MÁTRIX-INVERTÁLÁS, LAPLACE–KIFEJTÉS, A CAUCHY-BINET–FORMULÁK, CAYLEY TÉTELE

Az alábbiakban igazoljuk a mátrix-invertálás Gauss–eliminációs algoritmusát, a determináns Laplace–kifejtésére vonatkozó képletet, a Cauchy-Binet formulákat, és ennek alkalmazásaként belátjuk Cayley tételét, mely szerint az n pontú számozott fák száma n^{n-2}.

1. A determináns, mint mérték

Az alábbi felépítést érdemes összevetni azzal, ami Freud Róbert: *Lineáris algebra* című könyvének 9.8. szakaszában szerepel. A teljesség kedvéért a determináns összes alaptulajdonságát igazoljuk.

1.1. Definíció. Legyen T test és $V = T^n$. Egy $D : V^n \to T$ függvényt *előjeles mérték*nek nevezünk, ha

(1) *multilineáris,* azaz minden változójában lineáris, miközben a többi változó rögzített;
(2) *alternáló,* azaz ha két változó egyenlő, akkor a D értéke nulla.

1.2. Állítás. *A D előjeles mérték értéke nem változik, ha valamelyik változójához egy másik változó tetszőleges λ-szorosát adjuk, ahol $\lambda \in T$.* □

1.3. Lemma. *Ha D előjeles mérték, és két változóját felcseréljük, akkor D előjelet vált.*

Bizonyítás. Jelölje H azt a kétváltozós függvényt, amit D-ből úgy kapunk, hogy $n-2$ változóját tetszőlegesen rögzítjük. Ekkor H mindkét változójában lineáris, és $H(v, v) = 0$ minden $v \in V$-re. Azt kell belátnunk, hogy $u, v \in V$ esetén $H(u, v) = -H(v, u)$. Az előző következményt háromszor alkalmazva

$$H(u, v) = H(u + v, v) = H(u + v, v - (u + v)) = H(u + v, -u) = H(v, -u) = -H(v, u)\,,$$

hiszen H lineáris a második változóban. □

Az alternáló elnevezést az előző állítás indokolja. Megjegyezzük, hogy ha T karakterisztikája nem 2, akkor a mérték definíciójában használhatnánk azt a tulajdonságot is, hogy a mérték két változó cseréjekor előjelet vált.

1.4. Tétel. *Legyen $b_1, \ldots, b_n \in V$ tetszőleges, és $1 \le i \le n$ esetén*

$$v_i = \lambda_{1i} b_1 + \ldots + \lambda_{ni} b_n$$

valamilyen $\lambda_{ij} \in T$-re. Ekkor

$$D(v_1, \ldots, v_n) = D(b_1, \ldots, b_n) \sum_{f \in S_n} \mathrm{sg}(f) \lambda_{f(1),1} \ldots \lambda_{f(n),n}\,.$$

Az összegezés az $\{1, 2, \ldots, n\}$ összes f permutációin fut végig, és $\mathrm{sg}(f)$ az f előjele.

Bizonyítás. Ha a linearitást $D(v_1, \dots, v_n)$ minden változójában alkalmazzuk, akkor egy n^n tagú összeget kapunk, amelynek tagjai úgy keletkeznek, hogy mindegyik v_i-t előállító összegnek kivesszük egy-egy tagját, és ezeket D-be helyettesítjük. Jelölje $f(i)$ a v_i-ből kivett tagnak a sorszámát. Ekkor $f : \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}$ egy tetszőleges függvény, és

$$D(v_1, \dots, v_n) = \sum_f D(\lambda_{f(1),1} b_{f(1)}, \dots, \lambda_{f(n),n} b_{f(n)}) .$$

Mivel D multilineáris, a $\lambda_{f(i),i}$ együtthatókat kiemelhetjük. Ha f nem permutáció, akkor $D(b_{f(1)}, \dots, b_{f(n)}) = 0$, hiszen D alternáló. Ha f permutáció, akkor az argumentumokat cserélgetve azt kapjuk, hogy $D(b_{f(1)}, \dots, b_{f(n)}) = \mathrm{sg}(f) D(b_1, \dots, b_n)$, hiszen D két argumentum seréjekor előjelet vált. □

1.5. Következmény. *A D függvényt egyértelműen meghatározza V tetszőleges $b_1, \dots, b_n$ bázisán felvett értéke. Ha D nem azonosan nulla, akkor egyetlen bázison sem veheti fel a nulla értéket.* □

1.6. Tétel. *Legyen $v_i = (\lambda_{1i}, \dots, \lambda_{ni})^T$ (oszlopvektorok). Ekkor a*

$$D(v_1, \dots, v_n) = \sum_{f \in S_n} \mathrm{sg}(f) \lambda_{f(1),1} \dots \lambda_{f(n),n}$$

képlet előjeles mértéket definiál.

Bizonyítás. Az így definiált D nyilván multilineáris, hiszen a fenti szumma tagjai mindegyik v_i koordinátái közül pontosan egy tényezőt tartalmaznak. Tegyük fel, hogy $v_i = v_j$ de $i \neq j$. Ez azt jelenti, hogy $\lambda_{ki} = \lambda_{kj}$ minden k-ra. Párosítsuk az f permutációt $f \circ (i, j)$-vel. Elég megmutatni, hogy a fenti szumma két megfelelő tagja egyenlő, de az előjelezése ellentétes. Utóbbi következik a permutációk szorzástételéből, hiszen az (i, j) csere előjele -1. Mivel

$$\lambda_{[f \circ (i,j)](i),i} = \lambda_{f(j),i} = \lambda_{f(j),j}$$

és hasonlóan $\lambda_{[f \circ (i,j)](j),j} = \lambda_{f(i),i}$, végül $k \neq i, j$ esetén $\lambda_{[f \circ (i,j)](k),k} = \lambda_{f(k),k}$, ezért a két megfelelő szorzat tényleg egyenlő (csak az i-edik és j-edik tényező megcserélődött). □

1.7. Definíció. Ha $M = ((\lambda_{ij})) \in T^{n \times n}$, akkor M *determinánsa* az M oszlopainak az előző tételben definiált előjeles mértéke, azaz $\det(M) = \sum_{f \in S_n} \mathrm{sg}(f) \lambda_{f(1),1} \dots \lambda_{f(n),n}$.

1.8. Tétel. *Felső háromszögmátrix determinánsa a főátlóbeli elemek szorzata. Speciálisan az egységmátrix determinánsa* 1.

Bizonyítás. Ha $M = ((\lambda_{ij}))$ felső háromszögmátrix, akkor $i > j$ esetén $\lambda_{i,j} = 0$. Ezért ha a determinánst definiáló szumma f-hez tartozó tagja nem nulla, akkor minden i-re $f(i) \leq i$. Speciálisan $f(1) = 1$, de akkor $f(2) \leq 2$ miatt $f(2) = 2$, és így tovább, tehát f az identitás. Az ehhez tartozó tag a főátlóbeli elemek szorzata, $\mathrm{sg}(id) = 1$ előjellel szorozva. □

1.9. Tétel. *Transzponált mátrix determinánsa megegyezik az eredeti mátrix determinánsával.*

Bizonyítás. Ha $M = ((\lambda_{ij}))$, akkor $M^T = ((\lambda_{ji}))$, vagyis az indexek megcserélődnek. Ezért

$$det(M^T) = \sum_{f \in S_n} \mathrm{sg}(f) \lambda_{1,f(1)} \dots \lambda_{n,f(n)} .$$

Rendezzük ezt a szorzatot a második indexek szerint: $\lambda_{1,f(1)} \dots \lambda_{n,f(n)} = \lambda_{g(1),1} \dots \lambda_{g(n),n}$. Ebben λ_{ij} akkor és csak akkor szerepel, ha $f(i) = j$, akkor és csak akkor, ha $i = g(j)$. Azaz $g = f^{-1}$ (az f inverze). Az $f \leftrightarrow f^{-1}$ kölcsönösen egyértelmű megfeleltetés a permutációk között, és $\mathrm{sg}(f)\,\mathrm{sg}(g) = \mathrm{sg}(fg) = \mathrm{sg}(id) = 1$ miatt $\mathrm{sg}(f) = \mathrm{sg}(g)$. Ezért a fenti szumma $\det(M)$-mel egyenlő. □

1.10. Tétel. *Egy mátrix determinánsa akkor és csak akkor nulla, ha oszlopai lineárisan összefüggenek.*

Bizonyítás. Ha az M mátrix $v_1, \dots, v_n$ oszlopaira $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$, ahol valamelyik $\lambda_i \neq 0$, akkor szorozzuk meg az i-edik oszlopot λ_i-vel, és adjuk hozzá a többi oszlop megfelelő λ_j-szeresét. Ekkor a determináns értéke nulla lesz, hiszen van egy csupa nulla oszlopa. Másrészt ez az eredeti determináns λ_i-szerese. Ezért az eredeti determináns is nulla.

Most tegyük fel, hogy M oszlopai függetlenek, tehát bázist alkotnak. A determináns nem azonosan nulla mérték, mert az egységmátrixon az értéke 1. Az 1.5. Következmény miatt tehát $\det(M) \neq 0$. □

1.11. Tétel [Szorzástétel]. *Ha $M, N \in T^{n\times n}$, akkor* $\det(MN) = \det(M)\det(N)$.

Bizonyítás. Legyenek M oszlopai $b_1, \dots, b_n$ és $N = ((\lambda_{ij}))$. Ekkor az MN mátrix $v_1, \dots, v_n$ oszlopaira $v_i = \lambda_{1i} b_1 + \dots + \lambda_{ni} b_n$. Legyen $D = \det$, ez előjeles mérték, ezért az 1.4. Tétel miatt

$$\det(v_1, \dots, v_n) = \det(b_1, \dots, b_n) \sum_{f \in S_n} \mathrm{sg}(f) \lambda_{f(1),1} \dots \lambda_{f(n),n} \,.$$

A fenti szumma értéke $\det(N)$. □

2. A Laplace–kifejtés

Legyen T test és $M \in T^{n\times m}$. Az M mátrix egy $r \times r$-es aldeterminánsán azt értjük, hogy kiválasztjuk a sorok halmazának egy r elemű I részhalmazát, az oszlopok halmazának egy r elemű J részhalmazát, és vesszük a kapott sorok és oszlopok metszéspontjaiban álló elemek alkotta mátrix determinánsát. Ennek jele $M_{I,J}$. Ha $I = \{i_1, \dots, i_r\}$ és $J = \{j_1, \dots, j_r\}$, akkor az $M_{I,J}$ aldeterminánshoz tartozó $\mathrm{sg}(I, J)$ előjel legyen $(-1)^{i_1+\dots+i_r+j_1+\dots+j_r}$.

2.1. Tétel. *Legyen $M \in T^{n\times n}$. Rögzítsük az r elemű $I \subseteq \{1, 2, \dots, n\}$ halmazt tetszőlegesen, és jelölje I' az I komplementumát az $\{1, 2, \dots, n\}$ halmazra nézve. Ekkor az M determinánsának Laplace-féle kifejtése*

$$\det(M) = \sum_J \mathrm{sg}(I, J) M_{I,J} M_{I',J'} \,,$$

ahol az összegzés az oszlopok r elemű J részhalmazaira terjed ki (tehát $\binom{n}{r}$ tag van).

Bizonyítás. Elég megmutatni, hogy a jobb oldalon álló összeg az M oszlopainak előjeles mértéke, ami az egységmátrixon 1 (hiszen az 1.5. Következmény miatt ilyen mérték csak egyetlenegy van: maga a determináns). Ha mondjuk M első oszlopát λ-val szorozzuk, akkor az összeg J-edik tagja λ-val szorzódik mindegyik J-re, hiszen $1 \in J$ esetén $M_{I,J}$, ellenkező esetben pedig $M_{I',J'}$ szorzódik J-vel. Ugyanígy láthatjuk be az összegtartást is.

Ha M az egységmátrix, akkor egy $M_{I,J}$ aldetermináns csak úgy lehet nullától különböző, ha minden sorában és oszlopában van 1-es. Az $i \in I$-edik sorbeli 1-es az i-edik oszlopban van, tehát $J = I$, így az összegnek egyetlen nem nulla tagja van, ebben mindkét tényező az egységmátrix, az előjel pedig 1, hiszen a -1 kitevőjében minden szám kétszer szerepel.

Tehát már csak azt kell belátni, hogy ha M két oszlopa, mondjuk az u-adik és a v-edik egyenlő, akkor a jobboldali összeg nulla. Azokra a J halmazokra, melyekre mindkét oszlop J-ben, vagy J'-ben szerepel, a megfelelő tag nyilván nulla. Ha J nem ilyen halmaz, akkor tehát az egyik oszlop, mondjuk az u-hoz tartozó, benne van, a másik nincs. Legyen $\overline{J}$ az a halmaz, amely J-ből úgy keletkezik, hogy u-t kivesszük, és v-t betesszük. Ez kölcsönösen egyértelmű megfeleltetés, azt kell belátni, hogy a megfelelő két tag az összegben kiejti egymást. Mivel a két oszlop egyenlő, ezért az $M_{I,J}$ és $M_{I,\overline{J}}$ oszlopvektorai ugyanazok, csak esetleg más sorrendben szerepelnek, és ugyanez mondható a komplementerekről is.

Számoljuk meg, hány eleme van J-nek u és v között (u-t nem beleszámítva). Legyen ez a szám k. Ekkor $M_{I,J}$-ből $M_{I,\overline{J}}$ úgy kapható meg, hogy k oszlopcserét végzünk. A J' halmazban azok az u és v közötti számok vannak, amik nincsenek J-ben, ezek száma tehát $|v-u|-1-k$, ennyi oszlopcserét kell elvégezni a komplementer halmazhoz tartozó aldeterminánson. Ez összesen $|v-u|-1$ csere. Viszont $\mathrm{sg}(I,\overline{J})$ is megváltozott $\mathrm{sg}(I,J)$-hez képest, méghozzá $(-1)^{v-u}$-val. Mivel $|v-u|-1+(v-u)$ páratlan szám (hiszen $|v-u|$ és $v-u$ paritása ugyanaz), a két tag tényleg kiejti egymást. $\square$

3. A Cauchy-Binet formulák

3.1. Tétel. *Legyen T test $M \in T^{m\times k}$, $N \in T^{k\times n}$ és $K = MN$. Rögzítsük az r elemű $I \subseteq \{1,2,\ldots,m\}$ és $J \subseteq \{1,2,\ldots n\}$ részhalmazokat. Ekkor a Laplace–kifejtésnél használt jelöléssel*

$$K_{I,J} = \sum_{S\subseteq\{1,2,\ldots k\},\ |S|=r} M_{I,S}N_{S,J}\,.$$

Bizonyítás. Mivel az M mátrixnak csak az I-beli sorairól van szó, a többi sort elhagyhatjuk. Hasonló mondható N oszlopairól is, tehát feltehető, hogy $m = r = n$, és így $K_{I,J} = \det(K)$.

Legyenek $u_1,\ldots,u_r$ az M sorvektorai. Ha u_i-t λ-val szorozzuk, akkor a mátrixszorzás definíciója miatt a K mátrix i-edik sora is λ-val szorzódik. Hasonlóképpen, ha u_i-t két vektor összegére bontjuk, akkor a K mátrix i-edik sora is a megfelelő két vektor összegére bomlik. Ezért a tételben szereplő egyenlőség mindkét oldala lineáris mindegyik u_i változóban. Hasonló mondható el N oszlopvektorairól is. Ha tehát ezeket a vektorokat felírjuk $\sum \lambda_i e_i$ alakban (ahol e_i-ben az i-edik komponens 1, a többi nulla), és a keletkező összeget a linearitás szerint kibontjuk (ugyanúgy, mint az 1.4. Tétel bizonyításában), akkor mindkét oldal ugyanúgy bomlik szét, és így elég az állítást abban az esetben belátni, ha mind M soraiban, mind N oszlopaiban csak ilyen e_i vektorok szerepelnek.

Ha van két egyenlő sor M-ben, vagy két egyenlő oszlop N-ben, akkor mindkét oldal nulla. Mivel mind M-beli sorcsere, mind N-beli oszlopcsere esetében mindkét oldal előjelet vált, ezért még az is feltehető, hogy ezek az e_i vektorok „növekvő” sorrendben szerepelnek mindkét mátrixban. Azaz M sorai $e_{i_1}^T,\ldots,e_{i_r}^T$, és N oszlopai $e_{j_1},\ldots,e_{j_r}$, ahol $1 \le i_1 < \cdots < i_r \le k$ és $1 \le j_1 < \cdots < j_r \le k$.

Legyen $U = \{i_1,\ldots,i_r\}$ és $V = \{j_1,\ldots,j_r\}$. Ekkor $M_{I,S} = 1$ ha $S = U$ és 0 egyébként. Hasonlóan $N_{S,J} = 1$ ha $S = V$ és 0 egyébként. Ezért ha $U \ne V$, akkor a jobb oldali összeg nulla, de $\det(K)$ is nulla, mert K-ban van csupa nulla sor. Ha pedig $U = V$, akkor K az egységmátrix, és a jobb oldali összegnek egyetlen nem nulla tagja van, amikor $S = U = V$, és ennek értéke 1. $\square$

4. Cayley tétele

4.1. Tétel [Cayley]. *Az n csúcsú számozott fák száma n^{n-2}.*

A bizonyításhoz legyen G egy hurokélek és párhuzamos élek nélküli irányított gráf, melynek n csúcsa és m éle van. Készítsük el a következő M mátrixot. Az M sorai G éleinek, oszlopai G csúcsainak felelnek meg. Az e élnek megfelelő sorban $+1$-et írunk ahhoz a csúcshoz, ahonnan e kiindul, -1-et ahhoz a csúcshoz, ahová e érkezik, a többi helyre pedig nullát.

4.2. Lemma. *Ha a G gráf (az irányítást elfeledve) legalább két csúcsú fa, akkor az M minden $(n-1)\times(n-1)$-es aldeterminánsának értéke ± 1.*

Bizonyítás. Az állítás $n=2$-re nyilvánvaló, n szerinti indukcióval bizonyítunk. Legyen N egy $(n-1)\times(n-1)$-es részmátrix, ez egy v csúcshoz tartozó oszlop elhagyásával adódik. Mivel minden legalább két csúcsú fának legalább két elsőfokú csúcsa van, választhatunk egy v-től különböző, elsőfokú u csúcsot. Legyen e az u-ból kiinduló egyetlen él. Az u oszlopában minden elem nulla, kivéve az e sorában álló ± 1-et. Az ehhez tartozó aldetermináns értéke az indukciós feltevés szerint ± 1, mert az e és v elhagyásával kapott gráf is fa. Ezért az N determinánsát u oszlopa szerint kifejtve készen vagyunk. $\square$

4.3. Lemma. *Az M mátrix bizonyos sorvektorai akkor és csak akkor lineárisan függetlenek, ha a megfelelő élek által feszített részgráf nem tartalmaz (nem feltétlenül irányított) kört. Speciálisan $n-1$ él akkor és csak akkor független, ha feszítő fát alkotnak.*

Bizonyítás. Tekintsük egy kör éleit. Az egyes élek irányításának megfordítása a mátrix megfelelő sorainak függetlenségét nem befolyásolja, ezért feltehető, hogy ez a kör irányított. Ekkor viszont a kör éleihez tartozó sorok összege nulla.

Megfordítva, tegyük fel, hogy sorok egy halmaza lineárisan összefügg. Vegyünk egy nemtriviális lineáris kombinációt, ami nullát ad. Dobjuk el azokat az éleket, ahol a megfelelő együttható nulla. Elég megmutatni, hogy a megmaradó élek alkotta gráfban minden csúcs legalább másodfokú (mert akkor van benne kör). De ez igaz, mert ha a v csúcsból kiinduló egyetlen e élhez a lineáris kombinációban a λ együttható tartozik, akkor a v-hez tartozó (ezekre a sorokra vonatkozó) oszlopösszeg $\pm\lambda$ lesz, holott nullának kellene lennie.

A második állítás abból következik, hogy egy n csúcsú és $n-1$ élű gráf akkor és csak akkor fa, ha nem tartalmaz kört. $\square$

Legyen G irányítatlan egyszerű gráf, melynek n csúcsa van. Készítsük el azt az $n\times n$-es K mátrixot, amelynek sorai és oszlopai is G csúcsainak felelnek meg, a főátlóban a megfelelő csúcs fokszáma van, az i-edik sor j-edik eleme pedig -1, ha az i-edik és j-edik csúcs között megy él, és nulla ha nem megy él.

4.4. Tétel. *A K mátrixban a főátló bármelyik eleméhez tartozó nem előjelezett aldetermináns értéke megegyezik G feszítő fáinak számával.*

Bizonyítás. Könnyű látni, hogy az imént vizsgált M mátrixra $M^T M = K$ teljesül (az élek irányítása közömbös). Alkalmazzuk a Cauchy-Binet formulát arra az $I=J$ halmazra, amit úgy kapunk, hogy G-ből egy tetszőleges csúcsot elhagyunk. Ha S az élek tetszőleges $n-1$ elemű részhalmaza, akkor az előző két lemma miatt az $M_{I,S} = M^T_{S,J}$ determináns értéke ± 1 ha az S-beli élek feszítő fát alkotnak, és 0 ha nem. Ezért a Cauchy-Binet formula jobb oldalán álló összeg G feszítő fáit számolja meg. $\square$

A Cayley–tétel bizonyításához alkalmazzuk az előző tételt akkor, amikor G az n pontú teljes gráf. Ekkor K az a mátrix, amiben a főátló végig $n-1$, a többi elem pedig -1. A bal felső sarokban álló $(n-1)\times(n-1)$-es aldeterminánst a következőképpen számíthatjuk ki. Először az összes sort hozzáadjuk az elsőhöz. Ekkor az első sorban végig 1-esek állnak. Ezt a sort a többihez adjuk, ez eltünteti a főátló alatti elemeket, a főátló elemeinek szorzata pedig n^{n-2} lesz.

5. Invertálás Gauss–eliminációval

5.1. Állítás. *Legyen T test, $M \in T^{n\times n}$, és E^{ij} az a mátrix, amiben az i-edik sor j-edik eleme 1, a többi elem 0.*

1. *Az $E^{ij}M$ mátrix i-edik sora éppen az M mátrix j-edik sora, a többi sor pedig azonosan nulla.*
2. *Az $E+E^{ij}+E^{ji}-E^{ii}-E^{jj}=E^{ij}+E^{ji}+\sum_{k\neq ij}E^{kk}$ mátrixszal való balszorzás felcseréli M-ben az i-edik és a j-edik sort.*
3. *Az $E+\lambda E^{ij}$-vel való balszorzás az M mátrix j-edik sorának λ-szorosát hozzáadja az i-edik sorhoz.*
4. *Ha az egységmátrix főátlójának i-edik elemét 1-ről λ-ra változtatjuk, akkor az ezzel a mátrixszal való balszorzás az M mátrix i-edik sorát λ-szorosára változtatja.*

Balszorzás helyett jobbról szorozva ezekkel a speciális mátrixokkal, a sorokra vonatkozó átalakítások helyett a megfelelő oszlopokra vonatkozó átalakításokat kapjuk. □

5.2. Tétel. *Legyen $M \in T^{n\times n}$. Gauss elimináció segítségével (tehát az előző állítás (2), (3) és (4) átalakításait használva, ahol a (4) pontban $\lambda \neq 0$-t szabad csak használni) az M mátrix vagy egységmátrixszá alakítható, vagy pedig keletkezik benne egy csupa nulla sor. Utóbbi esetben M nem invertálható. Az előbbi esetben az egységmátrixszal ugyanezeket az átalakításokat elvégezve M inverze adódik.*

Bizonyítás. Az eljárás során az nem változik meg, hogy az átalakított mátrix determinánsa nulla-e. Ezért ha egy csupa nulla sor keletkezik, akkor $\det(M)=0$, és ezért M-nek nincs inverze. Tegyük fel, hogy egységmátrix keletkezik, és jelölje az egyes átalakításokhoz tartozó mátrixokat rendre $S_1,\dots,S_k$. Ekkor $S_k\dots S_1M=E$, ezért $M^{-1}=S_k\dots S_1=S_k\dots S_1E$. Vagyis az egységmátrixszal ugyanezeket az átalakításokat elvégezve M inverze adódik. □