

Néhány többváltozós kemometriai módszer alkalmazása műszeres analitikai vizsgálatok értékelésére

*Dalmadi István, Seregély Zsolt, Kaffka Károly, Farkas József**

Budapesti Corvinus Egyetem, Élelmiszertudományi Kar

*Központi Élelmiszer-tudományi Kutatóintézet

Érkezett: 2007. szeptember 4.

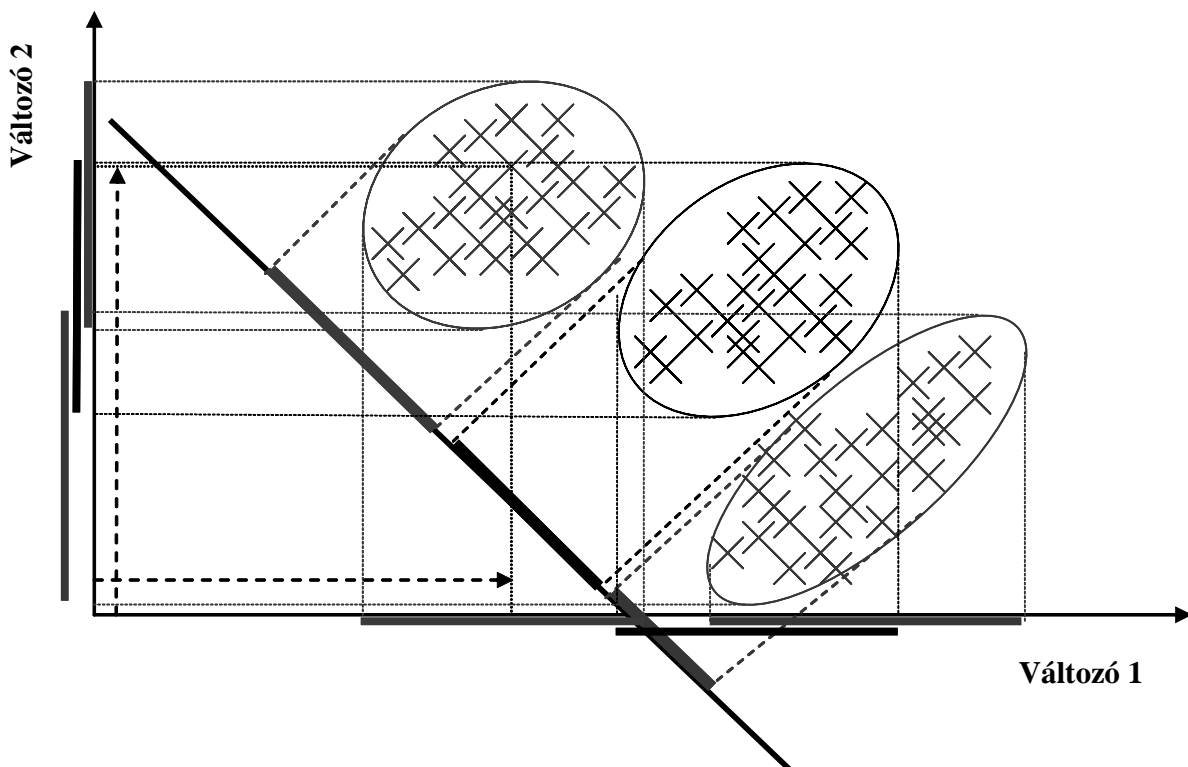
Mezőgazdasági termékek és élelmiszerek minőségét sokféleképpen, különböző jellemzők megadásával írhatjuk le. A technika rohamos fejlődésének köszönhetően a műszeres analitikai módszerek fejlesztése – megismerése – alkalmazása egyre inkább teret hódít. Egyrészt a klasszikus módszerek automatizálásával olyan mérőműszereket alakítanak ki, amelyekkel biztosítható a reprodukálhatóság, a pontosság és a gyors információszerzés. Másrészt indirekt módszerek terjednek el, amelyek másodlagos, legtöbb esetben fizikai paraméterek mérésén és azok kémiai jellemzőkkel történő korreláltatásán keresztül következtetnek a minta tulajdonságaira. E korrelatív módszernek egyik megtestesítője az ún. NIR technika (Ozaki, 2006), mely a közeli infravörös tartományban mért spektrumértékekből összetételi jellemzők gyors, pontos, roncsolásmentes meghatározására képes, környezetkímélő és minta-kímélő módon. A NIR technika az 1990-es évek második felére elérte azt az analitikai érettségi szintet, ahol a felhasználó számára inkább a vizsgálat eszközeként, mint tárgyaként szolgál, ugyanakkor még számos olyan terület van, ahol felhasználására még most folynak kísérletek. Az indirekt vizsgálati módszerek egy másik csoportját a még kevésbé ismert kémiai érzékelősorok alkotják (Pearce, 2003), amiket legtöbbször elektronikus orrként és nyelvként említ a szakirodalom. Az elektronikus orr érzékelői nagy érzékenységek, de kis specifitást mutatnak az egyes egyedi molekulákra és az illékony komponensekre különböző jelválaszokat adnak, a gázkeverék ún. „ujjlenyomatát” hozzák létre.

A NIR és az elektronikus orr technikára ugyanúgy jellemző, hogy a vizsgált mintáról rövid idő alatt nagymennyiségű információt szolgáltat, melynek értékelése elképzelhetetlen matematikai-statisztikai módszerek nélkül (Adams, 2004). A többváltozós statisztikai módszerek rutinszerű alkalmazását a hardver és a szoftver területeken bekövetkezett fejlődésnek kellett megelőznie, így ezek a módszerek csak mostanában válnak szélesebb körben ismertté. Jelen közleményünkben saját NIR és kémiai

szenzorsorokkal kapott eredményeinken keresztül mutatunk be néhány példát a többváltozós statisztikai módszerekből.

Kvalitatív módszerek

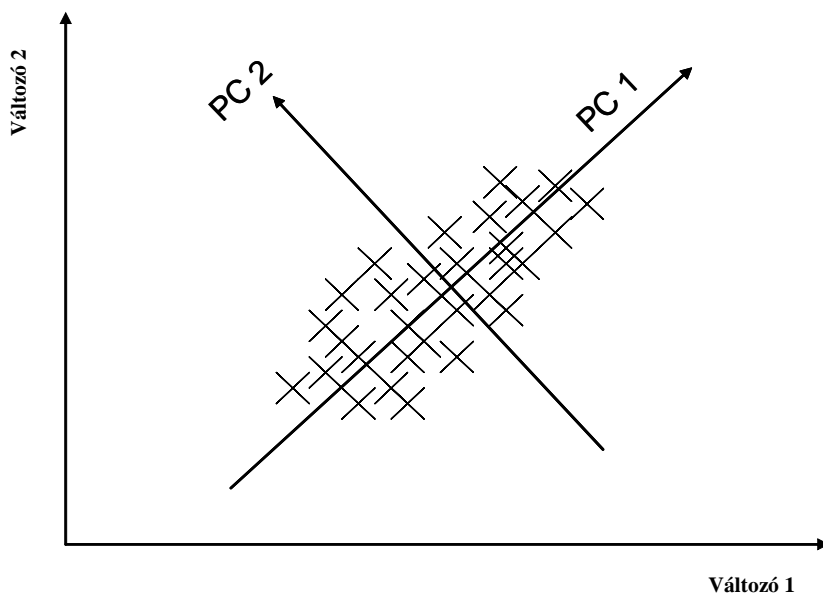
A tudomány területén is követelmény a megfigyelt dolgok - azok hasonlósága alapján történő - rendszerezése. A vizsgálati adatok csoportosítása nélkül az összegyűjtött nagy mennyiségű tudományos információ nem több az adatok egyszerű halmazánál. Habár az emberi agy kiválóan felismeri az alakokat és alakzatokat, nehezen birkózik meg a tudományos életben sokszor keletkező terjedelmes adattáblák értékelésével. Megoldást jelent a számsorokból képezett grafikon, azonban 2-3 változónál többet egyszerre nem lehet szemléletesen ábrázolni. A nagy mennyiségű információ feldolgozására olyan statisztikai technikákat kell keresni és alkalmazni, amelyek hatékonyan emelik ki a hasonlóságokat és a különbségeket az egyes csoportok között. A kvalitatív statisztikai módszerek alkalmazásánál olyan matematikai modellt kell származtatnunk az objektumok csoportosításához, hogy az egy csoporton belül lévők jobban hasonlítsanak egymáshoz, mint a másik csoport elemeihez (1. ábra). A hasonlóság/különbözőség megítéléséhez más-más algoritmus használható, s így eltérő statisztikai eljárások születtek.



1. ábra: A csoportok elkülönítése a többdimenziós térben

A különböző osztályozó technikákat csoportosíthatjuk aszerint, hogy előzetesen ismertek-e, hogy az egyes minták mely csoportba tartoznak vagy sem. Ez alapján beszélhetünk felügyelt (supervised) és nem felügyelt (unsupervised) módszerről.

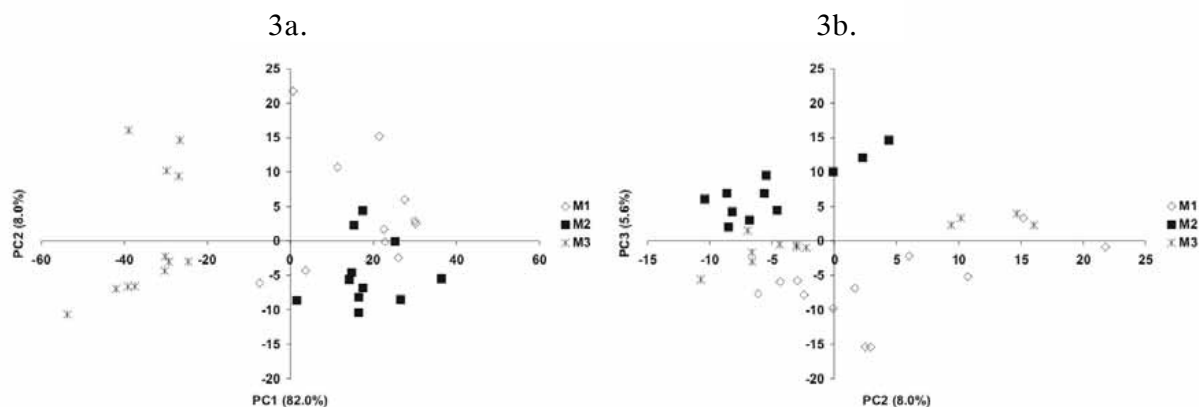
A **főkomponens-analízis** (*Principal Component Analysis, PCA*) egy nem felügyelt (unsupervised), lineáris módszer. Segítségével felmérhető, hogy a megfigyelési egységek között természetes módon alakulnak-e ki csoportok. Elvégzi a kiindulási adatmátrix dimenziójának csökkentését úgy, hogy a teljes adatmátrixot viszonylag jól reprezentáló, kisszámú háttérváltozó segítségével írja le. Ehhez a változók által kijelölt sokdimenziós térben olyan irányokat keres, amelyek irányában a mérési eredmények varianciája a legnagyobb (**2. ábra**). Az első főkomponens (PC) magyarázza meg a variancia legnagyobb részét. A második főkomponens ortogonális az elsőre, vagyis független attól. Grafikusan ez annyit tesz, hogy a főkomponensek merőlegesek egymásra. A második főkomponens a megmaradt variációból annyit magyaráz, amennyi csak lehetséges. Ez így folytatódik mindaddig, amíg az összes variancia el nem fogy. A műveletsorral kiemelhetők a releváns információt hordozó főkomponensek, melyek a legnagyobb variációt adják, és kizárhatjuk a véletlenszerű zajt.



2. ábra: A főkomponens-analízis elve

Számos esetben a képzett új változóknak – főkomponenseknek – egyedi jelentést adhatunk. A **3. ábrán** instant kávé beszállítói tételei között elektronikus orral mért különbségek láthatók. A csoportok elhelyezkedése alapján elmondható, hogy az PC1-PC2 főkomponens-síkon (**3a. ábra**) a három vizsgált tétel közül csak egy különült el (M3), a másik két csoport

átfedést mutatott egymással. Ugyanakkor a következő főkomponens-síkot (PC2-PC3) megvizsgálva (**3b. ábra**) megfigyelhető, hogy az M1 jelű minta jobb elkülönülése mellett a másik két csoport is viszonylag jól megkülönböztethető volt egymástól. Ez azt jelenti, hogy a mért illó komponensek belső szórásában lehetett találni olyan komponenst, melyet a PC3 főkomponens jellemez, amelynek alapján az azonosnak minősülő termékek illata is megkülönböztethető. Ez a főkomponens az összes variancia 5,6 %-át írja le és alkalmas arra, hogy a homogénnek minősülő tételekben levő esetleges különbségek beazonosíthatók legyenek.

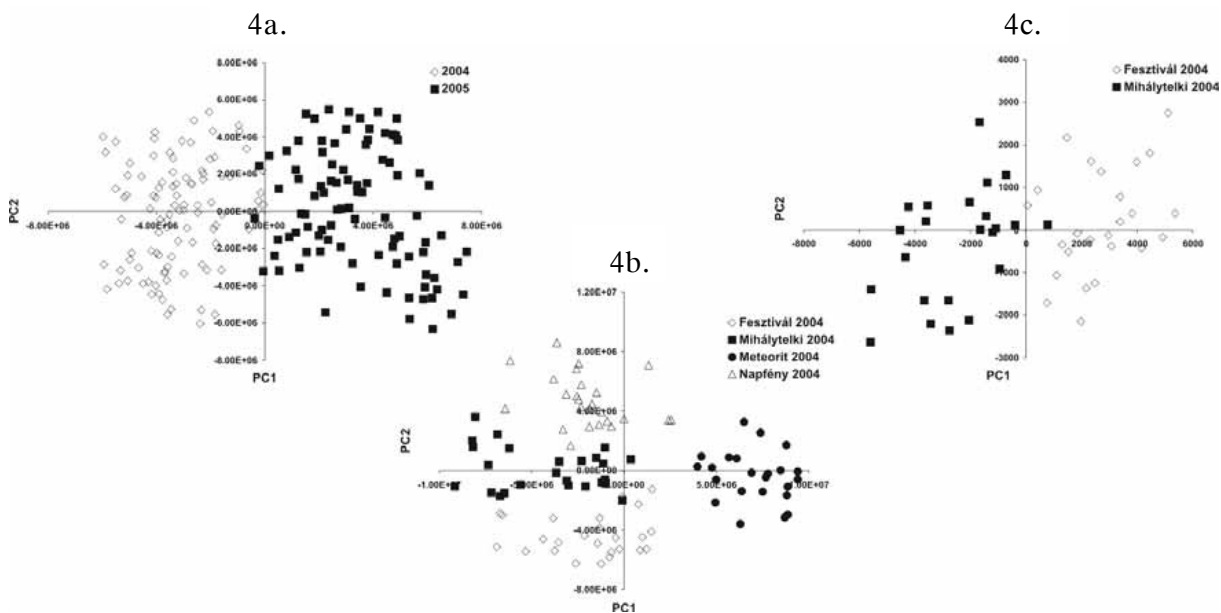


3. ábra: Főkomponens-analízis: Instant kávé különböző tételeinek összehasonlítása elektronikus orral 20°C-on
(3a.: 1.-2. főkomponens, 3b.: 2.-3. főkomponens) (Somogyi, 2005)

Ha nem az összes csoport válik szét az első főkomponensek szerint, a teljes variancia kisebb részét leíró főkomponensek próbálgatása helyett hierarchikus osztályozás is megvalósítható. A **4. ábrán** különböző évjáratú és fajtájú paprikamagok elkülönítése látható NIR spektrumuk alapján hierarchikus osztályozással. A főkomponens-analízist elvégezve első lépésben a két évjárat különült el (**4a. ábra**). Kiválasztva a 2004. év mintáit új projekciós síkot képeztünk a fajták elkülönítéséhez. Látható, hogy a Napfény és a Meteorit fajták jól elkülönülnek, de a Mihálytelki és a Fesztivál nem (**4b. ábra**). Ehhez egy harmadik sík szükséges. A **4c. ábrán** látható, hogy csak a két csoportra elvégzett főkomponens-analízissel e két csoport is jól elkülönült.

A **klaszter-analízis** (*Cluster Analysis*) szintén nem felügyelt módszer. Az eljárás a teljes mintából indul ki és a minta egyedeit próbálja viszonylag homogén csoportokba rendezni. Tehát az eljárás indulásánál még nem rendelkezünk csoportokkal, viszont az eljárás végére igen. Itt is szem előtt kell tartani, hogy a csoporton belüli objektumok jobban hasonlítsanak egymásra mint a más csoportba tartozókhöz. Ehhez meg kell határozni, hogy mit értünk távolságon/hasonlóságon (euklidészi, Mahalanobis-

távolság) és milyen döntési függvényt használunk a csoportok szétválasztásához (két legközelebbi pont távolsága; két legmesszebb lévő pont távolsága; csoportközéppontok távolsága stb.). Fontos, hogy a részhalmazokra történő bontásnál a csoportoknak ne legyenek közös elemei, valamint minden elem legyen besorolva valamelyik csoportba. A hierarchikus klaszterezés algoritmus egyelemű klaszterekből indul ki. Minden lépésben a két legközelebb fekvő klasztert összevonva csökkenti a klaszterek számát, amíg minden eset egyetlen klaszterbe nem kerül. A folyamatot ún. dendogrammon követhetjük nyomon. A módszer megfigyelési egységek és változók közötti összefüggések vizsgálatára is alkalmas.



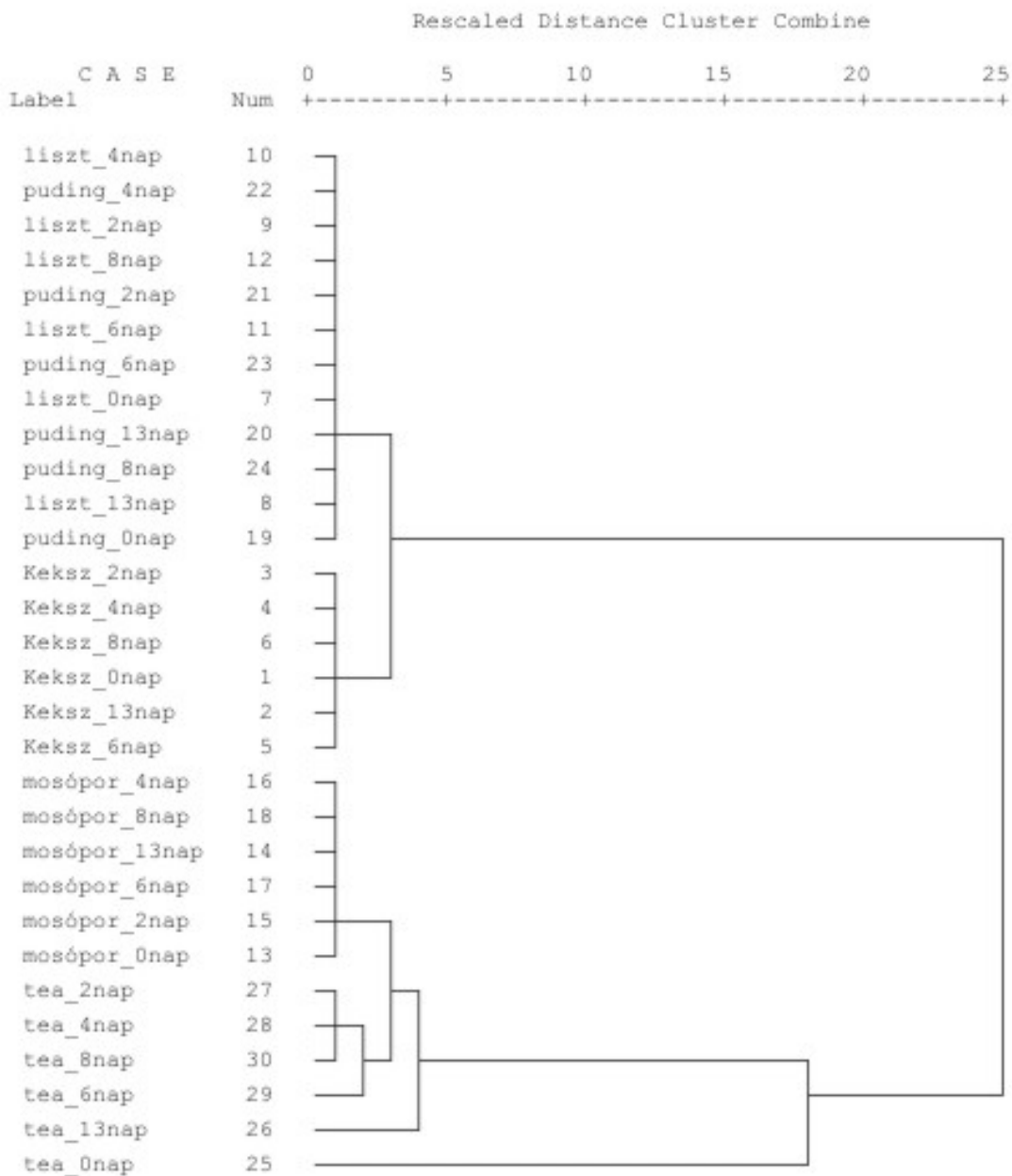
4. ábra: Főkomponens-analízis: paprikamagok hierarchikus osztályozása NIR spektrumuk alapján (Kaffka, 2007)

Vizgáltuk, hogy az élelmiszerek más – sok esetben erős illattal rendelkező - termék melletti tárolásakor megjelennek-e idegen illatok az élelmiszereken vagy sem. Az egymás mellé tett termékekből több tárolási napon is mintát vettünk és elektronikus orral vizsgáltuk. A dendogrammon (5. ábra) látható, hogy a becsomagolt élelmiszerek mindegyik tagja saját csoportjának más-más időpontban mért értékéhez hasonlít legjobban (azokkal vonta össze az algoritmus leghamarabb). Az egyetlen eltérés a mosópor-filteres tea kombinációnál született, ahol is a filteresdobozt védő celofánt a tárolás előtt eltávolítottuk. A teafilter jó illatmegkötőnek bizonyult, hiszen csak a 0. napi teaillat különült el határozottan. A teafilter valamennyi tárolt tagját előbb vonta össze az algoritmus a mosóporéval, mint a saját 0. napi értékével. Ez felhívja a figyelmet arra, hogy egyes élelmiszerek hajlamosabbak az illatanyagok megkötésére, és a helytelen

tárolás során (pl. a csomagolóanyag megsérülése) ezek a termékek képesek a környezetből adszorbeálni az illatanyagokat.

***** H I E R A R C H I C A L C L U S T E R A N A L Y S I S *****

Dendrogram using Average Linkage (Between Groups)

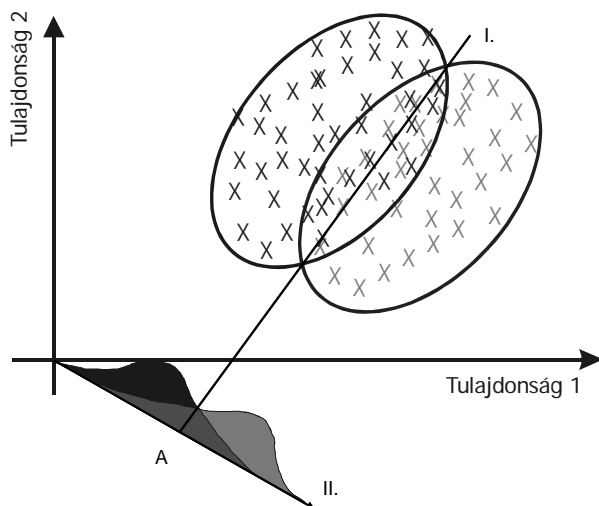


5. ábra: Klaszter analízis: különböző termékek egymás mellett tárolhatóságának vizsgálata elektronikus orral (Balla, 2006a.)

A nem felügyelt módszerekkel szemben hatékonyabb szétválasztás érhető el, ha felügyelt ún. supervised módszert alkalmazunk. A felügyelt módszerek végrehajtásakor előre definiált csoportok vizsgálatára kerül sor. Az egyes csoportokat kialakíthatjuk szakmai ismeret alapján vagy más többváltozós módszer outputja szerint (pl. klaszter-analízis).

Felügyelt módszer a szintén lineáris **kanonikus diszkriminancia-analízis** (*Canonical Discriminant Analysis, CDA*) is. A feladat a főkomponens elemzéssel szemben, hogy nem a teljes varianciát, hanem a csoportok közötti varianciát maximáljuk, miközben a csoporton belülit minimalizáljuk. Az algoritmus a megfigyeléseket a mintatérből egy olyan diszkrimináló térbe viszi át, ahol a csoportok a lehető legjobban elkülönülnek, és kiválasztja azokat a változókat, amelyek a csoportok különbözőségét határozottan magyarázzák. A szeparáló eljárás azokat a hiperfelületeket keresi, amelyek elválasztják egymástól a minta osztályait feltételezve, hogy az azonos osztályokban szereplő elemek “közel”, a különböző osztályokban szereplők pedig távol helyezkednek el egymástól.

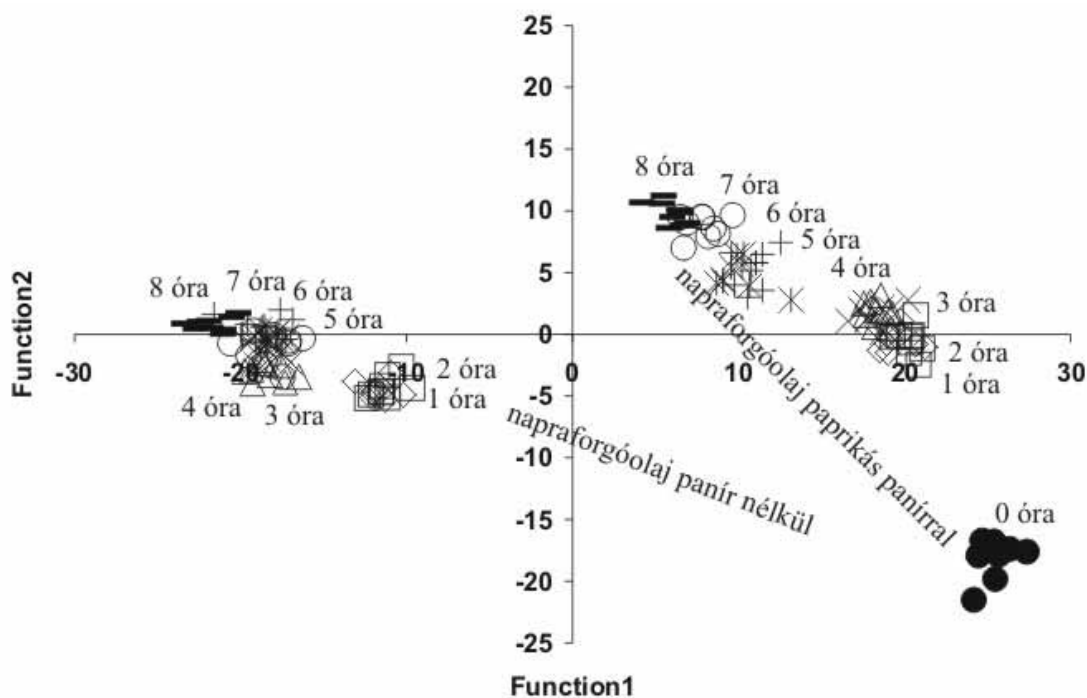
A módszer lényegében a következőképpen jár el (**6. ábra**): a két halmazt körülvevő szórás-ellipszis metszéspontjain át egyenest (I.) fektet, majd erre az origón átmenő merőleges egyenest illeszt. Ha a pontokat a II. egyenesre vetítjük, akkor a két csoport egyváltozós eloszlása közötti átfedés kisebb lesz, mint bármilyen más egyenes esetén. Az I. egyenes segítségével osztható a minta két csoportba. Ha a szórás-ellipszisek átfednek az alkotott modell nem adja vissza az összes minta eredeti csoportba tartozását. A szétválasztás jósága annak alapján vizsgálható, hogy az elemek hányad részét sikerült helyesen besorolni.



6. ábra: A diszkriminancia-analízis elve

A **7. ábra** a növényi olajok sütés közbeni változását mutatja elektronikus orral vizsgálva. Látható, hogy panírral és panír nélkül végzett sütés egyedi karakterű illóanyagokat szabadított fel az olajban, ugyanakkor a sütési idő előrehaladtával mindkét csoport azonos irányban mozdult el. Az ábra alatt feltüntetett %-os adatok a modellalkotás és a keresztvalidáció jóságát mutatják. Jelzik, hogy viszonylag pontosan soroltuk vissza a mintákat a nekik megfelelő csoportba. A keresztvalidációval ellenőrizhető a modell

megbízhatósága (robosztussága) abban az esetben, ha nem áll rendelkezésre független mintaszett a validáláshoz. Ekkor nem választjuk szét a kalibráló és a validáló mintákat, hanem a mintákat adott tagszámú csoportokba osztjuk, és ezen csoportok közül egyet kihagyunk a kalibrációból, míg a maradék mintákra elvégezzük a kalibrációt. Az így nyert kalibrációs egyenlettel úgy analizáltatjuk a visszatartott csoportot, mintha független, ismeretlen minták lennének, majd statisztikai értékelést követően a visszatartott csoport visszakerül a kalibráló mintaseregbe. A keresztvalidálás során az előbbieken ismertetett műveletsort az összes csoporttal megismétlődik. Amennyiben a csoport egytagú, a minták mindegyike egyszer kihagyásra kerül. Ezt hívjuk teljes keresztvalidációnak (full cross validation).



92.8% of original grouped cases correctly classified

80.4% of cross-validated grouped cases correctly classified

7. ábra: Diszkriminancia-analízis: Étkezési sütőolaj sütés közbeni változásának vizsgálata elektronikus orral (Dalmadi, 2006)

Az analízis további informatív eleme a *tévesztési mátrix* (confusion matrix). Különböző fizikai tartósító eljárások hatását elektronikus nyelv használatával tanulmányoztuk málnapüréken. Az **1. táblázatban** látható, hogy a teljes modellre vonatkoztatott újrabesorolási eredmények meglehetősen rosszak. Viszont jó elkülönüléseket figyelhetünk meg, ha az egyes csoportokat külön-külön értékeljük. Látható, hogy bár minden csoportból minden csoportba tévesztett a modell, mégis csak kevés nagy hidrosztatikus nyomással (HHP) kezelt minta került a kezeletlen mintába

és fordítva. Tehát elektronikus nyelvvel nagyobb különbség volt kimutatható HHP kezelés esetében, mint a hőkezelésnél.

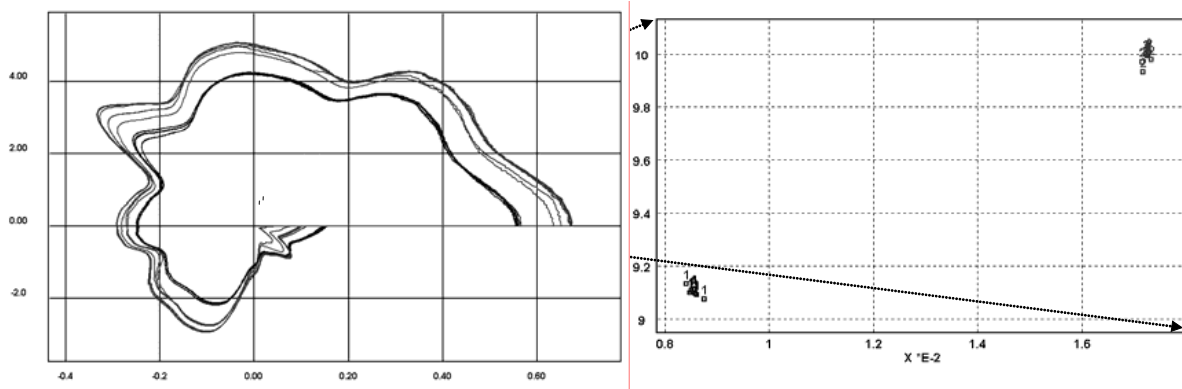
**1. táblázat: Diszkriminancia-analízis tévesztési mátrix:
Fizikai tartósító eljárások hatása málnapürén elektronikus nyelvvel vizsgálva (Dalmadi et al., 2006)**

| | | Predicted Group Membership | | |
|------------------|--------------------------|----------------------------|--------------------------|----------|
| | | HHP (600MPa, 5perc) | hőkezelt (80C, 5perc) | kontroll |
| Original | HHP (600MPa, 5perc) | 50,0 | 43,8 | 6,3 |
| | hőkezelt (80C, 5perc) | 18,8 | 62,5 | 18,8 |
| | kontroll | 6,3 | 18,8 | 75,0 |
| Cross-validation | HHP (600MPa, 5perc) | 25,0 | 68,8 | 6,3 |
| | hőkezelt (80C, 5perc) | 56,3 | 12,5 | 31,3 |
| | kontroll | 6,3 | 37,5 | 56,3 |

62,5% of original grouped cases correctly classified.

31,3% of cross-validated grouped cases correctly classified.

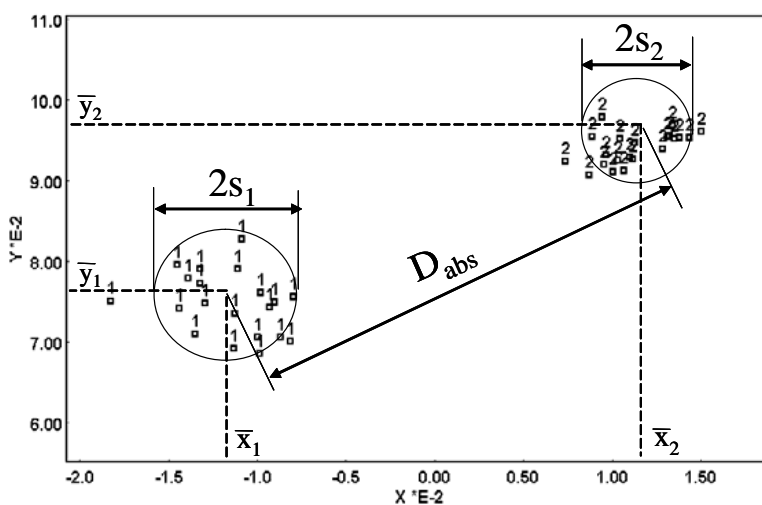
A **Polár Minősítő Rendszer** (*Polar Qualification System, PQS*) egy geometriai szemléleten alapuló adatredukciós módszer, mellyel drasztikusan csökkenthető a változók száma (Kaffka, 2001). Ehhez a derékszögű koordinátarendszerben ábrázolt spektrumot a polár koordinátarendszerben ábrázoljuk úgy, hogy a sugár a spektrumérték míg a szög a hullámhossz függvénye. Ezután meghatározzuk a polár spektrum középpontját (**8. ábra**).



8. ábra: Polár Minősítő Rendszer elve

Néha előfordulhat, hogy a spektrumot a polár koordinátarendszerben ábrázolva a különbségeket hordozó csúcsok 180°-ra kerülnek egymástól, gyengítve a minőségpontok elhelyezkedésére gyakorolt hatását. Ez a probléma az ún. „hullámhossztartomány-optimalizálással” oldható meg. Ekkor az a hullámhossztartomány kerül meghatározásra, amely a két minta közötti legjobb elkülönülést biztosítja a minták minőségpontjainak alapján (9. ábra). A legjobb elkülönítés meghatározására szolgáló célfüggvények:

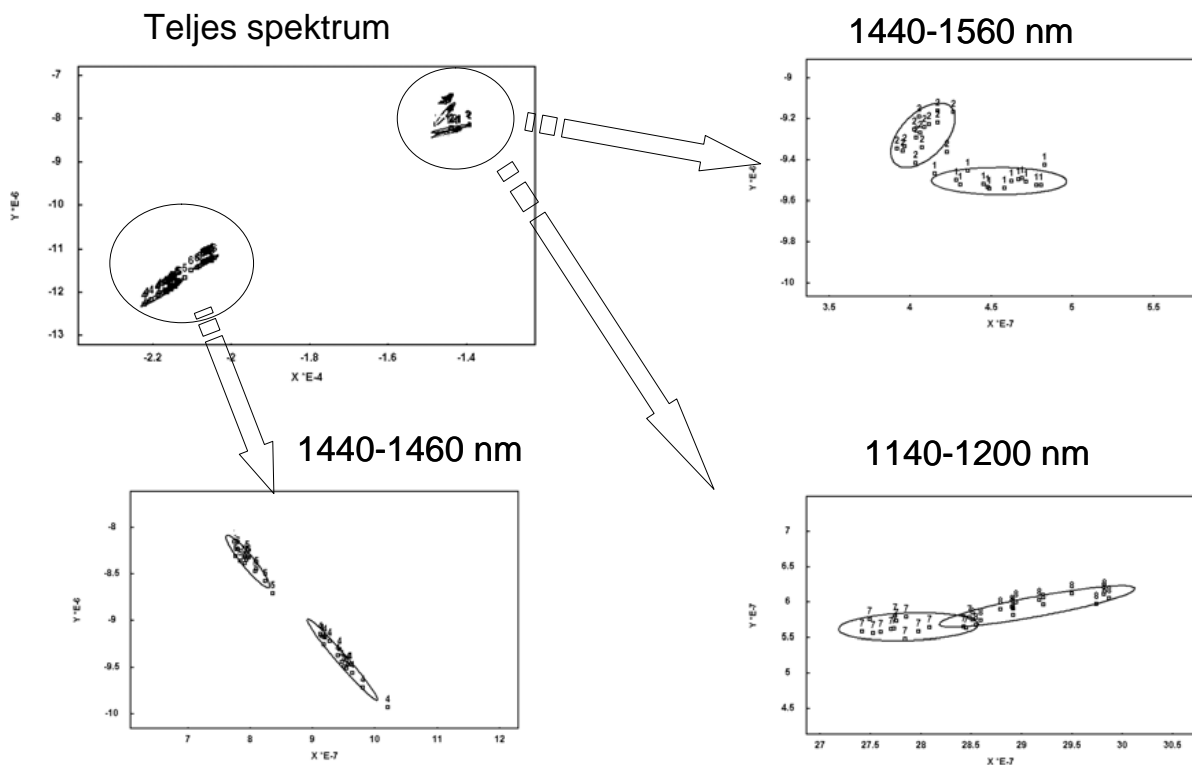
- az „abszolút távolság”: $D_{absz} = \sqrt{(\bar{x}_2 - \bar{x}_1)^2 + (\bar{y}_2 - \bar{y}_1)^2}$,
- a „normalizált távolság”: $D_{norm} = \frac{D_{absz}}{D_{absz} + s_1 + s_2}$,
- vagy az „érzékenység”: $S = \frac{D_{absz}}{s_1 + s_2}$.



9. ábra: Csoportok közötti távolság meghatározása PQS rendszerben (Kaffka, 2001)

Az optimalizációs algoritmus a spektrum elejétől egy keskeny kapuval elindulva, a spektrum végéig eltolva az optimalizálás valamennyi lépésében kiszámítja az érzékenység – vagy a másik két paraméter – értékét a kapu által lefedett spektrális pontok felhasználásával. Ezután az algoritmus visszatér a spektrum elejéhez és a kaput megszélesítve újra eltolja azt a spektrum végéig. Ez addig folytatódik, amíg a kapu az első és utolsó hullámhossz közötti távolságig nem szélesedik, azaz át nem öleli a teljes spektrumot.

A PQS módszert alkalmaztuk panírporok beszállítói tételeinek megkülönböztetésénél (10. ábra). A teljes NIR spektrumot felhasználva jól elkülöníthető volt a két különböző fajta panír egymástól. A hullámhossz optimalizációt elvégezve mindkét csoportnál található volt olyan tartomány, ahol a különböző tételek is elkülöníthetőek voltak egymástól.

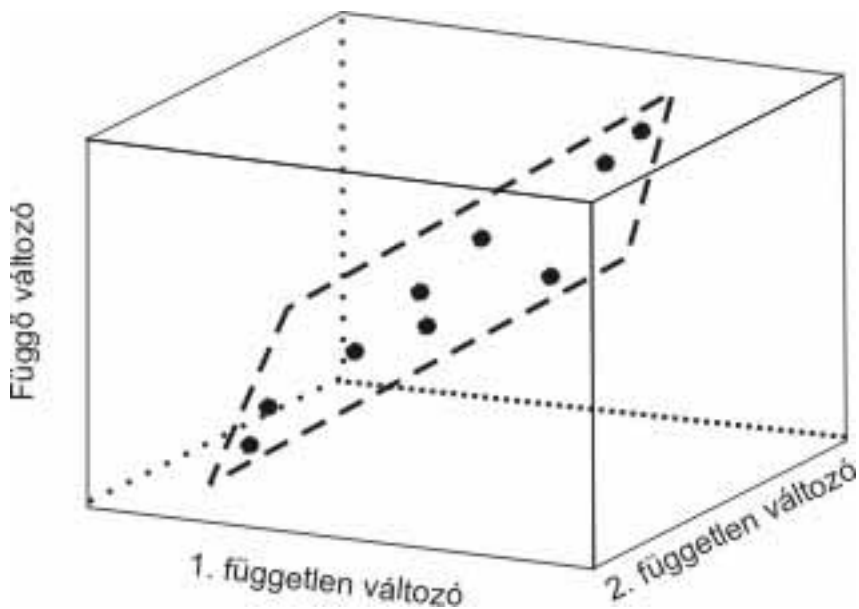


10. ábra PQS: Panírporok különböző tételeinek összehasonlítása NIR spektrumuk alapján (Balla, 2006b.)

Az előző módszerekkel ellentétben az **Spektrum (Minta) Felismerő Eszköz** nevű módszer (*Sample Recognition Tool, SRT*) egy nem lineáris kiértékelő módszer, mely alkalmas a spektrumok nem linearitásában rejlő információk kivonatolására. Továbbá míg a bemutatott klasszikus módszerek kiértékelésénél vizuálisan kell meghatározni, hogy az ismeretlen minta pontja mely csoportba sorolódott, addig az SRT módszer az ismeretlen mintát spektruma alapján egyértelműen a feltréningezett, betanított mintacsoportok valamelyikéhez rendeli.

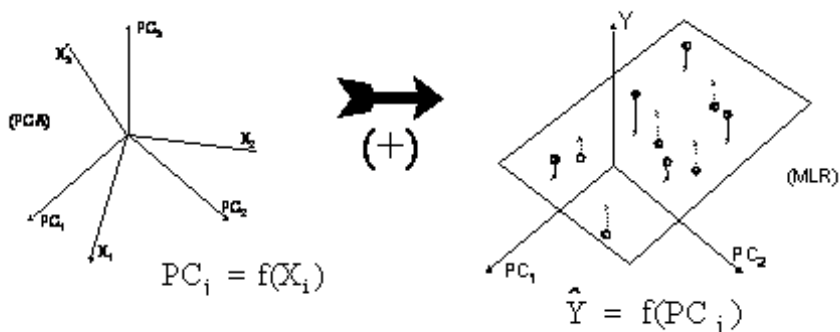
A tréning alapja, hogy az egyes osztályokhoz tartozó spektrumok minden hullámhosszhoz tartozó értékeinek sűrűségfüggvényét valószínűségi függvénynek tekintjük. Az ismeretlen minta spektrumértékeit a modellezett osztályok függvényei alapján minden hullámhosszon értékeli, és ezen valószínűségi értékek logaritmusainak összegét határozza meg, majd abba az osztályba sorolja, amelyhez az ismeretlen minta így kapott valószínűsége a legnagyobb. Eredménye hasonló a diszkriminancia-analízisnél látott tévesztési mátrixhoz.

A **11. ábrán** bemutatott mátrix olivaolaj napraforgóolajjal való hamisításának kimutatására irányuló kutatásunk eredménye. Látható, hogy a NIR vizsgálat SRT módszerrel alkalmas volt a keverési sor valamennyi elemének megbízható azonosítására.



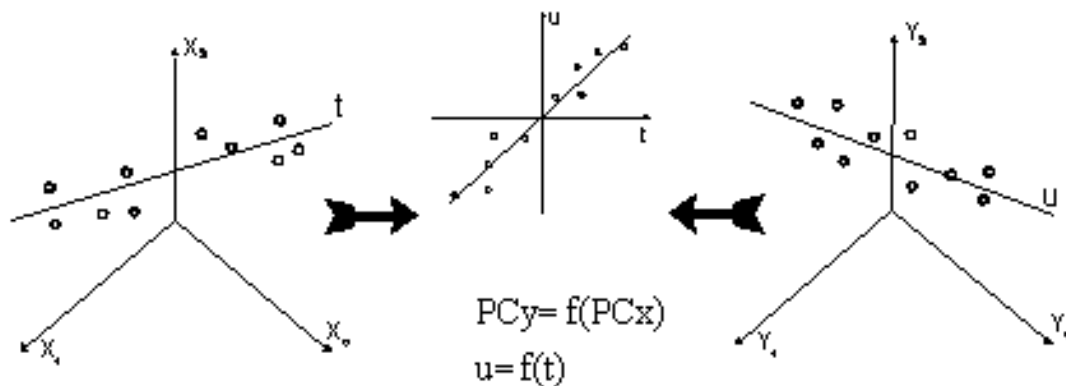
12. ábra Többváltozós lineáris regresszió

Ideális esetben a megfelelő változók kiválasztásával – vagyis elkerülve a keresztkorreláció jelenségét miközben a változók erősen korrelálnak a függő változóval – biztosíthatjuk, hogy egy-egy változó a referencia (y) érték különböző részeit magyarázza. Sajnos a gyakorlatban a változók egymástól ritkán függetlenek. Megoldást jelent, ha az eredeti változók helyett az azokból képzett, új változókkal végezzük el a regressziós feladatot. A **főkomponens-regresszió** (*Principal Component Regression, PCR*) algoritmus a eredeti változók főkomponens-elemzése során alkotott új változók és a függő változók között teremt kvantitatív kapcsolatot. Ezáltal olyan előrebecslést kapunk a függő változó értékére mely használja ugyan az összes eredeti változóban meglévő információt, de kiküszöböli a keresztkorrelációt. Így az előrebecslés (prediction) hibája kisebb lesz mintha az eredeti változókat használnánk, de csak annak az árán, hogy szisztematikus hiba kerül a becslésbe.



13. ábra Főkomponens regresszió elve

A PCR-t gyakran összehasonlítják a **részleges legkisebb négyzetek módszerével** (*Partial Least Squares Regression, PLS*) módszerével. Általában a PLS jobb eredményeket ad (kisebb hibát és jobb értelmezhetőséget is), de ez nem szükségszerű. A jobb eredmények oka, hogy a függő változóban (y) meglévő információt is fölhasználjuk a becslés során, egyidejűleg modellezi a független és függő változót. A modell a számított rejtett változók számának növelésével egyre nagyobb mértékben írja le az adathalmazok összefüggéseit. A PLS ezekre a változókra részleges kalibrációkat alkalmaz a variancia összegének modellezésére, amelyeket a művelet végén egy átfogó kalibrációs egyenletbe gyűjt (**14. ábra**). Az optimális tagszám meghatározása a PLS kalibráció része: túl kevés változó esetén a kalibráció kevés információt hordoz és nagy predikciós hibával dolgozik, míg túl sok változó alkalmazáskor a modell túlilleszti a kalibrációs adatokat, és az így elveszti a robusztusságát, stabilitását. Az optimális tagszámot rendszerint keresztvalidálással határozzuk meg.

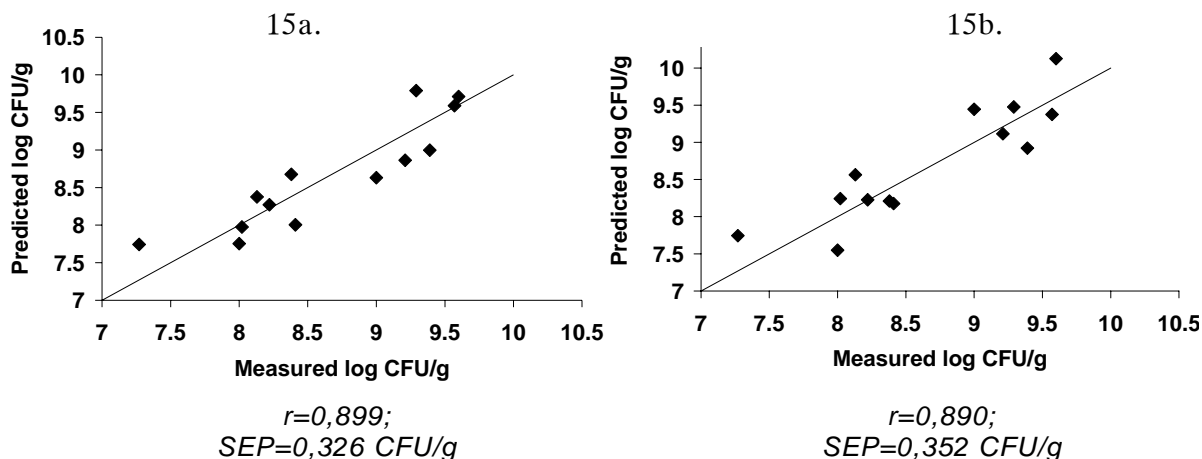


14. ábra: A PLS regresszió elve

PLS regresszióval vizsgáltuk, hogy az elektronikus orr jelválaszaiból becsülhető-e az aerob összcsíraszám sertéshússzeletek 4, 8 és 12 °C-os tárolása mellett. Azt tapasztaltuk, hogy a szenzorok jelválaszai csak bizonyos bakteriális szennyezettség felett mutattak jelentősebb korrelációt az aerob összcsíraszámmal. A 10^7 -nél nagyobb csíraszám-értékek figyelembevételével készült predikciós modell a **15a. ábrán** látható módon alakult.

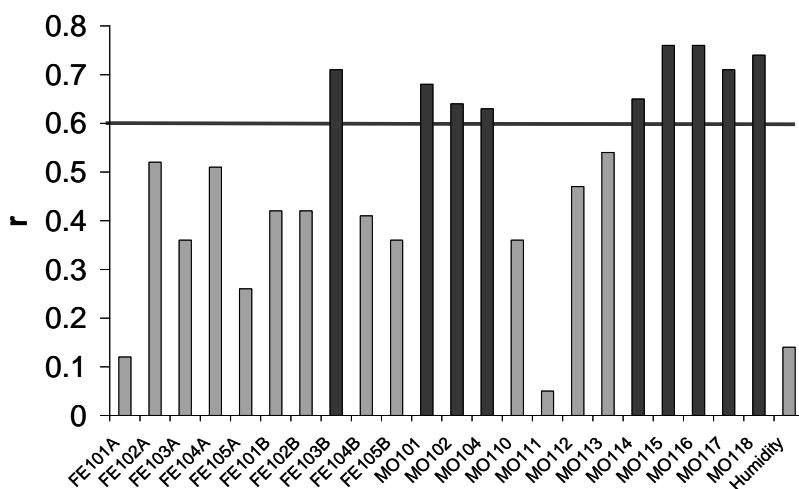
Szenzorselekció során kiválasztottuk azokat az érzékelőket, amelyek az aerob összcsíraszámmal szorosan korrelálnak ($R > 0.6$) (**16. ábra**), majd ezekre végeztük el a PLS kalibrációt. A szenzorselekciót követően a **15b. ábrán** látható eredményhez jutottunk. A 14 lényegi információt nem

tartalmazó érzékelő elhagyásával a korrelációs koefficiens és a predikciós hiba értéke alig változott.



15. ábra: PLS: sertéshús-szeletek aerob összcsíraszámának becslése elektronikus orral 4- 12°C- közötti tárolási kísérletek esetén

(15a: az összes szenzor adatainak felhasználásával alkotott modell; 15b: a kiválasztott 9 legérzékenyebb szenzor felhasználásával alkotott modell) (Horváth, 2007)



16. ábra PLS: elektronikus orr szenzor jelválaszainak vizsgálata sertéshús-szeletek aerob összcsíraszámának becsléséhez (Horváth, 2007)

Ez lehetőséget ígér kevesebb szenzorral működő, hordozható készülék kifejlesztésére.

Irodalom

- Adams, M.J. (2004): Chemometrics in Analytical Spectroscopy. 2nd Edition, The Royal Society of Chemistry, Athenaeum Press, Ltd., Gateshead, UK
- Balla, Cs. (2005): Sütőolajok és panírok vizsgálata. Jelentés innovációs kutatásról, BCE ÉTK Hűtő- és Állatitermék Technológia Tanszék
- Balla, Cs. (2006): Panírok vizsgálata. Jelentés innovációs kutatásról, BCE ÉTK Hűtő- és Állatitermék Technológia Tanszék

- Dalmadi, I. - Kántor, D.B. - Polyák-Fehér, K. - Fekete, A. - Farkas, J. (2006): Bogyógyümölcs-pürék érzékszervi tulajdonságainak vizsgálata műszeres módszerekkel. Műszaki Kémiai Napok '06, Veszprém, 2006. április 25-27. Konferencia kiadvány, 186-189.
- Dalmadi, I. - Seregély, Zs. - Balla, Cs. (2006): Monitoring the effect of crumb on sunflower oil during frying by NIR spectroscopy and chemosensor array. Book of Abstracts, 3rd Central European Congress on Food, 22-24 May 2006, Sofia, Bulgaria, 125.
- Horváth, M. K. - Seregély, Zs. - Dalmadi, I. - Andrásy, É. - Farkas, J. (2007): Estimation of bacteriological spoilage of pork cutlets by electronic nose. Acta Microbiologica et Immunologica Hungarica, **54** (2), 179-194.
- Kaffka, K. - Dalmadi, I. - Horváth, K. (2007): Classifying paprika seeds according to variety and harvest year using their NIR spectra The 13th International Conference on Near Infrared Spectroscopy (13th ICNIRS), Umeå-Vasa, Sweden & Finland 15-21 June 2007, NIR Abstracts: C-52_276
- Kaffka, K. - Seregély, Zs. (2001): PQS: egy új adatredukciós és minősítő rendszer. Magyar Kémiai Folyóirat, 107 (2.), 71-81.
- Ozaki Y. - McClure, W. F. - Christy, A.A. (eds.) (2006): Near-Infrared Spectroscopy in Food Science and Technology. John Wiley & Sons, Ltd., Chichester, England
- Pearce, T.C. - Schiffman, S.S. - Nagle, M.T. - Gardner, J.W. (eds.) (2003): Handbook of Machine Olfaction: Electronic Nose Technology, Wiley-VCH
- Somogyi, L. (2005): Italporok objektív műszeres minősítése. Jelentés innovációs kutatásról, BCE, ÉTK, Gabona- és Iparinövény Technológia Tanszék
- Seregély, Zs. (2006): A minőség nyomonkövetése az élelmiszerlánc különböző pontjaiban gyors fizikai módszerekkel, OTKA PosztDoktori Pályázat Zárójelentés (D-45953)

Néhány többváltozós kemometriai módszer alkalmazása műszeres analitikai vizsgálatok értékelésére

A korrelatív mérési technikák mára versenyképes társaivá váltak a hagyományos műszeres analitikai eljárásoknak. Kiemelkedő jelentőséggel bír az ún. NIR technika, mely a közeli infravörös tartományban mért spektrumértékekből összetételi jellemzők gyors, pontos, roncsolásmentes meghatározására képes. Az indirekt vizsgálati módszerek egy másik csoportját a kémiai érzékelősorok alkotják, amiket legtöbbször elektronikus orrként és nyelvként említ a szakirodalom. A korrelatív technikákra jellemző, hogy a vizsgált mintáról rövid idő alatt nagymennyiségű információt szolgáltat, melynek értékelése elképzelhetetlen matematikai-statisztikai módszerek nélkül. A többváltozós statisztikai módszerek rutinszerű alkalmazását a hardver és a szoftver területeken bekövetkezett

fejlődésnek kellett megelőznie, így ezek a módszerek csak mostanában válnak szélesebb körben ismertté. Jelen közleményünkben saját NIR és kémiai szenzorokkal kapott eredményeinken keresztül mutatunk be példákat néhány többváltozós statisztikai módszerekről. Bemutatásra kerülnek olyan nem-felügyelt (főkomponens-analízis, hierarchikus klaszter-analízis) és felügyelt (diszkriminancia-analízis, Polár Minősítő Rendszer és SRT módszer) kvalitatív matematikai-statisztikai módszerek, melyekkel az osztályozási feladatokat tudtuk megvalósítani. Továbbá kitérünk a többváltozós kvantitatív értékelés néhány esetére is, mint a többváltozós lineáris regresszió, a főkomponens-regresszió és a részleges legkisebb négyzetek módszere.

Application of some multivariate chemometric methods for the evaluation of instrumental analytical investigations

Correlative measurement techniques have become by now competitors of the traditional instrumental analytical procedures. The so-called NIR technique has gained outstanding importance. It is capable for the fast, accurate and non-destructive determination of the compositional characteristics from the spectral values measured in the near infrared region. An other group of indirect measurement methods include chemical sensor arrays, that are generally mentioned in the literature as electronic nose and electronic tongue. Correlative techniques provide in short time large amount of information about the sample, that has to be evaluated by mathematical-statistical methods. Development in the field of hardware and software had to precede routine application of relevant multivariate mathematical-statistical methods, so these methods have only nowadays become more widely known. In the present study we show some examples of such methods through our results obtained by NIR and chemical sensor arrays. Supervised (Principal Component Analysis, Hierarchical Cluster Analysis) and unsupervised (Canonical Discriminant Analysis, Polar Qualification System and SRT methods) qualitative mathematical statistical methods are presented here, by which classification tasks could be performed. Furthermore, some cases of multivariate quantitative evaluation, such as Multivariate Linear Regression, Principal Component Regression, and Partial Least Squares Regression are also described.