Adatfelvételi és adatfeldolgozási módszerek a PIM modellben

Csiszár Gábor

2003. november 14.

Tartalomjegyzék

1. Bevezetés

A 2. fejezet a PIM[1] modellről és azokról az adatfeldolgozási módszerekről szól, amelyekkel elő tudunk állítani a PIM számára megfelelő bemeneti adatokat a KSH[2] rendelkezésére álló különféle mérésekből. Először megvizsgáljuk, hogy milyennek kellene lennie a PIM modell számára rendelkezésre álló bemeneti adatoknak. Ezután mutatunk néhány példát arra, hogy milyen felmérések állnak rendelkezésre a KSH-ban. Megvilágítjuk, hogy milyen összefüggések használhatóak fel az adatok átalakítása folyamán. Ennek keretében ismertetjük a modell összefüggéseit. Beszélünk az értékcsökkenési függvényről, és ennek kapcsán a maximális élettartam szerepéről a modellben. Megvizsgáljuk, hogy a jövőben mely felmérések lennének a legmegfelelőbbek a PIM modell céljaira. Látni fogjuk azt is, mely területek szorulnak fejlesztésre a PIM modellben. Végül két példát mutatunk adatfeldolgozási módszerekre.

A 3. fejezet és a 4. fejezet a felhasznált algoritmusok matematikai hátterét tárgyalja. A 3. fejezetben egy egyszerű spline-függvény illesztésről lesz szó, amely segítségével az eszközcsoportnak, ha tudjuk az állagcsoportok megoszlását, meg tudjuk adni a kormegoszlását (évekre bontva) is. A 4. fejezetben egy olyan algoritmusról lesz szó, amely az eszközcsoport ún. állagmutatójából megadja az eszközcsoport állagcsoportjainak megoszlását. A 4. fejezetben tárgyalt algoritmus csak a PIM modell 'beindítására' használható, az ezzel a módszerrel készült adatokat később helyesbíteni kell - a 'Hibabecslés' részben ezt külön tárgyaljuk..

A függelékben bemutatjuk az algoritmusok egy *Mathematica* implementációját, valamint példákon keresztül szemléltetjük az algoritmusok működését.

[1]Perpentual Inventory Modell - Folyamatos leltárazás modellje. A PIM modell egy részletesebb ismertetése található [5] cikkében vagy [6]-ban

[2]Központi Statisztikai Hivatal

2. A PIM modell ismertetése

2.1. A PIM modellről általában

A PIM modellt a Kanadai Statisztikai Hivatal alkotta meg, hogy a segítségével nyomon követhesse a nemzeti tőkeállomány változását. A modell nemcsak az egyes ágazatokban jelenlevő tőke nagyságát, hanem a struktúráját is tárolja (mennyi elöregedett, illetve új eszköz van, mekkora lesz a várható selejtezések aránya, stb.). A modell segítségével könnyebben számíthatóak a nemzeti számlarendszer (SNA[0]) mutatói.

A Központi Statisztikai Hivatal az EU-csatlakozás Nemzeti Programjának teljesítése érdekében vezeti be a PIM modellt, az Állóeszköz-projekt keretein belül. A modell bevezetése során számos, csak a magyar nemzetgazdaságra jellemző problémát kellett figyelembe venni (a 90-es években lezajlott rendszerváltás és privatizáció, a piacgazdaság kialakulás; nem megfelelő módszertannal készült, hiányos és elavult felmérések). Ezeket figyelembe véve, a modellben számos módosítást (elhagytunk néhány kevésbé fontos részletet) és fejlesztést eszközöltünk.

Meg kell jegyezni, hogy a PIM modell működtetéséhez nem lesz elegendő csupán a 'program futtatása', hanem szükség lesz az egyes szektorok (háztartási, vállalati, pénzügy, nonprofit,stb.), és az ezen belüli ágazatok (gépipar, erdészet, távközlés, stb.) viszonyait jól ismerő szakértők közreműködésére. Az adatokat időről időre frissítjük (és a modell helyességét is teszteljük) a többévente elkészülő, átfogó felmérések segítségével (pl.: Általános Mezőgazdasági összeírás, BM nyilvántartás, stb.).

A modell a gazdaság egyes régióiban bekövetkező tőkeállomány változás nyomon követésére használható, hiteles képet nyújt a tőkeállomány struktúrájáról. A modellbe épített függvények segítségével előrejelzések is készíthetőek: mely ágazatokban, illetve földrajzi régióban indul meg a fejlődés, hol van szükség tőkebeáramlásra, mely iparágak épülnek le, stb. A modell segítségével összehasonlíthatóvá válik nemcsak az egyes nemzetgazdaságok nagysága, hanem struktúrája is. Elemezhetővé válik a gazdaságpolitikai döntések rövi- és hosszútávú hatása.

[0]SNA — System of National Accounts, vagyis a Nemzeti Számlák Rendszere

2.2. Milyen bemenő adatokat kell előállítanunk?

Az állóeszközöket szétválasztjuk aszerint, hogy

- jármű, gép, felszerelés vagy épített ingatlan;
- milyen szektorba tartozik (államháztartási, vállalati, háztartási, nonprofit, stb);
- milyen ágazatba tartozik (viziközművek, erdészeti létesítmények, stb);
- milyen megyébe tartozik(Pest megye);
- mennyi a maximális élettartama az eszköznek;
- mi az aktiválási éve az eszköznek;

2.1 Példa Tegyük fel, a vállalatok szektorában részletes felmérést végeztek a járművek állományában. Tehát a járműveket a következő ismérvek aléapján tagoljuk eszközcsoportokra:

szektora	vállalat
ágazat	erdészet
telephely	Pest megye
maximális élettartam	15 év
aktiválási év	1994

Más szavakkal, ideális adatbázisunkból lekérdezhető, hogy a vállalati szektorban az adott aktiválási évű, adott maximális élettartamú, adott ágazatú, adott székhelyű vállalatoknál pénzben kifejezve mennyi a járművek mostani bruttó értéke (újrabeszerzési érték, mennyit kellene fizetnünk, ha mindezt újonnan meg kellene vennünk), mostani nettó értéke (mennyit kapnánk érte, ha el akarnánk adni), valamint a történelmi értéke (mennyiért vettük mindezt az aktiválás évében)

Ilyen részletességű felmérések sajnos nem állnak rendelkezésre. A beruházási adatsorokat (amelyekben pontosan a kivánt formában vannak megadva az adatok) az állami gazdaság szétesése, a privatizáció és a piacgazdaság kialakulása közben bekövetkezett követhetetlen tőkemozgások miatt. A tíz évnél régebbi felmérések nem tükrözik a gazdaság jelenlegi szerkezetét, hiányosak és elavultak.

Ezért a KSH arra a meggyőződésre jutott, hogy rekonstruáljuk a beruházási adatsorokat (illetve megkonstruáljuk a PIM bemeneti adatait) a rendelkezésre álló, most elvégzett (elvégzendő) felmérésekből. Mivel ezek a felmérések nem az állóeszköz projekt keretében készültek, ezért ezek igen sokfélék. A feladatunk, hogy ezt a sokféle felmérést egységes, a PIM számára elfogadható adattá alakítsuk.

2.3. A rendelkezésre álló adatok

Lássunk tehát néhány példát a rendelkezésre álló felmérések sokféleségére:

- A vállalati szektorban 2000-ben végrehajtott mintavételes eszközfelvétel alapján ágazati tagolásban ismert a járművek teljes állományának korcsoportok szerinti összetétele, valamint a 2000 évi árakra számított újrabeszerzési (bruttó) értéke.
- A járművek esetén kiegészítő információkat szolgáltat a közúti járművek kormegoszlására a Belügyminisztérium teljes körű nyilvántartása, amely (az üzembe helyezett típusok szerint tagoltan) tartalmazza az egyes évjáratokhoz tartozó járművek darabszámokat. Rendelkezésre áll továbbá, hogy a személygépkocsit, illetve haszonjárművet "természetes személy" vagy "jogi személy" tartja üzemben. Lehetőség van az újrabeszerzési érték, valamint a jelenlegi (használt állapotnak megfelelő, vagyis nettó) érték meghatározására.

Az első példákban éves (tehát ismert, hogy adott évben mennyi az eszközök bruttó értéke) valamint korcsoportos (meg van adva, hogy pl.: mekkora a bruttó értéke a 0-5, 5-10, 10-15, 15-20, valamint 20-25 éves járműveknek) tagolással van dolgunk. Az éves tagoláson nem kell változtatnunk, a korcsoportos tagolást viszont tovább kell finomítanunk[3], hogy éves tagolást kapjunk.

Nehezebb dolgunk van viszont, ha az eszközök állagcsoportokra vannak osztva. Mit is jelent az, hogy állag? Az állag egy százalékban kifejezett érték,

[3] a függvény-illesztési algoritmust a 3. fejezet ismerteti

ami megadja, hogy egy eszköz hanyadrészét éri a bruttó értékének (amennyit új állapotban érne). Tehát az állagból és a bruttó értékből meg tudjuk állapítani, hogy mennyi az eszköz nettó értéke (a piaci értéke, amennyiért használt állapotban el tudjuk adni).

- Az Országos Vízügyi Főigazgatóság kezelésében lévő, árvízvédelmi létesítmények értékéről felmérés készült országos szinten, öt állagcsoportra bontva

Nyilván minél öregebb egy eszköz, annál rosszabb az állaga, s ez megfordítva is igaz: minél rosszabb az állaga, annál inkább elöregedett. Az állag és az életkor között összefüggés áll fenn. Ha ismerjük egy eszköz maximális élettartamát és életkorát, akkor tudunk készíteni becslést az állagára. Ugyanez igaz fordítva is: hogyha ismerjük egy eszköz maximális élettartamát és állagát, tudunk becslést készíteni az életkorára. Így az állagcsoportokból korcsoportokat tudunk készíteni, majd a korcsoportokat tovább tudjuk finomítani életkor szerinti tagolássá[4].

Vannak olyan esetek is amikor a homogén eszközcsoportnak csak a bruttó és a nettó értéke áll rendelkezésre. Ekkor is elkészítjük a PIM bemenő adatait, bár az így konstruált adatokat csak kiindulási adatokként használjuk fel. Ezeket aztán majd a később elkészítendő felmérések adataival helyettesítjük és vetjük össze.

- A járművek állományában rendelkezésre áll a bruttó érték és a nettó érték szektoros, megyei és ágazati tagolásban, így rendelkezésre áll az állagmutató (a nettó és a bruttó érték hányadosa).

A metódus ilyen adatok esetében: az átlagos állagmutatóból egy ún. "minimax" eljárással[3] állagmutató csoportokat képzünk, ezeket átalakítjuk korcsoportokká, majd erre fügvényt illesztünk (itt három átalakítás is történt, tehát a végeredményként kapott adatok minősége gyenge).

[4]Az összefüggés nem ilyen nyilvánvaló: a nagy maximális élettartamú eszközöket soha nem selejtezik le, hanem folyamatosan felújítják, így nem tükrözi a valóságos viszonyokat, ha aktiválási évek szerint adjuk meg az adatokat. Ilyen eszközök pl.: az utak, gátak, épületek

[3]Ezt a 4. fejezetben tárgyaljuk

2.4. Az értékcsökkenési függvény és a maximális élettartam kapcsolata

Adott évhez tartozó történelmi értékből kiszámíthatjuk a bruttó értéket, ha ismerjük az árindexet, azt hogy egy eszköz újrabeszerzési értéke mennyit változott az infláció, technológiai avulás, stb. miatt

Ha K a történelmi érték, G a bruttó érték, PI az árindex, now az aktuális évet, $past$ az üzembe helyezés évét jelöli, t pedig az aktiválástól eltelt évek száma, akkor a következő összefüggés áll fenn:

$$G_{now,t} = \frac{K_{past,t}}{PI_{past,t}} \tag{1}$$

Adott évhez tartozó történelmi értékből kiszámíthatjuk a nettó értéket, ha ismerjük

- Az árindexet
- Az évek során végbement értékcsökkenést

Legyen K,PI,now,$past$,t ugyanaz, mint (1)-ben, valamint N jelentse a nettó értéket,L a maximális élettartamot, $D(t, L)$ pedig azt, hogy az aktiválástól eltelt t-edik évben mennyi az eszköz nettó és bruttó értékének az aránya (értékcsökkenési függvény). Ekkor a következő összefüggés áll fenn:

$$N_{now,t} = \frac{K_{past,t}}{PI_{past,t}} D(t, L) \tag{2}$$

A kanadai PIM modelltől abban különbözik a KSH által használt modell, hogy a selejtezési függvényt és az avulási függvényt az adatkonverzió (állagból életkorba, illetve fordítva) miatt kivettük a modellből.

2.2 Megjegyzés Bár az algoritmusok *Mathematica* implementációjában a lehető legegyszerűbb értékcsökkenési függvény szerepel, az algoritmus könnyen átírható úgy, hogy a felhasználó adja meg az adott eszközhöz tartozó értékcsökkenési függvényt. A következő két képlet bármelyike megadhatná az értékcsökkenési függvényt:

$$D(t, L) = (1 - \frac{t}{L})^a \qquad D(t, L) = \frac{L - t}{L - at}$$

ahol $a > 1$ az első képletben, t és L jelentése a szokásos

2.3 Megjegyzés Természetesen az avulás nemcsak idő dimenziójú lehet, hanem pl.: megtett kilóméter, üzemi ciklusok száma, árvizek száma (és erőssége), stb.

2.4 Megjegyzés A modellben szerepelt az ún. avulási, illetve értékcsökkenési függvény. Azonban az állag-kor transzformáció, illetve ennek az inverze nem tudtuk megvalósítani sem diszkrét sem folytonos esetben, így ennek alkalmazásától eltekintettünk (egyenlőre) hogy egy lezárt, működőképes modell keretében tárgyaljuk az algoritmusokat (máskülönben csomó kérdést nyitva kellett volna hagynunk).

Látható, hogy modellünkben kiemelt szerepet kap a maximális élettartam. Ezért külön foglalkozunk a maximális élettartammal kapcsolatos problémákkal, valamint a maximális élettartam szempontjából homogén eszközcsoportok képzési módjával.

2.5. Esettanulmány a maximális élettartam fontosságának szemléltetésére

Példánkban azt akarjuk megmutatni, hogy olyan eszközöket (járműveket, gépeket, épületeket) nem szabad egy eszközcsoportba összevonni, amelyeknek jelentősen különbözik a maximális élettartama. Tegyük fel, 3 eszközünk van (A,B,C), amelyeknek ismert az állagmegoszlása (a szokásos 5 állagcsoportban).

Az eszközök megoszlása öt állagmutató-csoport szerint:

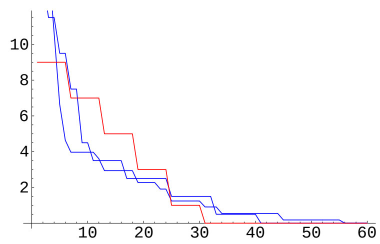
Eszközök	A szakértők becslése a maximális élettartamra				
jele	0 – 20%	20 – 40%	40 – 60%	60 – 80%	80 – 100%
"A"	20	25	15	22	18
"B"	5	15	30	30	20
"C"	10	15	20	25	30

Az esettanulmány céljaira a fenti 3 változathoz különböző maximális élettartamokat rendeltünk:

Eszközök	A szakértők becslése a maximális élettartamra		
jele	"1" változat	"2" változat	"3" változat
"A"	5 év	10 év	30 év
"B"	30 év	40 év	30 év
"C"	55 év	40 év	30 év

Esetünkben az összes megfigyelt eszköz átlagos (összesített) maximális élettartama mind a három súlyozási változatban 30 év. A grafikonon látható a selejtezések várható alakulása.

A vízszintes tengelyen a modell indításától eltelt évek vannak, míg a függőleges tengelyen a selejtezések mértékét látjuk a kezdeti eszközmennyiséghez képest, százalékban kifejezve.



Ha figyelembe vesszük a selejtezéssel együtt járó eszközpótlást, akkor adódik, hogy akkor lesz a legnagyobb mértékű'a selejtezés, amikor az eszközcsoport a maximális élettartam szempontjából legkevésbé homogén ('1' változat).

A legjobb volna, ha az eszközcsoportokat ugyanúgy tagolnánk maximális élettartam szerint, mint terület, szektor, vagy ágazat szerint[5]. A gyakorlatban általában nem lehet az eszközcsoportokat maximális élettartam szerint tovább bontani, vagy ha mégis, akkor a felmérések költsége a többszörösére növekszik A következő megoldáshoz folyamodunk tehát: megengedjük, hogy az eszközcsoport maximális élettartam szempontjából ne legyen teljesen homogén, ám azt nem hagyjuk, ám azt nem hagyjuk, hogy az eszközcsoporton

[5]A maximális élettartamot években (nem negyedében vagy hónapban) mérjük. Ez a legkisebb egység

belül az egyik eszköz maximális élettartama a másiknk többszöröse vagy törtrésze legyen.Például így:

	1 éves	2 éves	3éves	4-6 éves	8-14 éves	15-26 éves	24-47 éves	...
érték								

A csoportokban a maximális élettartamok alsó korlátjának és felső korlátjának aránya kb. $1,33^2 = 1,7689$[6]. Az egy csoportba eső eszközöket összegezzük (a különböző csoportba esőket nem), és kiszámítjuk az átlagos (összesített) maximális élettartamot.

2.6. Adatfelvételi módszerek

Láttuk, hogy az eredeti PIM modellben az adatok életkor szerinti tagolásúak. Sajnos az utólag elkészített felméréseknél nehéz ilyen pontosságban bekérni az adatokat (túlságosan költséges volna, vagy egyszerűen nem is lehetséges). Mégis, a jövőben milyen felméréseket végezzünk?

Természetesen a legjobb az éves tagolás, állagnál pedig a százalékos megoszlás (centilisek). Jól használhatóak azok a felmérések is, ahol az adatokat korcsoportokra (5-10 db) vagy állagcsoportokra bontjuk (szintén 5-10 db). Kis hibával ugyanis át tudjuk konvertálni az állagcsoportokat korcsoportokká, azokat pedig szintén kis hibával életkor szerinti tagolásúvá tudjuk finomítani.

A legrosszabb helyzetben akkor vagyunk, ha csupán egyetlen állagmutató (a bruttó érték, G és a nettó érték, N és a kettő hányadosa, az N/G=ALFA) áll rendelkezésre (márpedig a KSH-ban sajnos sok az ilyen fajta adatgyüjtés). Ilyenkor ezt a hányadost átalakítjuk öt állagmutató csoporttá, az állagmutató csoportokat korcsoportokká, majd a korcsoportokat életkor szerinti tagolássá finomítjuk. Mivel az átalakítások során a hiba igen nagy lesz (eleve kis információtartalmú adatokkal dolgozunk), az ilyen felméréseket csak 'gyorsmentésként' szabad igénybe vennünk.

Érdemes megjegyezni, hogy sokszor célszerűbb állagcsoport szerinti tagolásban tárolni az adatokat. Igaz, hogy a PIM modellben az adatokat életkor szerint csoportosítjuk. Azonban sok állóeszköz esetében (a nagy maximális élettartamúaknál, mint az épületek, gépek, utak) nem az életkor a döntő,

[6]Ezt az értéket érzékenységvizsgálattal majd ellenőrizni kell. Meg kell vizsgálni, hogy mekkora hatása van az intervallum hosszának a várható selejtezések alakulására

hanem az állag. Hiszen pl. az épületeket nem selejtezzük le maximális élettartamuk betöltésekor, hanem felújítjuk. Nyilván a felújítás nem az életkort, hanem az állagot befolyásolja. Amit sokszor újítunk fel, azt nem életkorban, hanem állagmutatóban érdemes tárolni.

2.7. Az algoritmus alkalmazásai

Lássunk néhány példát a 3 és 4. fejezetben ismertetett algoritmusok alkalmazására:

2.7.1. A járművek állományának feldolgozása

A rendelkezésre álló adatok:

A KSH szektorokban végzett felmérései, megyénkénti és ágazatonkénti tagolásban. A KSH adatállományai alapján ismert, hogy mennyi a járművek bruttó értéke (amennyiért újonnan meg kellene vennünk a járműveket, G), valamint nettó értéke (amennyiért most el tudjuk adni a járműveket, N) szektoronként, megyénként és ágazatonként felbontva. A bruttó illetve nettó érték alapján adott a $\frac{N}{G} = ALFA$ állagmutató is.

2.5 Definíció A továbbiakban $G_{i,j,k}$ fogja jelölni a bruttó értékét , $N_{i,j,k}$ pedig a nettó értékét a járműveknek az i-ik szektorban, j-ik megyében, k-ik ágazatban. A kettőnek a hányadosa $\frac{N_{i,j,k}}{G_{i,j,k}} = A_{i,j,k}$ pedig megadja az átlagos állagmutatót.

A BM adatállományára alapozott becslések. A BM adatai alapján (a forgalomba került összes jogi személyek által birtokolt jármű adata lekérdezhető: jármű neme, gyártmány, típus, gyártási év, valamint külső szakértők által meghatározott bruttó érték, G illetve nettó érték, N) összeállítható a járművek országos kormegoszlása, illetve korcsoportonkénti megoszlása. Az összes járművet szétoszthatjuk típusok szerint, így az egyes típusok országos kormegoszlása is becsülhető. Sajnos az adatokat ez alapján az adatbázis alapján nem tagolhatjuk szektoronként, megyénként illetve ágazatonként. A következő típusokat különítettük el:

tehergépjárművek;	vontatók;
lassú járművek;	pótkocsik;
utánfutók;	autóbuszok;
személygépkocsik	

2.6 Definíció A továbbiakban $KORMEGOSZLAS_{l,t}$ jelöli t életkorú járművek arányát az összes járműhöz képest az l-edik típusban. Tehát $KORMEGOSZLAS_l$ egy életkor szerinti vektor (kormegoszlás). Ez a BM adataiból származik (nem ugyanaz, mint $KORMEGOSZLAS_{i,j,k,l}$).

Szakértői becslések. Arra azonban, hogy adott megyében, szektorban illetve ágazatban milyen a járművek típus szerinti megoszlása, készíthetőek szakértői becslések.

2.7 Definíció A továbbiakban $w_{i,j,k,l}$ jelentse az i-ik szektorban, j-ik megyében és k-ik ágazatban az l-ik típus arányát. A típusokat már kellően homogén eszközcsoportnak tartjuk, így a típusokhoz maximális élettartamokat rendelünk.

2.8 Definíció Jelöljük $A_{i,j,k}$-val, illetve $A_{i,j,k,l}$-el az i-edik szektorban, j-edik megyében, k-adik ágazatban (és l-edik típusban) az állagmutatót

2.9 Definíció $ALLAGCSOPORT_{i,j,k}$ illetve ALLAGCSOPORT$_{i,j,k,l}$ jelölje az i-edik szektorban, j-edik megyében, k-adik ágazatban (és l-edik típusban) a KSH adatai alapján, illetve a szakértői becslések alapján konstruált állagcsoportot. Definiáljuk ugyanígy $KORMEGOSZLAS_{i,j,k}$-t és $KORMEGOSZLAS_{i,j,k,l}$-t az éves bontásra.

Az algoritmus a KSH adataira támaszkodik a leginkább, a szakértői becsléseket és a BM adatait csak korrekciós céllal használja. Az a kívánatos, hogy a járműveknek minél kevesebb típusát különböztessük meg. Az algoritmus főbb lépései a következők:

Először a szektoronként, megyénként, ágazatonként adott átlagos állagmutatóhoz 5 db állagmutató csoportot konstruálunk. Ugyanaz az állagmutató csoport megoszlása lesz az i-ik szektorban, j-ik megyében, k-ik ágazatban és l-ik típusban, mint az i-ik szektorban, j-ik megyében, k-ik ágazatban. Utána (immár a típus szerint is tagolt) állagmutató csoportokat korcsoportokká transzformáljuk, majd ezt tovább finomítjuk életkor szerinti tagolássá. Ekkor tudjuk, hogy

$$G_{i,j,k,l} = G_{i,j,k} w_{i,j,k,l} \qquad N_{i,j,k,l} = N_{i,j,k} w_{i,j,k,l}$$

$$\Rightarrow \begin{cases} A_{i,j,k} = A_{i,j,k,l} \\ ALLAGMUTATOCSOPORT_{i,j,k} = ALLAGMUTATOCSOPORT_{i,j,k,l} \\ KORMEGOSZLAS_{i,j,k} = KORMEGOSZLAS_{i,j,k,l} \end{cases} \tag{3}$$

A második lépésben következik a szakértői becslések és a BM adatai alapján a korrigálás. Azt szeretnénk, ha az állagcsoportok kielégítenének bizonyos egyenleteket. Azt szeretnénk, ha a típusonkénti összesített, átlagos kormegoszlások megegyezzenek azokkal a kormegoszlásokkal, amiket a BM adatbázisából kaptunk.

2.10 Definíció Jelöljük $VKORMEGOSZLAS_{i,j,k,l}$-el a korrigált adatokat. Ekkor:

$$\frac{\sum_{i,j,k} VKORMEGOSZLAS_{i,j,k,l} G_{i,j,k,l}}{G} = KORMEGOSZLAS_l$$

Ahhoz, hogy a két kormegoszlás a lehető legközelebb legyen egymáshoz, a következő célfüggvényt (távolságfüggvényt) kell minimalizálnunk:

$$\frac{\sum_{i,j,k} (KORMEGOSZLAS_{i,j,k,l} - VKORMEGOSZLAS_{i,j,k,l})^2 G_{i,j,k,l}}{100 G_{i,j,k,l}} \tag{4}$$

Ezzel készen is vagyunk.

2.11 Példa A kormegoszlás számítása szektoronként, megyénként, ágazatonként a következőképpen történik:

$$VKORMEGOSZLAS_{0,j,0} = \sum_{i,k} \big(\sum_{l} VKORMEGOSZLAS_{i,j,k,l}\big) G_{i,j,k}$$

2.7.2. A viziközműállomány felmérése

Rendelkezésre állnak önkormányzati adatgyűjtések a viziközművekről. Önkormányzatonként meg van adva a bruttó érték (újrabeszerzési érték), nettó érték, valamint a nettó és a bruttó érték hányadosa, az állagmutató. Értelemszerűen nincsen szektoros, vagy ágazati tagolás (az összes viziközmű az állam, illetve az önkormányzatok tulajdonában van, és a viziközművek maga egy ágazatot jelent). A regionális tagolás azonban mélyebb, mint a megyei tagolás, hiszen 1564 önkormányzat van (csupán azokat számolva, amelynek van polgármestere). Ezenkívül a víziközműveket 6 csoportra osztjuk (a maximális élettartamot L-el jelöljük):

Vízmű	$L = 50$	Víztároló	$L = 50$
Víztisztító	$L = 35$	Szennyvízgyüjtő	$L = 60$
Vízelosztó	$L = 60$	Szennyvíztisztító	$L = 35$

A lehetőségekhez képest megbízható adatok nyerhetőek úgy, hogy az állagmutatókból (önkormányzatonként csoportonként) állagcsoportokat képezünk. Ezeket összegezzük megyénként (kb. 70-80 önkormányzatról van szó megyénként). Ezután a megyei állagcsoportokat korcsoportokká alakítjuk, majd a korcsoportokat évekre bontjuk.

3. Spline függvény illesztés

3.1. Alapvető definíciók

3.1 Definíció *Eszköz*ről vagy *eszközfajtá*ról beszélünk, ha az eszközfogalom elég homogén, a bele tartozó eszközök maximális élettartama megegyezik, de semmilyen más (terület, szektorbeli hovatartozás vagy ágazat) tagolás nincs (pl.: járművek). Ellenkező esetben (ha van valamilyen tagolás) már *eszközcsoport*ról beszélünk.

3.2 Definíció A *használhatóság* vagy *állag* százalékos mérőszám. Egy eszköz használhatósága (állaga) azt méri, hogy az eszköz hanyadrészét éri az eredeti értékének.

3.3 Példa Eszközfajtának nevezzük pl. a személyszállító járműveket. De a személyszállító járművek Tatabányán a vállalati szektorban már eszközcsoportnak számít.Ha egy jármű 56%-osan használható, akkor eredeti értékének 56%-át éri.

Képezzünk *állagcsoportok*at, és adjuk meg, hogy az adott eszközcsoport mekkora hányada tartozik az adott állagcsoportokba (*állagcsoport megoszlás* vagy *állagmegoszlás*).

3.4 Példa Az alábbi példában megfigyelhető, hogy az állagmegoszlás-tagokat összegezve, százat kapunk.

	0–20%	20–40%	40–60%	60–80%	80–100%
Víztisztító művek	17	19	21	21	22

3.5 Definíció A *maximális élettartam* jelentése értelemszerű. Ha egy eszköz életkora eléri maximális élettartamát, értéke nullára csökken, leselejtezzük. Az eszközfajtába tartozó eszközök maximális élettartama (lényegtelen eltérésektől eltekintve) megegyezik.

3.6 Megjegyzés Ennél tovább is megyünk: feltételezzük, hogy az állag (A) és az életkor (E) között lineáris összefüggés áll fenn. Tehát:

$$A = 1 - \frac{E}{L}$$

3.2. A feladat megfogalmazása

Adott egy eszközcsoport k számú állagcsoportra bontva ($3 \leq k \leq 10$). Azt szeretnénk, hogy az eszközcsoport adatai évenként álljanak rendelkezésre. A megoldás lépései a következőek (algoritmusvázlat).

(1) Az állagcsoportokból korcsoportokat csinálunk (ez mindössze egy lineáris transzformációt jelent).

(2) Illesztünk egy megfelelő fokú polinomot (k csoport esetén $(k-2)$ fokút). Ezzel csupán egy algoritmus kiindulási adatait határozzuk meg, valamint ehhez a függvényhez próbálunk minél közelebb lévő spline függvényt illeszteni.

(3) Egy olyan másodfokú spline függvénnyel közelítjük meg a polinomot, amely nemcsak folytonos és differenciálható, hanem az is teljesül rá, hogy területtartó, vagyis a korcsoport által meghatározott intervallum feletti terület megegyezik a korcsoportba eső eszközök százalékos mennyiségével.

(4) Végül a spline függvényből integrálással már könnyen megkapjuk évenkénti bontásban az adatokat.

3.3. A megoldás

Első lépésként állagcsoportokból korcsoportokat kell kapnunk. Nyilván egy korcsoportban ugyanannyi eszköz lesz, mint a neki megfelelő állagcsoportban. Tehát csak azt kell meghatároznunk, hogy mik a korcsoportokhoz tartozó intervallumok. Kihasználva, hogy az állag és az életkor között lineáris összefüggés van:

Állagcsoport	$[100, 100K_{1,2}]$	$\ldots$	$[100K_{i,1}, 100K_{i,2}]$	$\ldots$	$[100K_{k,1}, 0]$
Korcsoport	$[0, L(1-K_{1,2})]$	$\ldots$	$[L(1-K_{i,1}), L(1-K_{i,2})]$	$\ldots$	$[L(1-K_{k,1}), L]$

Második lépésként függvényt kell illesztenünk az adatokra. Ehhez néhány dolgot tisztáznunk kell.

3.7 Definíció Az i-edik korcsoportban lévő eszköz (százalékos) mennyiségét jelölje E_i, a korcsoporthoz tartozó intervallum hosszát jelölje M_i.

3.8 Definíció Legyen f egy $k-2$-ed fokú polinom:

$$f(x) = c_0 + c_1 x + \cdots + c_{k-2} x^{k-2}$$

Tehát f maximum 8-adfokú.

A k állagcsoporthoz legközelebb eső függvényt az úgynevezett *legkisebb négyzetek módszerével* határozzuk meg. A módszer valójában egy szélsőértékszámítási feladat megoldását jelenti:

$$G(f) = \sum_{i=1}^{k-2} (M_i f(\frac{L(1-K_{i,1}) + L(1-K_{i,2})}{2}) - E_i)^2$$

Figyeljük meg jó alaposan a képletet. Nyilván $\frac{L(1-K_{i,1})+L(1-K_{i,2})}{2}$ a korcsoporthoz tartozó intervallum közepe. M_i azért kell, hogy az intervallum hosszának megfelelő súllyal vegyük figyelembe a függvényértéket.

A G funkcionál minimumát keressük, úgy, hogy feltesszük, hogy G változója meghatározott alakú függvény: k-adfokú polinom, bizonyos szabad paraméterekkel:

$$\begin{aligned} P_i &= \frac{L(1-K_{i,1}) + L(1-K_{i,2})}{2} \\ f(x) &= c_0 + c_1 x + \cdots + c_{k-2} x^{k-2} \\ G(c_0, c_1, \ldots, c_{k-2}) &= \sum_{i=1}^{k-2} (M_i (c_0 + c_1 P_i + \cdots + c_{k-2} P_i^{k-2}) - E_i)^2 \end{aligned} \tag{5}$$

A feladatot elméletben a következő módon oldjuk meg: megkeressük a G függvény pozitív lokális minimumhelyeit, és ezek közül kiválasztjuk a(z egyik) legkisebb függvényértékhez tartozót. A lokális minimum létezésének elégséges feltétele, hogy:

$$G'(c_0, c_1, \ldots, c_{k-2}) = (0, 0, \ldots, 0)$$

és a

$$D(c_0, c_1, \ldots, c_{k-2}) := \begin{pmatrix} G''_{c_0,c_0}(c_0, c_1, \ldots, c_{k-2}) & \cdots & G''_{c_{k-2},c_0}(c_0, c_1, \ldots, c_{k-2}) \\ \vdots & \ddots & \vdots \\ G''_{c_0 c_{k-2}}(c_0, c_1, \ldots, c_{k-2}) & \cdots & G''_{c_{k-2},c_{k-2}}(c_0, c_1, \ldots, c_{k-2}) \end{pmatrix}$$

mátrix pozitív definit.

3.9 Megjegyzés Ez persze nem algoritmus, csupán vázlat, amiből látható, hogy a probléma elméletben megoldható. A feladatot a *Mathematica* képes kezelni a beépített függvények segítségével: a gyakorlatban a *Fit*[7] és a *Regress*[8] függvénnyel dolgozunk.

3.10 Definíció A spline-függvény, amit illeszteni szeretnénk:

$s(x) := a_{i,0}+a_{i,1}x+a_{i,2}x^2$, ahol $\quad L(1-K_{i,1}) \leq x \leq L(1-K_{i,2}) \quad (i = 1, \ldots, k)$

A függvényre a következő tulajdonságokat követeljük meg:

(1) folytonos,

(2) egyszer differenciálható,

(3) a korcsoporthoz tartozó intervallum alatti terület megegyezik a korcsoportban lévő eszközök (százalékos) mennyiségével.

3.11 Megjegyzés A harmadik feltétel a többletfeltétel.

Ez egyenletekben kifejezve a következőket jelenti:

$$\begin{aligned} a_{i,0} + a_{i,1}(1-K_{i,2}) + a_{i,2}(1-K_{i,2})^2 &= \\ = a_{i+1,0} + a_{i+1,1}(1-K_{i+1,1}) + a_{i+1,2}(1-K_{i+1,1})^2 & \qquad i = \{1, \ldots, k-1\} \end{aligned} \tag{6}$$

Ez a folytonosság. A deriválhatóság pedig:

$$a_{i,0} + 2a_{i,1}(1-K_{i,2}) = a_{i+1,0} + 2a_{i+1,1}(1-K_{i+1,1}) \qquad i = \{1, \ldots, k-1\} \tag{7}$$

A harmadik egyenlet pedig a területtartás:

$$\begin{gathered} \int_{L(1-K_{i,1})}^{L(1-K_{i,2})} s(x) = E_i \\ a_{i,0}(L(1-K_{i,2}) - L(1-K_{i,1})) + \frac{1}{2}a_{i,1}(L(1-K_{i,2})^2 - L(1-K_{i,1})^2) + \\ + \frac{1}{3}a_{i,2}(L(1-K_{i,2})^3 - L(1-K_{i,1})^3) = E_i \qquad i = \{1, \ldots, k\} \end{gathered} \tag{8}$$

[7] a *Fit* a legkisebb négyzetek mószerével fügvényt illeszt

[8] Ugyanaz, mint a *Fit*, csak kiírja a függvény és a pontok távolságát és pár másik paramétert

A (6) és a (7) $(k-1)$—$(k-1)$ számú, míg (8) k számú egyenletet jelent. Összesen $3k$ ismeretlenünk és $3k-2$ lineáris egyenletünk van. Az egyenletek egymástól függetlenek, és ez azt jelenti, hogy két szabad paraméterünk marad. Könnyen látható, hogy bármely két együtthatóját választjuk ki $s(x)$-nek, a többi együttható ennek a kettőnek lineáris függvénye lesz (hiszen az egyenletrendszer lineáris egyenletrendszer). Az egyszerűség kedvéért válasszuk $a_{0,0}$-t és $a_{0,1}$-t szabad paraméternek.

Az s és az f függvény távolságát $\mathcal{L}^2[0,L]$-en értelmezett norma segítségével határozzuk meg (a távolság nagysága nem lényeges, a feladatunk csak ennek a távolságnak a minimalizálása):

$$D(s,f) = \int_0^L (s(x) - f(x))^2 dx$$

A gyakorlatban a *Mathematicával* ezt a feladatot a *Solve*[9] és a *FindMinimum*[10] segítségével oldjuk meg.

Végül már csak az évekre bontás maradt hátra. $[0 \ldots L]$-et L számú intervallumra bontjuk, és ezen intervallumok felett integráljuk s-et.

[9]egyenletrendszerek megoldására képes, megadja, hogy a kötött változók hogyan függnek a szabad paraméterektől, $c_{0,0}$-tól és $c_{0,1}$-től

[10]meghatározza a $D(s,f) = \int_0^L (s(x)-f(x))^2 dx$ minimumát úgy, hogy ezt a szélsőértéket a $c_{1,0} = f(0)$ és a $c_{1,1} = f'(0)$ kezdeti értékekkel kezdi el keresni

4. Állagcsoportok konstruálása állagmutatóból

4.1. A feladat megfogalmazása

4.1 Definíció Egy eszközcsoport *újrabeszerzési-* vagy *bruttóérték* az eszközcsoport eszközeinek értéke, ha mindegyik eszköz új volna. Egy eszközcsoport *nettó értéke* az eszközcsoport mostani értéke.

4.2 Definíció Egy eszközcsoport *állagmutatóján* azt a (százalékban megadott) számot értjük, amely megadja, hogy az összes eszköz mostani értéke hanyadrésze az újrabeszerzési értékének (amennyibe kerülnének az eszközök, ha újak lennének).

4.3 Megjegyzés Az eszközcsoport újrabeszerzési értékének és nettó értékének a hányadosának a 100-szorosa az állagmutató.

4.4 Definíció Legyen adott egy eszközcsoport állagcsoportokra osztva, úgy, hogy ismerjük, hogy az eszközök hogyan oszlanak meg az állagcsoportok között. Az ilyen eloszlásokat *állagmegoszlásnak* nevezzük.

4.5 Definíció Egy eszközcsoporthoz tartozó állagcsoportban lévő eszközök arányát nevezzük az állagmegoszlás egy *tagjának*.

Vannak eszközfajták, amelyeknél jól tagoltan állnak rendelkezésre az adatok (vagyis az eszközfajtán belül az eszközöket a szektorok, ágazatok és megyék szerint tagoljuk). Azonban az egy csoportba tartozó eszközöknél (ugyanaz a szektor, megye, ágazat) csak az újrabeszerzési érték, a nettó érték és az állagmutató ismert.

Az adatokat az előzőekben ismertetett PIM modell használja fel. Márpedig a PIM modellnek az adatok éves bontásban kellenek. A feladatunk az, hogy ezekből az adatokból elfogadható hibával rendelkező kormegoszlásokat állítsunk elő. Nyilvánvaló, hogy az egy csoportba tartozó eszközöknél nem tudjuk megbecsülni a kormegoszlás hibáját. Az összesített adatoknál a kormegoszlás hibája már elfogadható nagyságú.

4.6 Példa Tekintsünk a magyarországi víziközművekre egy fiktív (elképzelt) példát. Eszerint a víziközműveket hatféle csoportra osztjuk:

(1) vízmű $L = 50$	(4) víztároló $L = 50$
(2) víztisztító $L = 35$	(5) szennyvízgyűjtő $L = 60$
(3) vízelosztó $L = 60$	(6) szennyvíztisztító $L = 35$

A víziközműveknek van területi, de nincs szektoros (a víziközművek mindig az önkormányzati szektorba tartoznak) vagy ágazati tagolása (ellentétben a feladat általános megfogalmazásával). Ezenkívül az állagmutatókat (bizonyos szempontokat figyelembe véve) mesterségesen állítjuk elő. Itt csak az országos adatok hitelesek, csak ezeknek becsüljük meg a pontosságát.

4.7 Megjegyzés Az állagmutatókat (20, annyi, amennyi megye van) egy $[40, 80]$-on, míg a bruttó értéket (súlyokat) $[50, 150]$-on egyenletes eloszlás segítségével konstruáljuk.

4.2. A megoldás

Először is definiáljunk egy távolságot az állagmegoszlásokra (olyan állagmegoszlásról van szó, ahol egy eszközcsoporthoz öt állagcsoport tartozik mindig):

$$\sqrt{\frac{\sum_{k=1}^{5}(\alpha_{1,k}-\alpha_{2,k})^2}{5}}$$

ahol $\alpha_{i,j}\quad(i=\{1,2\}, j=\{1\ldots5\})$ az állagmegoszlás tagjait jelenti. Ez valóban távolság (származtatható az l^2 normából, annak speciális esete).

Az állagmegoszlás tagjaira adottak a következő egyenletek és egyenlőtlenségek:

$$\sum_{k=1}^{5}(0.2k-0.1)\alpha_{1,k}=A \qquad \sum_{k=1}^{5}\alpha_{1,k}=100 \qquad 0\le\alpha_{1,i}\le100 \tag{9}$$

4.8 Definíció A két egyenlet miatt három szabad paraméterünk van. Legyen g az állagmegoszlás sűrűségfüggvénye. Ekkor

$$g(\alpha_{1,2},\alpha_{1,3},\alpha_{1,4})=g(\alpha_{2,2},\alpha_{2,3},\alpha_{2,4}) \tag{10}$$

ahol α_1 és α_2 tetszőleges állagmegoszlások, azzal a feltétellel, hogy kielégítik a két egyenletet. Ekkor azt tesszük fel, az eloszlásunk a középső három változóban egyenletes.

4.9 Állítás *Teljesül, hogyha α valószinűségi állagmegoszlás $\alpha_{i_1},\alpha_{i_2},\alpha_{i_3}$ változóban egyenletes, akkor egyenletes $\alpha_{j_1},\alpha_{j_2},\alpha_{j_3}$ változókban is (nyilván $0\le i_1,\ldots,i_3,j_1,\ldots,j_3\le3$).*

Bizonyítás. Tudjuk, hogy a (9)-ben meghatározott egyenletek fennállnak. Így három állagmegoszlás-tag meghatározza az állagmegoszlást, vagyis bármely három állagmegoszlás-tagból bármely három állagmegoszlás tagot meg tudjuk a két egyenlet segítségével. Mivel (9)-beli egyenletekben az állagmegoszlás-tag lineáris, ezért az $\{\alpha_{j,1}, \alpha_{j,2}, \alpha_{j,3}\}$ tagok lineárisan függnek $\{\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}\}$-tól. Ha az eloszlás sűrűségfüggényét (g) $\{\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}\}$ helyett $\{\alpha_{j,1}, \alpha_{j,2}, \alpha_{j,3}\}$ változók függvényeként akarjuk megadni, akkor a függvényt meg kell szorozni az $\{\alpha_{i,1}, \alpha_{i,2}, \alpha_{i,3}\}\{\alpha_{j,1}, \alpha_{j,2}, \alpha_{j,3}\}$ változótranszformációhoz tartozó Jacobi-determinánssal. Ez azonban konstans (mivel a változók lineárisan függnek egymástól). Tehát ha $g(\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}) = g(\beta_{i_1}, \beta_{i_2}, \beta_{i_3})\ \forall \alpha, \beta$, akkor $g(\alpha_{j_1}, \alpha_{j_2}, \alpha_{j_3}) = g(\beta_{j_1}, \beta_{j_2}, \beta_{j_3})\ \forall \alpha, \beta$

A feltételezett eloszlás alapján generálunk 1000 állagmegoszlást minden állagmutatóhoz. Ezután

- Állagmutatónként összegezzük a véletlen állagmegoszlásokat: ez lesz a tapasztalati várható állagmegoszlás. Ez már jól közelíti a várható állagmegoszlást.
- Súlyozva összegezzük a véletlen állagmegoszlásokat, így megkapjuk a összesített véletlen állagmegoszlásokat
- Ezeket összegezve megkapjuk az összegzett várható állagmegoszlást
- Az összegzett véletlen és összegzett várható állagmegoszlások távolságát átlagolva megkapjuk az átlagos távolságot

4.10 Megjegyzés Semmilyen más eloszlás (exponenciális, normális, stb.) nem definiálható 'jól' a (9) által meghatározott téren

4.11 Megjegyzés Nem szabad elfelejteni: a feltételezést (hogy tudjuk, hogy az állagmutatóhoz milyen eloszlás alapján rendelünk állagmegoszlást) semmilyen tapasztalati tény nem támasztja alá. Másfajta eloszláshoz másfajta várható állagmegoszlás tartozik. Ekkor az átlagos távolság nem csökkenhet egy bizonyos korlát alá[11].

[11] Ha nem elfajult eloszlással van dolgunk, akkor ez így is közel lesz az egyenletes eloszlás alapján generált várható állagmegoszláshoz (a távolság kb. $2 - -5\%$)

4.3. Szemléltető példák

Az első példában 5 állagmutatóhoz készítettünk állagmutatócsoportokat (állagmegoszlásokat), majd ezeket összegeztük. Az állagmutatókat egy $[40, 80]$ közötti egyenletes eloszlás segítségével konstruáltuk. Mindegyik állagmutatót ugyanakkora súllyal vettük figyelembe. Az eredmény:

	$0-20\%$	$20-40\%$	$40-60\%$	$60-80\%$	$80-100\%$
várható állagcsoport	13.01	16.57	19.98	20.86	29.55
átlagos távolság:	5.06		gyorsaság[12]:		35.48 sec

A második példában 15 állagmutatóhoz konstruáltunk állagcsoportokat, majd ezeket összegeztük (mindegyiket azonos súllyal vettük figyelembe).Az állagmutatók ugyanazok, csak mindegyiket háromszor vettük figyelembe. Az eredményként kapott összegzett állagcsoportok:

	$0-20\%$	$20-40\%$	$40-60\%$	$60-80\%$	$80-100\%$
várható állagcsoport	12.91	16.88	19.46	21.40	29.33
átlagos távolság:	2.95		gyorsaság[13]:		108.7 sec

A harmadik példában 17 állagmutatóhoz konstruálunk állagmegoszlásokat, de az összegzésnél nem azonos súllyal vesszük őket figyelembe: 2-t egyharmadnyi, 15-t pedig $\frac{1}{45}$-nyi súllyal veszünk figyelembe. Az állagmutatókat [40,80]-on értelmezett egyenletes eloszlás segítségével konstruáljuk. Az állagmutatók:

	$0-20\%$	$20-40\%$	$40-60\%$	$60-80\%$	$80-100\%$
várható állagcsoport	7.43	10.72	14.89	23.18	43.76
átlagos távolság:	4.72		gyorsaság[14]:		117.98 sec

Látható, hogy a legalacsonyabb átlagos hibát akkor nyertük, amikor 15 azonos súlyú adatot összegeztünk. Rosszabbul jártunk, ha 5, de azonos súlyú állagcsoportot összegeztünk, mint ha 17-et, ami közül kettőnek a súlya egyharmadnyi.

Itt látható a fejezet elején említett viziközművekre vonatkozó példa:

	$0-20\%$	$20-40\%$	$40-60\%$	$60-80\%$	$80-100\%$
vízmű	12.88	16.42	18.78	22.48	29.41
víztisztító	12.54	16.60	19.26	20.97	30.61
vízelosztó	10.38	14.23	18.00	23.04	34.33
víztároló	11.78	15.91	19.22	23.12	29.94
szennyvízgyüjtő	13.45	16.90	18.57	22.41	28.64
szennyvíztisztító	10.55	14.45	18.36	24.33	32.29

Az átlagos hiba:

	vízmű	víztisztító	vízelosztó
átlagos távolság	2.67	2.58	2.52
	víztároló	szennyvízgyüjtő	szennyvíztisztító
átlagos távolság	2.82	2.67	2.63

Az algoritmus gyorsasága 899.48 sec. Látható, hogy a végeredményként kapott várható állagmegoszlások igen közel vannak egymáshoz. Ez csuppán annak a következménye, hogy mesterségesen generáltuk a súlyokat és az állagmutatókat. Mivel mind a hat esetben ugyanolyan módon generáltuk az adatokat, a végeredmény is hasonló lett.

5. Összefoglalás, további tervek

5.1. A PIM fejlesztésének további irányai

Folytatódik a PIM modell algoritmusainak megírása, a meglevőek hatékonyabb megvalósítása, majd végül az algoritmusok programcsomagba szervezése.

Gondolkozunk a PIM-modell olyan megvalósításán, ahol az adatokat nemcsak éves (korcsoportos), hanem állag szerinti tagolásban is tárolni lehetne (egyes eszközfajtáknál, pl.:utak,épületek,gátak,stb. sokkal valósághűbben adja vissza a végbemenő folyamatokat)

Alkalmassá kell tenni a modellt arra, hogy szakértők szükség esetén módosíthassák az adatokat. Ezzel kezelhetővé válnának a szektorközi nagyobb tőkemozgások (pl.: iskolák egyházi tulajdonba juttatása), valamint a nagyarányú fejlesztések (épületek, utak, gátak,stb. esetében).

A hibabecslési módszerek nincsenek egységesen kidolgozva. Tisztázni kell a becslési hiba mibenlétét (pl.: az állagcsoportok közötti távolság hogyan hozható összefüggésbe a kormegoszlások távolságával, a PIM modellben a hibás selejtezések mértékével, stb). Tapasztalati mérések alapján kell tesztelnünk, hogy az algoritmusaink alkalmazása mennyire gyengíti az adatok minőségét (gondolok a függvény-illesztésre, az állag-kor transzformációra, az állagcsoportok konstruálására).

Érzékenységvizsgálattal kell eldönteni, hogy a maximális élettartam szempontjából mennyire kell egységesnek lennie az eszközcsoportnak.

Jó volna, ha empirikus felmérések alapján dönthetnék el, az egyes eszközfajtáknál milyen értékcsökkenési függvényt alkalmazhatunk.

5.2. Összegzés

Ha tömören szeretném megfogalmazni, hogy milyen végkövetkeztetésre jutottam a KSH-nak végzett munka során, akkor a következőket mondanám: a közgazdaságtanban - ugyanúgy, ahogy a fizikában, a kémiában és más természettudományokban - fontos szerep jut a méréseknek. Mérés nélkül nem tudjuk elvégezni megfigyeléseinket, felépíteni modelljeinket. Csakhogy a közgazdaságtan társadalomtudomány, alapfogalmai sokkal bizonytalanabbak, határozatlanabb körvonalúak, mint azok, amelyekkel bármely más természettudomány dolgozik. A fizikában a tömeget, a sebességet, a gyorsulást, az erőt és a közöttük lévő összefüggéseket könnyen elvégezhető és sokszor megismételhető kisérletekkel tudjuk mérni. A közgazdaságtanban ugyanakkor máig viták folynak arról , hogy miképpen mérjék a GDP-t, az inflációt, a külkereskedelmi mérleget, a munkanélküliséget, stb. Ráadásul a mérésekből kevés áll rendelkezésre, kisérleteket pedig egyáltalán nem lehet végezni, csak egyszeri jelenségeket megfigyelni.

Ennek ellenére (mindaddig, amíg méréseink eredményét számok formájában akarjuk viszontlátni) nem térhetünk ki a matematikai módszerek alkalmazása elől. Ám az alkalmazás során érdemes szem előtt tartani pár alapvető elvet: törekednünk kell arra, hogy modellünk felépítése minél egyszerűbb legyen, minél kevesebb feltevésre épüljön. Törekednünk kell arra, hogy a modell feltevései ne csak önmagukban, hanem egymással összefüggésben is megállják a helyüket[5].

Ezért a munka során meg kellett kötni a szükséges kompromisszumokat, amellett, hogy a modell lényeges sarokpontjaiból nem engedtünk. A PIM modell kezdeti adatainak egy része előállítható a 3 és a 4. fejezetben ismertetett két algoritmus segítségével. De tudnunk kell, hogy adataink bizonytalanok. A méréseket helyesbíteni kell, akár pár éven belül elvégzett új mérésekkel, akár úgy, hogy a nemzetgazdaságba bekerülő, valamint kilépő eszközökről folyamatosan (évente) készítünk új felméréseket. Az adatgenerálás érdekében lemondtunk az avulási és a selejtezési függvényekről, ami - bár érdekes koncepcióra épült - lehetetlenné tette volna az adatkonverziót az állagcsoportok és a korcsoportok között. Megmaradt viszont a modellben a maximális élettartam (nem hagytuk, hogy ezt a fogalmat átlagos maximális élettartammal) helyettesítsék.

[5]A múlt század első felében J. M. Keynes lázasan sürgette a kormányokat, hogy a gazdasági fejlődés beindítása érdekében nagy állami beruházásokkal csökkentsék a munkanélküliséget, növeljék a fogyasztást (jó ideig persze senki nem hallgatott rá) Ezért ma az EU a lisszaboni program néven ismert beruházásélénkítő és infrastruktúrafejlesztő csomag segítségével próbálja beindítani a gazdasági növekedést. Lehet azonban, hogy többet lendít az ügyön a szoros Schröder-Chirac eggyüttműködés, valamint a Németországban mostanában megszavazott gazdasági megszorító intézkedéscsomag.

A. Függvény-illesztés implementációja

Itt látható a 3.fejezetben ismertett spline-függvény illesztés *Mathematica* implementációja.

```
FvIlleszt[AllagCsoportok_?MatrixQ,VarhatoElettartam_]:=
  Module[{x,i,j,k,n,kcs,ih,ik,ef,efp,lkcs,espt,esp,desp,
  tesp,spt,sp,splf,minimum,tav,megoldas,EVT,ertmp,eval,
  psplf
,
```

megvizsgáljuk, hogy a bemeneti adatok megfelelnek-e a feltételeknek

```
  n=Length[AllagCsoportok[[1]]];
  If[(n<3||n>10||n!=Length[AllagCsoportok[[2]]] ||
    Sum[AllagCsoportok[[1,i]],i,1,n]==100 || Sum[AllagCsoportok[[2,i]],i,1,n]==100)
    ,Print["Hibásak a bemeneti adatok"]];
```

Végrehajtjuk a transzformációt állagcsoportokból korcsoportokba, majd a korcsoportokra 2- vel kisebb fokú polinomot illesztünk, mint ahány korcsoport van.

```
  kcs={Reverse[AllagCsoportok[[1]]],Reverse[AllagCsoportok[[2]]]/100*VarhatoElettartam};
  ik=Table[Sum[(kcs[[2,j]]-kcs[[2,i]])/2,{i,1,n}]];
  ef[x_]=Fit[Transpose[{ik, Table[kcs[[1,i]]/kcs[[2,i]],{i,1,n}]}], Table[x^i,{i,0,n-2}],x];
```

A polinomhoz másodfokú spline-függvényt illesztünk, ami folytonos , differenciálható,és a korcsoportokhoz tartozó eszközmennyiség megegyezik a függvénynek korcsoporthoz tartozó intervallum alatti területével

```
  For[i=1,i<=n,If[i!=1,ih=Append[ih,ih[[i-1]]+kcs[[2,i]]]//Flatten,ih={kcs[[2,1]]}];i++];
  spt=Array[sp,{n,3}];
  ih=Prepend[ih,0];
  esp={Table[spt[[i-1,1]]+spt[[i-1,2]]*ih[[i]]+spt[[i-1,3]]*ih[[i]]^2,{i,2,n}],
    Table[spt[[i,1]]+spt[[i,2]]*ih[[i]]+spt[[i,3]]*ih[[i]]^2,{i,2,n}]};
  desp={Table[spt[[i-1,2]]+2*spt[[i-1,3]]*ih[[i]],{i,2,n}],
    Table[spt[[i,2]]+2*spt[[i,3]]*ih[[i]],{i,2,n}]};
```

```
tesp={Table[((ih[[i+1]]-ih[[i]]))*spt[[i,1]]+((ih[[i+1]]^2-ih[[i]]^2))*spt[[i,2]]/2+
  ((ih[[i+1]]^3-ih[[i]]^3))*spt[[i,3]]/3,{i,1,n}],kcs[[1]]};
ertmp=Flatten[Transpose/@{esp,desp,tesp},1];
er=Table[ertmp[[i,1]]==ertmp[[i,2]],{i,1,3n-2}];
eval=Drop[Flatten[spt],2];
megoldas=Simplify[Expand[Solve[er,eval]]];
For[i=0,i<3n-2,(i++);
spt[[Quotient[i+4,3],Mod[i+2,3,1]]]=megoldas[[1,i,2]]];
splf[x_]:=[Sum[If[IntervalMemberQ[Interval[{ih[[i]],ih[[i+1]]}],x],1,0]*
  ((spt[[i,1]]+spt[[i,2]]x+spt[[i,3]]x^2)),{i,1,n}];
tav=((((((splf[0]-ef[0]))^2+
  ((splf[VarhatoElettartam]-ef[VarhatoElettartam]))^2))*VarhatoElettartam/2000+
  Sum[((((splf[2*m/2000*VarhatoElettartam]-ef[2*m/2000*VarhatoElettartam]))^2+
    ((splf[((2*m+1))/2000*VarhatoElettartam]-
    ef[((2*m+1))/2000*VarhatoElettartam]))^2))/((3*2000))*2*VarhatoElettartam,
    {m,1,1000}]))//Simplify;
megoldas=FindMinimum[tav,{sp[1,1],Evaluate[ef[0]]},{sp[1,2],
Evaluate[ef'[0]]}];
sp[1,1]=megoldas[[2,1,2]];
sp[1,2]=megoldas[[2,2,2]];
```

Kirajzolja a végeredményként kapott függvényeket és kiírja a kormegoszlást (évekre bontva).

```
efp=Plot[ef[x],{x,0,VarhatoElettartam},DisplayFunction->Identity];
lkcs=ListPlot[Transpose[{ik,Table[kcs[[1,j]]/kcs[[2,j]],{j,1,n}]}],
  PlotStyle->{RGBColor[0,1,0],PointSize[0.02]},DisplayFunction->Identity];
psplf=Plot[splf[x],{x,0,VarhatoElettartam},DisplayFunction->Identity];
Print["Az életkor szerinti sűrűségfüggvény és a spline-függvény és a korcsoportok:"];
Show[{efp,lkcs,psplf},AxesOrigin->{0,0},PlotRange->All,DisplayFunction:>$DisplayFunction];
EVT=Table[NIntegrate[splf[x],{x,i-1,i}],{i,1,VarhatoElettartam}];
Print["Az évekre bontott adatok:"];
Print[EVT];
Print["Megjegyzés: Mindegyik adat százalékos"];
EVT];
```

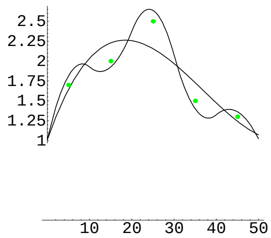
Íme két szemléltető példa a program futására. az első példa bemenete:

```
ClearAll[AllagCsoportok, VarhatoElettartam, EletkorLista];
```

```
AllagCsoportok = {13, 15, 25, 20, 17}, {20, 20, 20, 20, 20}};
VarhatoElettartam = 50;
EletkorLista = FvIlleszt[AllagCsoportok, VarhatoElettartam];
```

Majd a kimenet:

```
Az életkor szerinti sűrűségfüggvény és a spline-függvény és a korcsoportok:
```



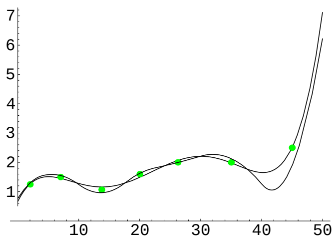

```
Az évekre bontott adatok
{1.12851, 1.32814, 1.50053, 1.64569, 1.76361, 1.85429, 1.91773, 1.95394, 1.96291, 1.94465,
1.9067, 1.87926, 1.86989, 1.87858, 1.90533, 1.95015, 2.01304, 2.09398, 2.193, 2.31007,
2.4367, 2.53882, 2.60792, 2.644, 2.64706, 2.6171, 2.55412, 2.45812, 2.3291, 2.16706,
1.98151, 1.81046, 1.66344, 1.54042, 1.44143, 1.36645, 1.31548, 1.28853, 1.2856, 1.30668,
1.34527, 1.37532, 1.39034, 1.3903, 1.37523, 1.34511, 1.29995, 1.23975, 1.16451,1.07422}
Megjegyzés: Mindegyik adat százalékos
```

A második példa bemenete (figyeljük meg, hogy az állagcsoportoknál az intervallumok nem ugyanolyan hosszúak):

```
ClearAll[AllagCsoportok, VarhatoElettartam, EletkorLista];
AllagCsoportok = 25, 20, 15, 8, 8, 9, 5, 20, 20, 15, 10, 15, 12, 8;
VarhatoElettartam = 50;
EletkorLista = FvIlleszt[AllagCsoportok, VarhatoElettartam];
```

Majd a kimenet:

```
Az életkor szerinti sűrűségfüggvény és a spline-függvény és a korcsoportok:
```




```
Az évekre bontott adatok
{0.925964, 1.18794, 1.38099, 1.5051, 1.56594, 1.58617, 1.57144, 1.52176, 1.43713, 1.31755,
1.17595, 1.06409, 0.994905, 0.96839, 0.984549, 1.04338, 1.14489, 1.28767, 1.43966, 1.56879,
1.67366, 1.75427, 1.81118, 1.85717, 1.905, 1.95522, 2.00786, 2.06289, 2.12032, 2.18015,
2.23677, 2.26768, 2.26727, 2.23554, 2.17248, 2.0781, 1.9524, 1.79538, 1.60703, 1.38736,
1.16647, 1.06475, 1.1123, 1.30913, 1.65523, 2.1506, 2.79524, 3.58915, 4.53234, 5.6248}
Megjegyzés: Mindegyik adat százalékos
```

B. Állagcsoport konstruálás implementációja

Az állagmutatóból állagcsoportot konstruáló algoritmus *Mathematica* implementációja. Az algoritmus kiszámolja az állagmutatókhoz tartozó várható állagmegoszlásokat, az összesített állagmegoszlást, valamint az átlagos hibát.

```
AllagKeszito[buv_ ?MatrixQ, eo_?VectorQ] :=
  Module[
    {i, j, k, l, a, suly = {0.1, 0.3, 0.5, 0.7, 0.9}, n = Length[buv], acs,
      alfa, meg, neo = Complement[Range[5], eo], er, tvsz, jk, m = 1000,
      TR = Table[True, {i, 1, 5}], vacs, ovacs, evacs, kcs, atv},
```

Azoknak az állagmegoszlásoknak a kiszámítása, amelyek kielégítik az egyenleteket (hogy az állagmegoszlás-tagok összege 100 és a súlyozott összege az állagmutató, és mindegyik állagmegoszlás-tag 0 és 100 közé esik).

```
    alfa = Array[(buv[[#, 1]]/buv[[#, 2]]) &, n]*100;
    acs = Array[a[#1, #2] &, {Length[buv], 5}];
    meg = Array[
        Solve[{Plus @@ acs[[#]] == 100, Plus @@ (acs[[#]]*suly) == alfa[[#]]},
             Table[a[#, neo[[i]]], {i, 1, 2}]] &, n];
    Array[(acs[[#1, neo[[#2]]]] = Simplify[meg[[#1, 1, #2, 2]]]) &, {n, 2}];
    jk = Table[{}, {i, 1, n}];
    For[l = 1,
      l <= n, {Clear[a]; a[l, eo[[1]]] := 2i; a[l, eo[[2]]] := 2j;
        a[l, eo[[3]]] := 2k;
        jk[[l]] =
          Select[Flatten[
              Table[Table[
                  Table[acs[[l]] , {k, 1, 50 - i - j}], {j, 1, 50 - i}], {i,
                  1, 50}],
              2], (Table[#[[i]] >= 0, {i, 1, 5}] == TR &&
                  Table[#[[i]] <= 100, {i, 1, 5}] == TR) &]}; l++];
    tvsz[x_, y_] := Sqrt[Sum[(x[[i]] - y[[i]])^2, {i, 1, 5}]/5];
```

A véletlen állagmegoszlások megkonstruálása. A várható állagmegoszlás kiszámítása. Az átlagos hiba kiszámítása.

```
  m = 1000; vacs = Table[{}, {i, 1, n}];
  For[j = 1,
    j <= n, (tm = jk[[j]];
      vacs[[j]] = Table[tm[[Random[Integer, {1, Length[tm]}]]], {i, 1, m}]);
     j++];
  ovacs =
    Plus @@ N[Transpose[buv][[2]]*vacs]/Plus @@ N[Transpose[buv][[2]]];
  evacs = Table[Sum[vacs[[j, i]], {i, 1, m}]/m, {j, 1, n}];
  kcs = Plus @@ ovacs/m; atv = Sum[tvsz[ovacs[[i]], kcs], {i, 1, m}]/m;
```

A végeredmény kiírása.

```
  Print[
    TableForm[{{"", "0-20%", "20-40%", "40-60%", "60-80%",
          "80-100%"}, {"várható acs.:", N[kcs]} // Flatten}]];
  Print[TableForm[{{"átlagos távolság a várható acs.-tól:", atv}}]];
  Print["Megjegyzés: az eloszlás ", eo, "-ban egyenletes."];
  kim = Table[evacs[[i]]*buv[[i, 2]], {i, 1, n}];
  Evaluate[{kcs*(Plus @@ Array[buv[[#, 2]] &, n]/100), atv, kim/100}]
]
```

A példák, amiken az algoritmust teszteltük, megtalálhatóak a 4. fejezetben a 'Szemléltető példák' részben.

Hivatkozások

[1] Szili László, Tóth János, *Matematika és Mathematica*, ELTE Eötvös kiadó, Budapest, 1996.

[2] B.N.Pschnechny, Yu. M. Danilia, *Numerical Methods in Extremal Problems*,MIR publishers, Moscow, 1972

[3] Stoyan Gisbert, Takó Galina, *Numerikus módszerek*, ELTE Eötvös kiadó, Budapest, 19993.

[4] Vajda János, *Kémiai számítástechnika - A legkisebb négyzetek módszere*, ELTE TTK Kémiai kibernetika laboratórium, Budapest, 1976.

[5] Becskei Péter, *Gazdaság és Statisztika: Folyamatos leltárazási módszer (PIM) alkalmazása Magyarországon*, Budapest, 2003.

[6] OECD, *Measurement of Capital Stocks, Consumption of Fixed Capital and Capital Services.*