frac{\partial T(x)}{\partial x}\right| dx$$

azonosságot. Ezt az eredményt, illetve a hozzá szükséges definíciót fogalmazom meg az alábbiakban.

Definiáljuk először a k-dimenziós tér sima transzformációinak a Jacobian-ját.

Jacobian definicíója. *Legyen $y_j = T_j(x_1, \dots, x_n)$, $1 \le j \le k$, a k-dimenziós tér egy A tartományának sima leképezés a k-dimenziós tér egy másik B tartományába. Jelölje*

$\mathbf{T}(x_1,\dots,x_k) = (T_1(x_1,\dots,x_n),\dots,T_k(x_1,\dots,x_n))$ ezt a transzformációt. E transzformáció $\mathcal{J}(\mathbf{T}(x_1,\dots,x_k))$ Jacobiánját egy $(x_1,\dots,x_k) \in A$ pontban úgy definiáljuk, hogy tekintjük először a $\mathbf{T}$ leképezés

$$\left(\frac{\partial T_j(x_1,\dots,x_k)}{\partial x_l}\right), \quad 1 \le j,l \le k,$$

deriváltját az $(x_1,\dots,x_k)$ pontban, ami egy $k \times k$ méretű mátrix. Ezután a $\mathbf{T}$ leképezés $\mathcal{J}(\mathbf{T}(x_1,\dots,x_k))$ Jacobianját az $(x_1,\dots,x_k)$ pontban úgy definiáljuk mint ezen (derivált) mátrix determinánsának az abszolut értékét.

(A Jacobian szemléletes tartalma: Ez adja meg, hogy az $(x_1,\dots,x_k)$ pont kis környezetének a térfogatát a $\mathbf{T} = (T_1,\dots,T_k)$ transzformáció hányszorosára nagyítja ki.)

Az integráltranszformációról szóló tételt először abban (a korábban tárgyalt) speciális esetben mondom ki, amikor a $\mathbf{T}$ transzformáció invertálható, majd megfogalmazom az általánosabb esetben érvényes eredményt is.

Integráltranszformációról szóló képlet speciális esete. *Legyen adva a k-dimenziós tér egy A tartományának egy sima $y_j = T_j(x_1,\dots,x_k)$, $1 \le j \le k$, transzformáltja a k-dimenziós tér egy másik B tartományába, amelyik invertálható, azaz az $y_j = T_j(x_1,\dots,x_k)$, $1 \le j \le k$, egyenletrendszernek egyetlen $(x_1,\dots,x_k) \in A$ megoldása van minden $(y_1,\dots,y_k) \in B$ pontra. Legyen továbbá adva az A tartományon egy (integrálható) $f(x_1,\dots,x_k)$ függvény. Ekkor*

$$\begin{aligned}&\int_B f(y_1,\dots,y_k)\,dy_1\dots dy_k\\ &\qquad = \int_A f(T_1(x_1,\dots,x_k),\dots,T_n(x_1,\dots,x_k))\mathcal{J}(\mathbf{T}(x_1,\dots,x_k))\,dx_1\dots dx_k,\end{aligned}$$

ahol $\mathcal{J}(\mathbf{T}(x_1,\dots,x_k))$ jelöli a $\mathbf{T}(x_1,\dots,x_k)$ leképezés Jacobianját.

Az általános eset, amikor a $\mathbf{T}$ leképezés az A tartománynak egy nem feltétlenül invertálható leképezése egy B tartományra hasonlóan tárgyalható. A különbség az, hogy most meg kell fogalmaznunk azt a megszorítást, mely szerint csak olyan függvényeket tekintünk az A tartományban, melyekre $f(x_1,\dots,x_k) = f(\bar{x}_1,\dots,\bar{x}_k)$ minden olyan $(x_1,\dots,x_k)$ és $(\bar{x}_1,\dots,\bar{x}_k)$ pontpárra, melyre $\mathbf{T}(x_1,\dots,x_k) = \mathbf{T}(\bar{x}_1,\dots,\bar{x}_k)$. Ezt úgy biztosítjuk, hogy először egy a $\mathbf{T}$ transzformáció B képterén definiált $f(y_1,\dots,y_k)$ függvényt tekintünk, majd megadjuk ennek a függvénynek a $\mathbf{T}$ transzformáció által indukált ősképét az A tartományon. Az integráltranszformációs képletben ezek a függvények jelennek meg.

Továbbá, ha tekintjük az $(y_1,\dots,y_k)$ pont kis környezetét, akkor ennek hatása a keresett azonosság másik oldalán a környezet teljes ősképén jelenik meg. Ennek az a következménye, hogy a $\mathbf{T}$ leképezés Jacobianjának értékét együtt kell tekintenünk minden olyan $(x_1,\dots,x_k)$ pontban, melyre $\mathbf{T}(x_1,\dots,x_k) = (y_1,\dots,y_k)$ valamely rögzített $(y_1,\dots,y_k) \in B$ pontban. Megfogalmazom az eredményt.

Integráltranszformációról szóló képlet. *Legyen adva a k-dimenziós tér egy A tartományának egy sima $y_j = T_j(x_1, \dots, x_k)$, $1 \le j \le k$, transzformáltja a k-dimenziós tér egy másik B tartományába. Legyen továbbá adva a B tartományon egy $f(y_1, \dots, y_k)$ függvény. Ezen $f(y_1, \dots, y_k)$ függvénynek a $\mathbf{T} = (T_1, \dots, T_k)$ transzformáció által meghatározott ősképén azt az A tartományon értelmezett $g(x_1, \dots, x_k) = \mathbf{T}^{-1} f(x_1, \dots, x_k)$ függvényt értjük az A tartományon, melyre*

$$g(x_1, \dots, x_k) = f(T_1(x_1, \dots, x_k), \dots, T_k(x_1, \dots, x_k))$$

minden $(x_1, \dots, x_k) \in A$ pontban. Ekkor

$$\begin{aligned}&\int_B f(y_1, \dots, y_k)\, dy_1 \dots dy_k \\ &\quad = \int_A \frac{\mathbf{T}^{-1} f(x_1, \dots, x_k)}{\displaystyle\sum_{\substack{\text{olyan } (z_1, \dots, z_k) \in A \text{ pontok} \\ \text{melyekre } T_j(z_1, \dots, z_k) = T_j(x_1, \dots, x_k) \\ j = 1, \dots, k}} \frac{1}{\mathcal{J}(\mathbf{T}(z_1, \dots, z_k))}}\, dx_1 \dots dx_k.\end{aligned}$$

Rövid, informális magyarázatot adok arra, hogy miért jelenik meg egy ilyen formula a fenti tételben. Legyenek az $x^{(1)} \in A, \dots, x^{(s)} \in A$ pontok az $y = \mathbf{T}(x)$ egyenlet megoldásai valamely rögzített $y \in B$ pontra, és tekintve az y pont egy kis C környezetét legyen ennek az ősképe olyan $A_1, \dots, A_s$ halmazok uniója, melyek az $x^{(1)}, \dots, x^{(s)}$ pontok kis környezetei. Ekkor $\lambda(C) \sim \lambda(A_r) \mathcal{J}\left(\mathbf{T}\left(x^{(r)}\right)\right)$ minden $r = 1, \dots, s$ számra, ahonnan $\sum\limits_{r=1}^{s} \lambda(A_r) \sim \sum\limits_{r=1}^{s} \dfrac{\lambda(C)}{\mathcal{J}(\mathbf{T}(x^{(r)}))}$, azaz $\lambda(C) \sim \dfrac{1}{\sum\limits_{r=1}^{s} \dfrac{1}{\mathcal{J}(\mathbf{T}(x^{(r)}))}} \sum\limits_{r=1}^{s} \lambda(A_r)$, és ez informális megfogalmazásban az $f(y)\, dy = \sum\limits_{p=1}^{s} \dfrac{f(\mathbf{T}(x_p))}{\sum\limits_{r=1}^{s} \dfrac{1}{\mathcal{J}(\mathbf{T}(x^{(r)}))}}\, dx_p$ formulát sugallja, ahol az $x_1, \dots, x_s$ pontok az y pont ősképei a $\mathbf{T}$ leképezés szerint, és ezekre teljesül az $f(y) = f(\mathbf{T}(x_p))$, $p = 1, \dots, s$ azonosság. Ebből a formulából következik a tételben kimondott azonosság.

Lássuk az előbb tárgyalt integráltranszformációnak legfontosabb alkalmazását, azt hogy hogyan lehet kétváltozós integrálokat polár koordinátarendszerben kiszámolni.

Tekintsük az $A = \{(r, \varphi) \colon r > 0, -\pi < \varphi < \pi\}$ tartománynak az $x = r\cos\varphi$ és $y = r\sin\varphi$ képletekkel megadott kölcsönösen egyértelmű leképezését a síkra. (Pontosabban, ez a leképezés az A halmazt arra a B halmazra képezi, melyet úgy kapunk, hogy a síkból kihagyjuk az $\{(x, y) \colon x = 0, y \le 0\}$ félegyenest, de egy síkbeli integrál értéke nem változik meg, ha egy félegyenest kihagyunk az integrálási tartományból.

Számoljuk ki a tekintett leképezés Jacobianját, majd mutassuk meg, hogy a kapott eredmény megfelel a szemléletes képnek. Egyszerű számolással, $\dfrac{\partial x}{\partial r} = \cos\varphi$, $\dfrac{\partial x}{\partial \varphi} =$

$-r\sin\varphi$, $\dfrac{\partial y}{\partial r} = \sin\varphi$, $\dfrac{\partial y}{\partial \varphi} = r\cos\varphi$, ezért a Jacobian értéke

$$\mathcal{J}(r\cos\varphi, r\sin\varphi) = \left|\det \begin{vmatrix} \cos\varphi, & -r\sin\varphi \\ \sin\varphi, & r\cos\varphi \end{vmatrix}\right| = r(\cos^2\varphi + \sin^2\varphi) = r,$$

ahonnan

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \int_{-\pi}^{\pi}\int_{0}^{\infty} rf(r\cos\varphi, r\sin\varphi)\,dr\,d\varphi. \qquad (**)$$

Annak érdekében, hogy szemléletesen is megértsük miért jelenik meg a polár koordinátarendszerre való áttéréskor az r-rel való szorzás az integrandusban, tekintsük a következő feladatot:

Ha adva van egy $T = [r, r+\Delta r] \times [\varphi, \varphi + \Delta\varphi]$ téglalap, ahol Δr és $\Delta\varphi$ kis számok és alkalmazzuk az $A(r,\varphi) = (x,y)$, $x = r\cos\varphi$, $y = r\sin\varphi$, leképezést, és tekintjük a T halmaz $A(T)$ képét, akkor mi lesz az $A(T)$ és T halmaz területének az aránya?

A T halmaz területe $\lambda(T) = \Delta r\Delta\varphi$. Az $A(T)$ halmaz azon pontokat tartalmazza az (x,y) síkon, melyeknek az abszolut értéke r és $r+\Delta r$ közé, az abszcissza tengellyel bezárt szöge pedig φ és $\varphi+\Delta\varphi$ közé esik. Ezért ennek területe $A(T) = (r+\Delta r)^2\dfrac{\Delta\varphi}{2} - r^2\dfrac{\Delta\varphi}{2} = r\Delta r\Delta\varphi + \dfrac{(\Delta r)^2\Delta\varphi}{2}$, ahonnan $\dfrac{\lambda(A(T))}{\lambda(T)} = r + \dfrac{\Delta r}{2}$. Innen következik, hogy a $\dfrac{\lambda(A(T))}{\lambda(T)}$ hányados az r számhoz tart, ha $\Delta r \to 0$, és $\Delta\varphi \to 0$. Ebből a formulából és az integráloknak a szokásos integrálközelítő összegek segítségével felírt approximációjából következik, hogy a polár-koordinátarendszerbe való áttéréskor, az (x,y) koordinátáknak az $r\cos\varphi$, $r\sin\varphi$ változókkal való helyettesítésekor az integrandust az r számmal kell megszorozni, azaz a $(**)$ formula érvényes.