A Cochran–Fisher tételről

A matematikai statisztika egyik fontos eredménye a Cochran–Fisher tétel, amely a variancia analízisben játszik fontos szerepet. Ugyanakkor ez a tétel lényegét tekintve valójában egy lineáris algebrai állítás. Érdemes megmutatni, hogy hogyan lehet ezt az eredményt, illetve néhány ehhez kapcsolódó a matematikai statisztikában szintén hasznos állítást a lináris algebra módszereivel bebizonyítani.

E jegyzet ezt a bizonyítást és a Cochran–Fisher tétel néhány alkalmazását tartalmazza. Először megfogalmazom magát az eredményt.

Cochran–Fisher tétel. *Legyen adva n független, standard normális eloszlású valószínűségi változó, $\xi_1,\dots,\xi_n$, és definiáljuk segítségükkel a $Q=Q(\xi_1,\dots,\xi_n)=\sum\limits_{j=1}^n \xi_j^2$ kifejezést. Legyen adva ezenkívül még k darab $Q_j=Q_j(\xi_1,\dots,\xi_n)$, $Q_j=\sum\limits_{i=1}^n\sum\limits_{l=1}^n c_{i,l}^{(j)}\xi_i\xi_l$, $1\le j\le k$, alakú (quadratikus) kifejezés a $c_{i,l}^{(j)}=c_{l,i}^{(j)}$ azonosságot minden $1\le i,l\le n$, $1\le j\le k$, indexre teljesítő együtthatókkal, amelyekre érvényes a*

$$Q=Q_1+Q_2+\cdots+Q_k \tag{1}$$

azonosság. Jelölje $\operatorname{rang}(Q_j)$ *a Q_j kifejezés rangját, amelyet úgy definiálunk, mint a Q_j kifejezés definiciójában szereplő együtthatók segítségével definiált $(c_{i,l}^{(j)})$, $1\le i,l\le n$, mátrix rangját. Ekkor teljesül az*

$$n\le \operatorname{rang}(Q_1)+\operatorname{rang}(Q_2)+\cdots+\operatorname{rang}(Q_k) \tag{2}$$

egyenlőtlenség. Ha a (2) formulában egyenlőség érvényes, akkor a Q_j, $1\le j\le k$, kifejezések független χ^2 eloszlású valószínűségi változók $\operatorname{rang}(Q_j)$ *szabadságfokkal. Sőt, ebben az esetben léteznek olyan független $\zeta_{j,1},\dots,\zeta_{j,r_j}$, $r_j=\operatorname{rang}(Q_j)$, $1\le j\le k$, standard normális eloszlású valószínűségi változók, amelyekre $Q_j=\sum\limits_{s=1}^{r_j}\zeta_{j,s}^2$, $1\le j\le k$, és $Q=\sum\limits_{j=1}^{k}\sum\limits_{s=1}^{r_j}\zeta_{j,s}^2$.*

Fel fogom idézni egy mátrix rangjának a definicióját és legfontosabb tulajdonságait. Ezelőtt megfogalmazom azt a lineáris algebrai állítást, amelyből a Cochran-Fisher tétel következik.

Tétel egy Euklideszi térbeli identitás transzformáció felbontásáról. *Tekintsük az $I=I_n$ identitás transzformációt az n-dimenziós Euklideszi térben. Legyen adva ennek egy*

$$I=Q_1+Q_2+\cdots+Q_k \tag{3}$$

alakú előállítása, ahol $Q_1,\dots,Q_k$ szimmetrikus leképezések. Ekkor

$$n\le \operatorname{rang}(Q_1)+\operatorname{rang}(Q_2)+\cdots+\operatorname{rang}(Q_k). \tag{4}$$

Ha a (4) formulában egyenlőség teljesül, akkor a $Q_1, \dots, Q_k$ leképezések egymásra merőleges projekciók.

Felidézem, hogy egy mátrix rangját úgy definiáljuk, mint a mátrixból kiválasztható maximális számú lineárisan független sorok számát. Tekintsünk egy szimmetrikus bilineáris függvényt meghatározó lineáris transzformációt egy Euklideszi térben. Egy ilyen transzformáció tetszőleges ortonormált koordinátarendszerben felírt mátrixának ugyanaz a rangja, mert a transzformáció két különböző koordinátarendszerben felírt A és B mátrix reprezentációja között az $A = UBU^*$ reláció érvényes alkalmas U unitér mátrix-szal. Innen következik, hogy beszélhetünk egy szimmetrikus bilineáris függvény vagy az őt egyértelműen meghatározó kvadratikus alak rangjáról is. Egy (nem feltétlenül szimmetrikus) transzformáció rangja megegyezik a transzformáció képterének dimenziójával, és egy kvadratikus alak rangja jellemezhető úgy is, mint a leképezés főtengely transzformációs alakjában szereplő nem zéró sajátértékek száma.

Érdemes tudni a következő eredményeket a lineáris algebrából egy mátrix vagy lineáris leképezés rangjáról. Egy mátrix lineárisan független sorainak és lineárisan független oszlopainak száma megegyezik, és ezek egyike tekinthető egy mátrix rangjának. Ha egy mátrixot megszorzunk (akár balról akár jobbról) egy másik mátrix-szal, a szorzatmátrix rangja kisebb vagy egyenlő, mint az eredeti mátrix rangja. Ha egy invertálható mátrix-szal szorzunk, akkor egyenlőség van. Ha adva van egy lineáris transzformáció, akkor ennek mátrixa függ attól, hogy mely bázis segítségével írtuk fel, de a kapott mátrix rangja már nem függ ettől. Ezen eredemény alapján beszélhetünk egy lineáris transzformáció rangjáról, mint az általa meghatározott mátrix rangjáról.

Annak érdekében, hogy az identitás transzformációnak egy Euklideszi térben történő felbontásáról szóló tételt jobban megértsük, érdemes felidézni egy Euklideszi tér projekciójának fogalmát és legfontosabb tulajdonságait. Geometriai módon egy E Euklideszi tér (ortogonális) projekcióját egy $B \subset E$ altérre úgy definiáljuk, mint az Euklideszi tér $x \in E$ pontjainak azt az $x \to Px$ transzformációját, amely az x pontnak e pont orthogonális vetületét felelteti meg a B altérre, azaz $Px \in B$, és $x - Px$ merőleges minden $y \in B$ vektorra, azaz $(x - Px, y) = 0$, ha $y \in B$. Két P és Q projekció merőleges egymásra, ha a Px és Qy vektorok merőlegesek egymásra minden $x, y \in E$ vektorra, azaz $(Px, Qy) = 0$ minden $x \in E$ és $y \in E$ vektorra. Érdemes a projekciók és projekciók merőlegességének fogalmát algebrai nyelven is megfogalmazni. A következő lemma a projekciók ilyen jellemzését fogalmazza meg.

Lemma egy Euklideszi tér projekcióinak a tulajdonságairól. *Egy E Euklideszi tér P projekciója szimmetrikus operátor, és a projekció definiciójában szereplő B altér, ahová a projekció vetít, megegyezik a P transzformáció $B = \{Px\colon\ x \in E\}$ képterével. Egy $x \in B$ vektorra $Px = x$, és egy a B altérre merőleges x vektorra $Px = 0$.*

Egy Euklideszi tér P önmagába való lineáris leképezése akkor és csak akkor projekció, ha szimmetrikus, és teljesíti a $P = P^2$ azonosságot. Két P és Q projekció akkor és csak akkor merőleges egymásra, ha $PQ = 0$. Ekkor a $QP = 0$ azonosság is teljesül. Speciálisan, ha P projekció, akkor $I - P$ egy rá merőleges projekció. Ezenkívül érvényes a $0 \le P \le I$ azonosság, azaz $0 \le (Px, x) \le (x, x)$ minden $x \in E$ vektorra.

A lemma bizonyítása: Legyen P projekció az E Euklideszi tér egy B alterére. Lássuk be először, hogy a B altér megegyezik a P operátor képterével. Valóban, egyrészt definició szerint $Px \in B$ minden $x \in E$ vektorra. Másrészt nincs olyan $y \in B$ vektor, amelyet nem tartalmaz a P operátor képtere. Ugyanis egy ilyen y vektor létezése esetén $y - Py$ olyan nem zéró vektor lenne, amelyre $y - Py \in B$, ezért merőleges önmagára, és ez ellentmondás. Ha $x \in B$, akkor $z = Px = x$. Valóban, ekkor a $z = Px$ vektorra teljesül mind az $x - z \in B$, mind az $(x - z, y) = 0$ minden $y \in B$ vektorra. Hasonlóan $Px = 0$, ha x merőleges a B altérre, mert ekkor $(x - 0, y) = 0$ minden $y \in B$ vektorra, és $0 \in B$, tehát az x vektor $x = (x - 0) + 0$, és $x = (x - Px) + Px$ felbontása egy B-re merőleges és egy a B altér által tartalmazott vektorra megegyezik.

Mutassuk meg azt is, hogy egy P projekció olyan szimmetrikus operátor, amelyre $P^2 = P$. Valóban, a $P = P^*$ reláció bizonyításához elég megmutatni, hogy $(Px, y) = (x, Py)$, ha $y \in B$ vagy ha y merőleges a B altérre. Ha $y \in B$, akkor $y = Py$, ezért $(Px, y) - (x, Py) = (Px, y) - (x, y) = (Px - x, y) = 0$, és ha y merőleges a B altérre, akkor $Py = 0$, ezért $(x, Py) = 0$ és $(Px, y) = 0$, mert $Px \in B$. A $Px = P^2x$ bizonyítását is elég ellenőrizni akkor, ha $x \in B$ vagy ha x merőleges a B altérre. Az első esetben viszont $x = Px = P^2x$, a második esetben pedig $0 = Px = P^2x$.

Lássuk be, hogy ha $P = P^*$, és $P = P^2$, akkor P projekció a P transzformáció $B = \{Px\colon\ x \in E\}$ képterére. Valóban, $Px \in B$ minden $x \in E$ vektorra, és mivel $y = Pz = P^2z = Py$ valamely $z \in E$ vektorral, ha $y \in B$, ezért $(x - Px, y) = (x, y) - (Px, y) = (x, Py) - (x, P^*y) = 0$ minden $x \in E$ és $y \in B$ vektorra.

Két P és Q projekció akkor és csak akkor merőleges, ha $(Px, Qy) = 0$ minden $x \in E$ és $y \in E$ vektorra. Viszont $(Px, Qy) = (P^*x, Qy) = (x, PQy)$, ezért $(Px, Qy) = 0$ minden $x \in E$ és $y \in E$ vektorra akkor és csak akkor, ha $PQ = 0$. Ha $PQ = 0$, akkor $0 = (PQ)^* = Q^*P^* = QP$. Ha P projekció, akkor $(I - P)^* = I^* - P^* = I - P$, $(I - P)^2 = I - 2P + P^2 = I - P$, és $P(I - P) = 0$, azaz $I - P$ a P projekcióra merőleges projekció. Végül, ha P projekció, akkor $(x, Px) = (x, P^2x) = (Px, Px) \geq 0$, és alkalmazva ezt az összefüggést az $I-P$ projekcióra azt kapjuk, hogy $(x, (I-P)x) \geq 0$, azaz $(x, Px) \leq (x, x)$.

Független normális eloszlású valószínűségi változók véges sorozata többváltozós normális eloszlású vektor, és egy normális eloszlású véletlen vektor lineáris transzformáltja szintén normális eloszlású. Ezenkívül tudjuk, hogy egy többváltozós normális eloszlású vektor koordinátái akkor és csak akkor függetlenek, ha korrelátlanok. Megmutatom e tények felhasználásával, hogy az egy Euklideszi térbeli identitás transzformáció felbontásáról szóló tételből következik a Cochran–Fisher tétel.

A $Q_j(x) = \sum\limits_{i=1}^{n} \sum\limits_{l=1}^{n} c_{i,l}^{(j)} \xi_i \xi_l$ kifejezésnek feleltessük meg a $C^{(j)} = (c_{i,l}^{(j)})$, $1 \leq i, l \leq n$, szimmetrikus mátrixot minden $1 \leq j \leq k$ indexre. Ezenkívül vezessük be az E_n n-dimenziós Euklideszi teret, amelynek elemei az $x = (x_1, \dots, x_n)$ valós szám n-esek a szokásos $(x, y) = \sum_{s=1}^{n} x_s y_s$ skalárszorzással, ha $x = (x_1, \dots, x_n)$ és $y = (y_1, \dots, y_n)$. Megmutatom az Euklideszi térbeli identitás transzformáció felbontásáról szóló tétel segítségével, hogy a Cochran-Fisher tétel feltételeinek teljesülése esetén a fenti $C^{(j)}$ mátrixok, illetve a nekik megfelelő Q_j operátorok az E_n térben, $1 \leq j \leq k$, egymásra merőleges projekciók.

A fenti jelölésekkel az (1) formula a

$$\xi I\xi^* = \xi C^{(1)}\xi^* + \cdots + \xi C^{(k)}\xi^* = \xi(C^{(1)} + \cdots + C^{(k)})\xi^* \tag{5}$$

azonosság alakjában írható, ahol I az identitás mátrixot jelöli, és a fenti azonosság egy valószínűséggel teljesül. Ez viszont csak úgy lehetséges, ha

$$xIx^* = x(C^{(1)} + \cdots + C^{(k)})x^* \tag{6}$$

minden $x \in E_n$ vektorra, ahonnan $I = C^{(1)} + \cdots + C^{(k)}$, és a $C^{(j)}$, $1 \le j \le k$, mátrixok által meghatározott Q_j szimmetrikus operátorok az E_n térben teljesítik a (3) relációt. Valóban a $T\colon (\Omega, \mathcal{A}, P) \to E_n$, $T(\omega) = (\xi_1(\omega), \dots, \xi_n(\omega))$, transzformáció abból a valószínűségi mezőből ahol a ξ_k, $1 \le k \le n$, valószínűségi változók definiálva vannak az E_n Euklideszi térbe, és azon a $(2\pi)^{-n/2} \exp\{-(x_1^2 + \cdots + x_n^2)/2\}\, dx_1 \dots dx_n$ mértékkel mértéktartó leképezés. Ezért az (5) azonosságból következik, hogy a (6) azonosság érvényes majdnem minden $x = (x^{(1)}, \dots, x^{(n)}) \in E_n$ pontban a standard normális vagy a Lebesque mérték szerint. Ez viszont csak úgy lehetséges, ha a (6) azonosság minden $x \in R^n$ vektorra igaz, tehát $I = Q_1 + \cdots + Q_n$.

Továbbá, ha a (2) formula érvényes egyenlőséggel, akkor a (4) formula is érvényes egyenlőséggel. Tehát a Cochran–Fisher tétel feltételeinek teljesülése esetén a Q_j, $1 \le j \le k$, operátorokra teljesülnek az Euklideszi térbeli identitás transzformáció felbontásáról szóló tétel feltételei. Ezért e tétel alapján a Q_j operátorok egymásra merőleges projekciók.

Azt a tényt, hogy a Q_j operátorok egymásra merőleges projekciók felírom számunkra hasznosabb alakban.

Vezessük be az $S_0 = 0$ és $S_j = \sum_{l=1}^{j} \operatorname{rang}(Q_l)$, $1 \le j \le k$, számokat, és jelölje B_j, a Q_j operátor képterét. Válasszunk $u_s = (u_{s,1}, \dots, u_{s,n}) \in E_n$, $1 \le s \le n$, egységvektorokat úgy, hogy az $S_{j-1} < s \le S_j$ indexekre, $1 \le j \le k$, $u_s \in B_j$, és az u_s, $S_{j-1} < s \le S_j$, vektorok merőlegesek egymásra. Ez lehetséges. Ekkor a Q_j projekciók merőlegessége miatt az u_s, $1 \le s \le n$ vektorok egy ortonormált bázist alkotnak az E_n térben, és az u_s vektorok koordinátáiból álló $U = (u_{s,t})$, $1 \le s, t \le n$, mátrix ortogononális. Jelölje $I_j = (a_{s,t}^{(j)})$, $1 \le s, t \le n$, azt a diagonális mátrixot, amelyre $a_{s,s}^{(j)} = 1$, ha $S_{j-1} < s \le S_j$, és $a_{s,t}^{(j)} = 0$ egyébként. Azt állítom, hogy

$$Q_j = U^* I_j U \quad \text{minden } 1 \le j \le k \text{ indexre.} \tag{7}$$

Ehhez elegendő belátni, hogy $u_s Q_j u_t^* = u_s U^* I_j U u_t^*$ minden $1 \le s, t \le n$ és $1 \le j \le k$ indexre, ahol u_s, $1 \le s \le n$, az előbb konstruált ortonormált bázis az E_n Euklideszi térben. Viszont $u_s Q_j u_t^* = 1$, ha $s = t$, és $S_{j-1} < s \le S_j$, és $u_s Q_j u_t^* = 0$ egyébként az u_s vektorok definiciója miatt. Másrészt, mivel U ortogonális transzformáció ezért $U^* = U^{-1}$, ahonnan az U mátrix definiciója miatt $u_s U^* = e_s$ minden $1 \le s \le n$ indexre, ahol e_s az a vektor, amelyiknek az s-ik koordinátája 1, és az összes többi koordinátája 0. Innen az is következik, hogy $U u_t^* = e_t^*$, és $u_s U^* I_j U u_t^* = e_s I_j e_t^*$. Viszont $e_s I_j e_t^* = 1$, ha $s = t$, és $S_{j-1} < s \le S_j$, és $e_s I_j e_t^* = 0$ egyébként. A fenti relációkból következik az (7) állítás.

Bebizonyítom az (7) formula segítségével a Cochran–Fisher tételt. Ennek érdekében definiáljuk az $\eta = (\eta_1, \dots, \eta_n) = \xi U^* = (\xi_1, \dots, \xi_n) U^*$ vektort. Ekkor η n-dimenziós normális eloszású vektor, amelynek koordinátái korrelátlanok és 1 szórásúak, ezért független standard normális eloszlású valószínűségi változók. Másrészt az (7) formula alapján

$$Q_j(\xi_1, \dots, \xi_n) = \xi Q_j \xi^* = \xi U^* I_j U \xi^* = \eta I_j \eta^* = \sum_{s=S_{j-1}+1}^{S_j} \eta_s^2 \quad \text{minden } 1 \le j \le k \text{ indexre.}$$

Ezzel a $Q_j = Q_j(\xi_1, \dots, \xi_n)$ kifejezéseknek megtaláltuk a kívánt $Q_j = \sum\limits_{s=1}^{r_j} \zeta_{j,s}^2$, $1 \le j \le k$, alakú előállítását, ahol $\zeta_{j,1}, \dots, \zeta_{j,r_j}$, $r_j = \text{rang}\,(Q_j)$, $1 \le j \le k$, független standard normális eloszlású valószínűségi változók. Egyszerűen, legyen $\zeta_{j,s} = \eta_{S_{j-1}+s}$ minden $1 \le j \le k$, $1 \le s \le r_j$ indexre. Végül vegyük észre. hogy

$$\sum_{j=1}^{k} Q_j(\xi_1, \dots, \xi_n) = \sum_{j=1}^{k} \sum_{s=1}^{r_j} \zeta_{j,s}^2 = \sum_{s=1}^{n} \eta_s^2 = (\eta, \eta) = \xi U^* U \xi^* = \sum_{s=1}^{n} \xi_s^2 = Q(\xi_1, \dots, \xi_n).$$

Egy Euklideszi térbeli identitás transzformáció felbontásáról szóló tétel bizonyítása. Tekintsük minden $1 \le j \le k$ indexre a Q_j leképezés képterének egy $r_j = \text{rang}\,(Q_j)$ elemből álló $e_{j,1}, \dots, e_{j,r_j}$ bázisát. Ha a $\sum\limits_{j=1}^{k} r_j < n$ reláció teljesülne, akkor létezne olyan $e \in E_n$ nem zéró vektor, amelyre $(e, e_{j,s}) = 0$ minden $1 \le j \le k$, $1 \le s \le r_j$ indexre. Egy ilyen e vektorra a $(e, Q_j e) = 0$ azonosság teljesül minden $1 \le j \le k$ indexre. Ez viszont ellentmond a (3) azonosságnak, mivel innen azt kapjuk, hogy $0 < (e, Ie) = \sum\limits_{j=1}^{k} (e, Q_j e) = 0$, ami ellentmondás.

Ha a $\sum\limits_{j=1}^{k} r_j = n$ azonosság érvényes, akkor tekintsük a Q_j leképezések képterének egy $e_{j,1}$, ... ,e_{j,r_j} bázisát minden $1 \le j \le k$ indexre. Vegyük e vektorok közül az $e_{j,s}$, $1 \le j \le k-1$, alakúakat, és vezessük be e vektorok által generált B alteret, valamint annak C ortogonális kiegészítő alterét az E_n Euklideszi térben. Azt állítom, hogy a C altér dimenziója r_k. Valóban, ha tekintjük az összes $e_{j,1}$, ... , e_{j,r_j}, $1 \le j \le k$, alakú vektort, akkor ezek egy n elemű generátor rendszert, ezért bázist alkotnak az E_n térben. Ezért az $e_{j,1}$, ... , e_{j,r_j}, $1 \le j \le k-1$, alakú vektorok lineárisan függetlenek, így az általuk generált B altér $n - r_k$, és annak C ortogonális kiegészítője r_k dimenziós.

Mivel mind a Q_k mátrix rangja, mind a C altér dimenzója r_k, és $Q_j x = 0$, ha $1 \le j \le k-1$, és $x \in C$, ezért $Q_k x = x - \sum\limits_{j=1}^{k-1} Q_j x = x$, ha $x \in C$, továbbá $Q_k x \in C$

minden $x \in R^n$ vektorra, tehát speciálisan akkor is, ha $x \in B$. Azt állítom, hogy ebből következik, hogy $Q_k x = 0$ $x \in B$ esetén. Valóban, ha egy $x \in B$ vektorra $Q_k x = y$ lenne valamely $y \neq 0$ (és $y \in C$) vektorral, akkor a Q_k mátrix szimmetrikus tulajdonsága miatt $(y, y) = (Q_k x, y) = (x, Q_k y) = (x, y) = 0$ lenne. (A számolás utolsó lépésében kihasználtam, hogy $x \in B$ és $y \in C$, ezért az x és y vektorok merőlegesek egymásra.)

Innen következik, hogy a Q_k operátor megegyezik a C altérre való ortogonális vetítéssel. Hasonlóan bizonyítható, hogy az összes Q_j, $1 \leq j \leq k$, operátor egy r_j dimenziós projekció. Továbbá, mivel $Q_j x \in B$ minden $x \in E_n$ vektorra, ha $1 \leq j \leq k-1$, és $Q_k y \in C$ minden $y \in E_n$ vektorra, ezért $(Q_j x, Q_k y) = 0$. Ez azt jelenti, hogy a Q_k projekció merőleges minden Q_j, $j \neq k$ projekcióra. Hasonlóan látható a többi projekció ortogonalitása is. A tételt bebizonyítottuk.

A következő tételben megfogalmazok és bebizonyítok néhány további a variancia analízisben hasznos állítást.

Tétel normális eloszlású véletlen vektorok kvadratikus alakjainak néhány hasznos tulajdonságáról. *Legyen adva egy n-dimenziós $X = (X_1, \dots, X_n)$ standard normális eloszlású valószínűségi változó.*

a.) Tekintsünk egy A szimmetrikus $n \times n$-es mátrixot. Az XAX^* kifejezés akkor és csak akkor χ-négyzet eloszlású valószínűségi változó, ha teljesül az $A^2 = A$ azonosság, azaz A projekció. Ha ez a feltétel teljesül, akkor XAX eloszlása rang(A) *szabadságfokú χ-négyzet eloszlás.*

b.) Ha A és B két szimmetrikus mátrix, amelyre $A^2 = A$, $B^2 = B$, és $AB = 0$, akkor az XAX^* és XBX^* valószínűségi változók függetlenek, és χ-négyzet eloszlásúak rang(A) *és* rang(B) *szabadságfokkal.*

c.) Ha Q, Q_1, Q_2 szimmetrikus $n \times n$ méretű mátrixok, XQX^* és XQ_1X^* χ-négyzet eloszlású valószínűségi változók, $Q = Q_1 + Q_2$, és Q_2 pozitív szemidefinit mátrix, akkor XQ_1X^* és XQ_2X^* független χ-négyzet eloszlású valószínűségi változó.

A tétel bizonyítása. Ha az A szimmetrikus mátrix teljesíti az $A^2 = A$ feltételt, akkor projekció, és az XAX^* kifejezést a főtengelytranszformáció segítségével felírhatjuk $\sum\limits_{j=1}^{\mathrm{rang}\,(A)} Y_j^2$ alakban, ahol Y_j, $1 \leq j \leq \mathrm{rang}\,(A)$ független, standard normális eloszlású valószínűségi változók. Általános szimmetrikus A mátrix esetében az XAX^* kifejezést $\sum\limits_{j=1}^{n} \lambda_j Y_j^2$ alakban írhatjuk a főtengelytranszformáció segítségével, ahol Y_j, $1 \leq j \leq n$, független, standard normális eloszlású valószínűségi változók, és $\lambda_1, \dots, \lambda_n$ az A mátrix sajátértékei. Ez a kifejezés akkor és csak akkor χ-négyzet eloszlású, ha az összes λ_j sajátérték 0 vagy 1 értéket vesz fel. (Ez például az eloszlás karakterisztikus függvényének felírásából látszik.) Ez a tulajdonság viszont akkor és csak akkor teljesül, ha A projekció, azaz $A^2 = A$.

A tétel b) pontjának bizonyításához elég azt észrevenni, hogy az adott feltételek mellett A és B két ortogonális projekció. Innen a Cochran–Fisher tétel bizonyításának alkalmazásával be lehet bizonyítani a b) pont állítását, de lehet a Cochran–Fisher tételt közvetlenül is alkalmazni az n-dimenziós E_n tér I identitás transzformációjának alkalmas $I = C+B+A$ alakú reprezentációjának felhasználásával. Ebben a reprezentációban C az $A+B$ projekció képterének az E_n tér $n-\operatorname{rang}(A)-\operatorname{rang}(B)$ dimenziós ortogonális kiegészítésére vett projekció.

A tétel c) pontjának bizonyítása érdekében mutassuk meg először azt, hogy az adott feltételek mellett a Q projekció B képtere tartalmazza a Q_1 projekció A képterét. Valóban, ha létezne egy $x \in A \setminus B$, $x \neq 0$, vektor, akkor erre az x vektorra teljesülne a $(Q_2x, x) = (Qx, x)-(Q_1x, x) = (y, x)-(x, x) = (y, y)-(x, x) < 0$ egyenlőtlenség, ahol y az x vektornak (az x vektornál rövidebb) Qx vetülete a B altérre. Ez az egyenlőtlenség viszont ellentmond annak a feltételnek, hogy Q_2 pozitív szemidefinit mátrix.

Abból, hogy a Q_1 projekció képtere része a Q projekció képterének következik, hogy $Q_1 = Q_1Q = QQ_1$, ezért $Q_2^2 = (Q-Q_1)^2 = Q+Q_1-2Q_1 = Q-Q_1 = Q_2$, és $Q_1Q_2 = Q_1(Q-Q_1) = 0$. Ez azt jelenti, hogy Q_1 és Q_2 ortogonális projekciók, és a b) rész állításából következik a c) rész állítása.

A Cochran–Fisher tétel egy alkalmazásaként megmutatom, hogy hogyan következik ebből az eredményből és a többváltozós centrális határeloszlástételből a matematikai statisztika egyik fontos határeloszlástétele, amely a χ-négyzet módszer alapjául szolgál. A következő eredmény bizonyítását fogom ismertetni.

Egy χ-négyzet határeloszlástétel. *Legyen adva r darab urna, és dobjunk ezekbe egymástól függetlenül n darab golyót egymástól függetlenül úgy, hogy mindegyik golyó p_j valószínűséggel esik a j-ik urnába, ahol $p_j \geq 0$, $j = 1, \dots, r$, $\sum\limits_{j=1}^{r} p_j = 1$ előírt számok. Jelölje $\nu_n(j)$, $1 \leq j \leq r$, a j-ik urnába eső golyók számát. Ekkor a $\sum\limits_{j=1}^{r} \frac{(\nu_n(j)-np_j)^2}{np_j}$ valószínűségi változók eloszlásban konvergálnak az $r-1$ szabadságfokú χ-négyzet eloszláshoz, ha $n \to \infty$.*

E tételt a Cochran–Fisher tételből és a többváltozós centrális határeloszlástétel alábbi következményéből fogjuk levezetni.

A többváltozós centrális határeloszlástétel egy következménye. *Tekintsük az előbb megfogalmazott χ-négyzet határeloszlástételben leírt modellt és az abban definiált $\nu_n(j)$, $1 \leq j \leq r$, valószínűségi változókat. Az $\frac{1}{\sqrt{n}}(\nu_n(1)-np_1, \dots, \nu_n(r)-np_r)$ véletlen vektorok eloszlásban konvergálnak egy olyan normális eloszlású $(Z_1, \dots, Z_r)$ véletlen vektorhoz, amelyre $EZ_j = 0$, $EZ_j^2 = p_j(1-p_j)$, $1 \leq j \leq r$, és $EZ_iZ_j = -p_ip_j$, $1 \leq i < j \leq r$, ha $n \to \infty$.*

Az előbb megfogalmazott χ-négyzet határeloszlástétel következik a fenti határeloszlástételből és abból az állításból, hogy egy olyan normális eloszlású $(Z_1, \dots, Z_r)$ véletlen vektorra, amelyre $EZ_j = 0$, $EZ_j^2 = p_j(1-p_j)$, $1 \leq j \leq r$, és $EZ_iZ_j = -p_ip_j$, ha $1 \leq i < j \leq r$, a $\sum\limits_{j=1}^{r} \frac{Z_j^2}{p_j}$ valószínűségi változó χ-négyzet eloszlású $r-1$ szabadságfokkal. Ez utóbbi állítás a Cochran–Fisher tételből és egy a $(Z_1, \dots, Z_r)$ véletlen vektorral megegyező eloszlású véletlen vektor alább ismertetett konstrukciójából következik.

Alkalmas kovarianciájú véletlen normális vektorok konstrukciója. *Legyenek adva $Y_1,\dots,Y_n$ független, standard normális valószínűségi változók és olyan pozitív p_j, $1\le j\le r$, valós számok, amelyekre $\sum\limits_{j=1}^{r} p_j=1$. Definiáljuk az*

$$U=\sqrt{p_1}Y_1+\sqrt{p_2}Y_2+\cdots+\sqrt{p_r}Y_r$$

és

$$Z_j=\sqrt{p_j}Y_j-p_jU,\quad 1\le j\le r,$$

valószínűségi változókat. Ekkor $EZ_j=0$, $EZ_j^2=p_j(1-p_j)$, és $EZ_iZ_j=-p_ip_j$, ha $1\le i<j\le r$. Ezenkívül teljesül a

$$\sum_{j=1}^{r}\frac{Z_j^2}{p_j}+U^2=\sum_{j=1}^{r}Y_j^2 \tag{8}$$

és $\sum\limits_{j=1}^{r} Z_j=0$ azonosság.

A fenti eredményből és a Cochran–Fisher tételből következik az, hogy az előbb felírt $\sum\limits_{j=1}^{r}\frac{Z_j^2}{p_j}$ összeg χ-négyzet eloszlású $r-1$ szabadságfokkal. Valóban, a (8) összegben szereplő $\sum\limits_{j=1}^{r}\frac{Z_j^2}{p_j}$ kifejezés a független standard normális eloszlású $Y_1,\dots,Y_r$ valószínűségi változók kvadratikus kifejezése, amelynek rangja a $\sum\limits_{j=1}^{r} Z_j=0$ azonosság miatt legfeljebb $r-1$. Ez a rang legfeljebb $r-1$, mert az előbb említett azonosságból következik, hogy $(Z_1,\dots,Z_r)=(Y_1,\dots,Y_r)D$ egy olyan D mátrix-szal, amelynek rangja kisebb vagy egyenlő, mint $r-1$. Ezért a a $Z=(Z_1,\dots,Z_r)$ vektor által definiált tetszőleges ZBZ^* kvadratikus alak $ZBZ^*=YDBD^*Y^*$ alakban írható, ahol $Y=(Y_1,\dots,Y_r)$, és $\operatorname{rang}(DBD^*)\le\operatorname{rang}(D)\le r-1$. Ezenkívül U^2 egy olyan kvadratikus kifejezése az $Y_1,\dots,Y_r$ valószínűségi változóknak, amelynek rangja 1. Ezért a (8) azonosságból és a Cochran–Fisher tételből következik, hogy a $\sum\limits_{j=1}^{r}\frac{Z_j^2}{p_j}$ kifejezés $r-1$ szabadságfokú χ-négyzet eloszlású valószínűségi változó. (Ezenkívül azt is tudjuk, hogy ez független az U^2 valószínűségi változótól. Sőt a Z_j, $1\le j\le r$, valószínűségi változók függetlenek az U valószínűségi változótól. De erre a tényre nincs szükségünk).

Felmerülhet a kérdés, hogy hogyan találhatjuk meg a $(Z_1,\dots,Z_r)$ véletlen vektor fenti, számunkra hasznos reprezentációját természetes módon. Ezt magyarázom el a következő megjegyzésben.

Megjegyzés. Tekintsünk egy $B(t)$, $0\le t\le 1$, Brown bridge-t, és vezessük be az $t_0=0$, $t_j=\sum\limits_{l=1}^{j}p_l$, $1\le j\le r$, számokat. E mennyiségek segítségével a keresett eloszlású $(Z_1,\dots,Z_r)$ vektort $Z_j=B(t_j)-B(t_{j-1})$, $1\le j\le r$, alakban állíthatjuk elő. Legyen adva a $B(t)$ Brown bridge-nek egy $B(t) = W(t) - tW(1)$, $0 \le t \le 1$, alakú előállítása egy $W(t)$, $0 \le t \le 1$, Wiener folyamat segítségével. Ekkor bevezetve az $Y_j = p_j^{-1/2}(W(t_j) - W(t_{j-1}))$, $1 \le j \le r$, független, standard normális eloszlású valószínűségi változókat, felírhatjuk az $U = W(1) = \sum_{j=1}^{r} \sqrt{p_j} Y_j$ és $Z_j = \sqrt{p_j} Y_j - p_j U$, $1 \le j \le r$ azonosságokat. Ezt az eljárást követtük a fenti konstrukcióban.

Az alkalmas kovarianciájú véletlen normális vektorok konstrukciójában megfogalmazott eredmények bizonyítása. Az $EZ_j = 0$, $1 \le j \le r$, reláció nyilvánvaló.

$$
\begin{aligned}
EZ_j^2 &= (p_j^{1/2} - p_j^{3/2})^2 EY_j^2 + \sum_{l:\ 1 \le l \le r,\ l \ne j} p_j^2 p_l EY_l^2 \\
&= p_j(1-p_j)^2 + \sum_{l:\ 1 \le l \le r,\ l \ne j} p_j^2 p_l = p_j(1-p_j)^2 + (1-p_j)p_j^2 = p_j(1-p_j),
\end{aligned}
$$

$$
\begin{aligned}
EZ_iZ_j &= p_ip_j \sum_{l:\ 1 \le l \le r,\ l \ne i,j} p_l - p_ip_j(1-p_i)EY_i^2 - p_ip_j(1-p_j)EY_j^2 \\
&= p_ip_j(1-p_i-p_j) - p_ip_j(2-p_i-p_j) = -p_ip_j,
\end{aligned}
$$

ha $i \ne j$, és

$$
\begin{aligned}
\sum_{j=1}^{r} \frac{Z_j^2}{p_j} + U^2 &= \sum_{j=1}^{r} (Y_j - \sqrt{p_j}U)^2 + U^2 = \sum_{j=1}^{r} Y_j^2 + \sum_{j=1}^{r} p_j U^2 - 2U \sum_{j=1}^{r} \sqrt{p_j} Y_j + U^2 \\
&= \sum_{j=1}^{r} Y_j^2 + U^2 - 2U^2 + U^2 = \sum_{j=1}^{r} Y_j^2.
\end{aligned}
$$

Végül

$$
\sum_{j=1}^{r} Z_j = \sum_{j=1}^{r} \sqrt{p_j} Y_j - \sum_{j=1}^{r} p_j U = U - U = 0.
$$

A kívánt eredményeket bebizonyítottuk.