

Equacions estructurals amb lavaan

Jose M Sallan *

Resum

L'anàlisi d'equacions estructurals és una tècnica per avaluar models amb relacions complexes entre variables (*path analysis*), amb variables observables i variables latents. Un cas especialment interessant d'aplicació és l'anàlisi factorial confirmatòria, que permet avaluar la validesa de constructe d'escala de mesura. En aquest article es presenta el paquet de R `lavaan`, que permet avaluar amb facilitat models d'equacions estructurals. Es reporta informació sobre la sintaxi del model, així com de les possibilitats del paquet.

1 Introducció

L'anàlisi d'equacions estructurals és una tècnica de gran utilitat en l'avaluació de models amb relacions complexes entre variables, com els definits en la tècnica de *path analysis*. També és la tècnica a utilitzar quan els models inclouen variables latents, mesurades a partir de dues o més variables observables. Un cas particular especialment rellevant d'equacions estructurals és l'anàlisi factorial confirmatòria, utilitzat per avaluar la validesa de constructe d'escala de mesura.

R és un entorn de programari per càlcul matricial i de generació de gràfics, que també pot usar-se per a càlculs propis de mètodes estadístiques [9]. En ser programari lliure, R es pot ampliar fàcilment mitjançant paquets creats per la comunitat. Aquest fet confereix a R un gran potencial per al seu ús en anàlisi de dades, però en contrapartida té una interfície d'usuari poc amigable, i obliga als seus usuaris a escollir el paquet més adequat d'entre els diferents desenvolupats per a una mateixa tècnica.

En aquest document pretenem presentar `lavaan`, un paquet de R per a anàlisi d'equacions estructurals amb funcionalitats suficients per realitzar les anàlisis requerides en gran nombre d'estudis que empenen aquesta tècnica. Començarem amb una breu descripció de les tècniques d'equacions estructurals. Seguidament descriurem la sintaxi amb què s'entren models amb `lavaan`, i les possibilitats d'anàlisi del paquet.

2 Models d'equacions estructurals

Els models d'equacions estructurals (SEM) ens permeten avaluar si un model que expressa una determinada relació entre les variables observables s'ajusta a les dades empíriques [10]. L'avaluació d'un SEM requereix una primera fase d'especificació del model, seguida de les fases d'identificació i d'estimació. En la fase d'*especificació* es defineixen les relacions a avaluar entre variables. Aquestes relacions poden ser més complexes que les definides en els models lineals, en què hi ha una única variable dependent i poden existir múltiples variables independents. En SEM podem tenir tant variables observables com a variables latents (no observables directament). De fet, amb SEM podem avaluar, entre d'altres models:

- *Path analysis*: Models que estableixen relacions entre variables observables. Aquestes relacions poden incloure relacions recíproques (equacions simultànies) i recursives.
- *Anàlisi factorial confirmatòria*: model en què s'estableixen unes relacions entre variables observables i unes variables latents. Generalment, es permet que les variables latents correlacionen entre si.

*Department of Management, ETSEIAT, UPC <http://josemsallan.blogspot.com>

- *Equacions estructurals*: model en què s'avalua conjuntament un model de mesura (relació entre variables latents i observables) i un model estructural (relacions entre variables latents).

En la fase d'*identificació* s'estableix la relació entre els paràmetres del model i les covariàncies de les variables observables. En la identificació s'obtenen els *graus de llibertat* del model, definits com la diferència entre el nombre de covariàncies diferents entre variables observables, i el nombre de paràmetres a estudiar en el model. Només es pot avaluar l'ajust del model a les dades en els models sobreidentificats, amb un valor de graus de llibertat positiu.

Finalment, en la fase d'*estimació*, s'obtenen estimadors dels paràmetres del model, i indicadors de l'ajust del model a les dades. El mètode d'estimació més emprat és el de màxima versemblança (*maximum likelihood*), que exigeix que les variables observables segueixin una distribució normal multivariant. Per a mostres no normals hi ha altres mètodes, com el de mínims quadrats ordinaris o l'asimptòtic lliure de distribució.

3 El paquet lavaan

En R existeixen diversos paquets per a anàlisi d'equacions estructurals. Els més coneguts són `sem` [3, 4], `OpenMx` [1] i `lavaan` [8, 6, 7]. `lavaan` permet escriure els models de mesura i estructurals més senzillament que els altres dos paquets, i permet realitzar anàlisis relativament complexes, com l'anàlisi de grups d'equacions estructurals, o l'ús de mètodes d'estimació alternatius al de màxima versemblança. Per instal·lar `lavaan`, cal obrir una consola de R i escriure:

```
install.packages("lavaan")
```

Per comprovar que el paquet s'ha instal·lat correctament, i sempre que vulguem usar el paquet en R, cal escriure a la consola:

```
library(lavaan)
```

4 Especificació de models SEM amb lavaan

Les referències disponibles actualment per a la sintaxi d'especificació de models en `lavaan` són el text introductori de [8] i el manual de referència [7]. Seguidament mostrarem dos exemples d'especificació de models: una anàlisi factorial confirmatòria i un model estructural.

4.1 Un model d'anàlisi factorial confirmatòria

L'exemple d'anàlisi factorial confirmatòria s'obté a partir de nou de les 26 variables d'HolzingerSwineford1939, un conjunt de dades disponible en R i que s'usa freqüentment en documents sobre equacions estructurals. Es tracta dels resultats d'un test de capacitat mental, en el qual s'estableix com a hipòtesi que les nou variables mesuren tres factors, tal com es mostra a la figura 1, associats a diferents capacitats:

- Les variables x_1 , x_2 i x_3 mesuren el factor *visual* representat per la variable latent v
- Les variables x_4 , x_5 i x_6 mesuren el factor *textual*, representat per la variable latent t
- Les variables x_7 , x_8 i x_9 mesuren el factor *veloc*, representat per la variable latent s

El model d'equacions estructurals associat a aquest conjunt d'hipòtesis és:

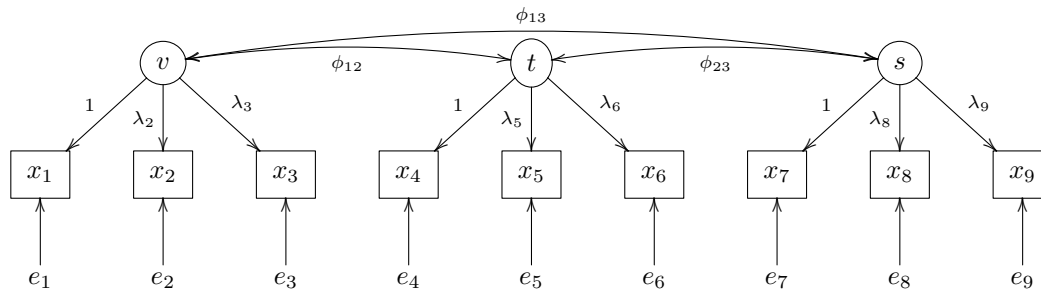


Figura 1: Representació gràfica del model d'anàlisi factorial confirmatòria

$$\begin{aligned}
 x_1 &= v + e_1 \\
 x_2 &= \lambda_2 v + e_2 \\
 x_3 &= \lambda_3 v + e_3 \\
 x_4 &= t + e_4 \\
 x_5 &= \lambda_5 t + e_5 \\
 x_6 &= \lambda_6 t + e_6 \\
 x_7 &= s + e_7 \\
 x_8 &= \lambda_8 s + e_8 \\
 x_9 &= \lambda_9 s + e_9
 \end{aligned}$$

On v, t, s són les variables latents, i les e_i els errors associats a cada variable observable (és a dir, la variabilitat de les x_i no explicada per les variables latents). Usualment, es considera per defecte que no existeix correlació entre els e_i , i que les variables latents tenen correlacions $\phi_{ij} \neq 0$.

L'especificació d'aquest model en sintaxi `lavaan` és:

```

model1 <- 'visual =~ x1 + x2 + x3
textual =~ x4 + x5 + x6
veloc =~ x7 + x8 + x9'
```

Aquesta especificació fa les consideracions per defecte següents:

- Tal i com s'ha especificat a les equacions estructurals, el coeficient que relaciona una de les variables observables associada una variable latent es fixa igual a 1
- Es deixen lliures les correlacions entre variables latents
- Els termes d'error de les variables observables són independents (les correlacions entre errors són zero)

Alternativament, podem fixar el valor d'algun dels paràmetres que per defecte són lliures. Per exemple, fer les correlacions entre variables latents iguals a zero, fent així que aquestes variables siguin ortogonals:

```

model1b <- 'visual =~ x1 + x2 + x3
textual =~ x4 + x5 + x6
veloc =~ x7 + x8 + x9
visual =~ 0*veloc
textual =~ 0*veloc
visual =~ 0*textual'
```

També podem deixar lliure un paràmetre que és fixe per defecte. Per exemple, podem definir un model confirmatori on totes les λ_i siguin lliures, i la variança de les variables latents sigui igual a 1:

```

model1c <- 'visual =~ NA*x1 + x2 + x3
textual =~ NA*x4 + x5 + x6
veloc =~ NA*x7 + x8 + x9

visual =~ 1*visual
textual =~ 1*textual
veloc =~ 1*veloc'

```

Aquestes especificacions per defecte són vàlides per a les funcions `cfa()` i `sem()`, les usades més freqüentment. En [8] pot trobar-se exemples i indicacions per fer les modificacions necessàries en cada cas. Alternativament, podem optar per especificar el model complet, sense elements per defecte, fent servir la funció `lavaan` [6].

4.2 Un model d'equacions estructurals

Per il·lustrar l'especificació de models d'equacions estructurals, farem servir una altre exemple clàssic en la literatura d'equacions estructurals, el model d'estabilitat d'alienació de [11]. Aquest model inclou un *model de mesura* (relació entre variables latents i observables) i un *model estructural* (relacions entre variables latents), tal com es mostra a la figura 2. A [3] o a [6] es poden trobar les dades originals, en aquest cas el nombre d'observacions i la matriu de covariàncies entre variables observables. Les equacions d'aquest model són:

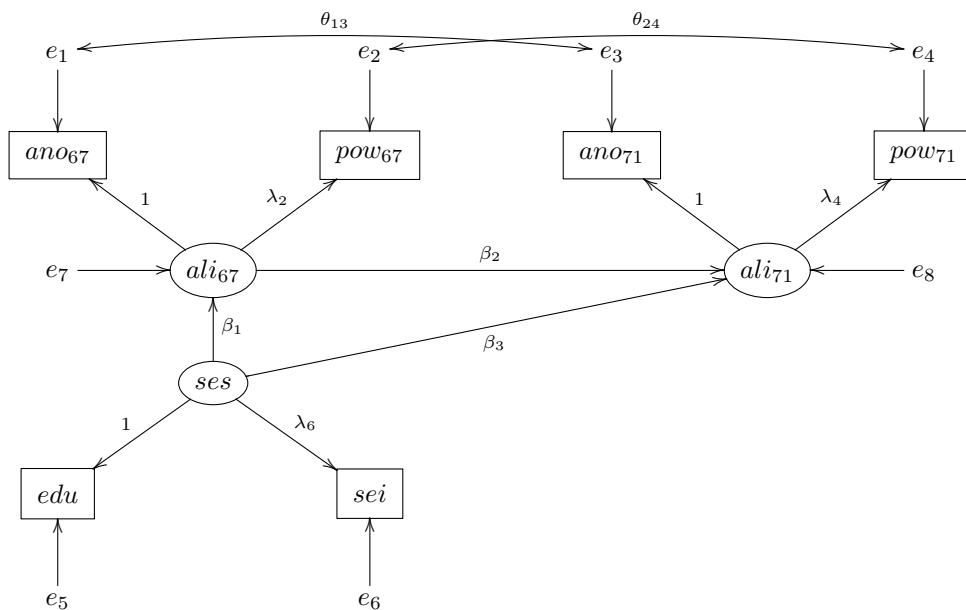


Figura 2: Modelo estructural para el modelo de alienación

$$\begin{aligned}
ano_{67} &= ali_{67} + e_1 \\
pow_{67} &= \lambda_2 ali_{67} + e_2 \\
ano_{71} &= ali_{71} + e_3 \\
pow_{71} &= \lambda_4 ali_{71} + e_4 \\
edu &= ses + e_5 \\
sei &= \lambda_6 ses + e_6 \\
ali_{67} &= \beta_1 ses + e_7 \\
ali_{71} &= \beta_2 ali_{67} + \beta_3 ses + e_8
\end{aligned}$$

A més, el model afegeix correlacions entre les variables d'error e_1 i e_3 (associades a ano_{67} i ano_{71}), i les variables e_2 i e_4 (associades a pow_{67} i pow_{71}). En sintaxi `lavaan`, aquest model s'especifica com:

```
wheaton.model <- '
ali67 =~ ano67 + pow67
ali71 =~ ano71 + pow71
ses =~ edu + sei
ali71 ~ ali67 + ses
ali67 ~ ses
ano67 =~ ano71
pow67 =~ pow71'
```

5 Estimació de models SEM amb lavaan

En `lavaan` disposem de tres funcions per estimar models d'equacions estructurals. La funció `cfa()` permet avaluar models d'anàlisi factorial confirmatòria, que poden ser estimats amb una funció més senzilla que els models estructurals genèrics que són estimats amb la funció `sem()` [5]. La funció `lavaan()` també permet estimar models genèrics d'equacions estructurals, en què s'especifica el model complet. En `sem()` i `cfa()` especifiquem primer el model, i després la font de les dades (podem utilitzar les dades brutes, o bé la matriu de covariàncies i el nombre d'observacions).

5.1 Anàlisi factorial confirmatòria

Començarem per estimar el model de la secció 4.1. Com les dades del model estan incorporades a R, obtindrem l'estimació fent:

```
fit.model1 <- cfa(model1, data = HolzingerSwineford1939)
```

Tal com s'ha formulat, a `fit.model1` tindrem els resultats de l'estimació del model de manera similar a com ho fa Mplus. Si volem que `lavaan` faci l'estimació del model com EQS, LISREL o AMOS, haurem de fer:

```
fit.model1.Wishart <- cfa(model1, data = HolzingerSwineford1939, likelihood="wishart")
```

Per a més informació sobre les diferències entre les dues estimacions, veure [6].

Per a obtenir els resultats, farem:

```
summary(fit.model1, fit.measures=TRUE)
```

El resultat que obtenim és:

```
Lavaan (0.4-9) converged normally after 36 iterations

Number of observations              301

Estimator                           ML
```

```

Minimum Function Chi-square          85.306
Degrees of freedom                   24
P-value                              0.000
10 Chi-square test baseline model:

Minimum Function Chi-square          918.852
Degrees of freedom                   36
P-value                              0.000
15 Full model versus baseline model:

Comparative Fit Index (CFI)          0.931
Tucker-Lewis Index (TLI)            0.896
20 Loglikelihood and Information Criteria:

Loglikelihood user model (H0)        -3737.745
Loglikelihood unrestricted model (H1) -3695.092
25 Number of free parameters          21
Akaike (AIC)                         7517.490
Bayesian (BIC)                       7595.339
Sample-size adjusted Bayesian (BIC)  7528.739
30 Root Mean Square Error of Approximation:

RMSEA                                0.092
90 Percent Confidence Interval       0.071 0.114
P-value RMSEA <= 0.05                0.001
35 Standardized Root Mean Square Residual:

SRMR                                  0.065
40 Parameter estimates:

Information Standard Errors          Expected Standard
45 Estimate Std.err Z-value P(>|z|)

Latent variables:
visual =~
50 x1          1.000
x2          0.554 0.100 5.554 0.000
x3          0.729 0.109 6.685 0.000
textual =~
55 x4          1.000
x5          1.113 0.065 17.014 0.000
x6          0.926 0.055 16.703 0.000
veloc =~
60 x7          1.000
x8          1.180 0.165 7.152 0.000
x9          1.082 0.151 7.155 0.000

Covariances:
visual ~
65 textual    0.408 0.074 5.552 0.000
veloc       0.262 0.056 4.660 0.000
textual ~
veloc       0.173 0.049 3.518 0.000

Variances:
70 x1          0.549 0.114 4.833 0.000
x2          1.134 0.102 11.146 0.000
x3          0.844 0.091 9.317 0.000
x4          0.371 0.048 7.779 0.000
x5          0.446 0.058 7.642 0.000
x6          0.356 0.043 8.277 0.000
75 x7          0.799 0.081 9.823 0.000

```

```

x8          0.488    0.074    6.573    0.000
x9          0.566    0.071    8.003    0.000
visual     0.809    0.145    5.564    0.000
textual    0.979    0.112    8.737    0.000
veloc      0.384    0.086    4.451    0.000

```

5.2 Model d'equacions estructurals

Per a estimar el model de la secció 4.2, hem d'entrar la matriu de covariàncies:

```

wheaton.cov <- matrix(c(
11.834,    0,    0,    0,    0,    0,
6.947,    9.364,    0,    0,    0,    0,
6.819,    5.091,    12.532,    0,    0,    0,
4.783,    5.028,    7.495,    9.986,    0,    0,
-3.839,   -3.889,   -3.841,   -3.625,    9.610,    0,
-21.899,  -18.831,  -21.748,  -18.775,  35.522,  450.288),
6, 6, byrow=TRUE)
colnames(wheaton.cov) <- rownames(wheaton.cov) <- c("ano67", "pow67", "ano71", "
pow71", "edu", "sei")

```

Llavors estimarem el model a partir de la matriu de covariàncies, i tenint el compte que la mostra compta amb $N = 932$ observacions:

```
fit.wheaton <- sem(wheaton.model, sample.cov=wheaton.cov, sample.nobs=932)
```

Obtindrem els resultats tot fent:

```

#instruccio per a obtenir resultats
summary(fit.wheaton, standardized=TRUE, fit.measures=TRUE)

#resultat obtingut
Lavaan (0.4-9) converged normally after 82 iterations

  Number of observations              932

  Estimator                          ML
  Minimum Function Chi-square         4.735
  Degrees of freedom                   4
  P-value                              0.316

Chi-square test baseline model:

  Minimum Function Chi-square         2133.722
  Degrees of freedom                   15
  P-value                              0.000

Full model versus baseline model:

  Comparative Fit Index (CFI)         1.000
  Tucker-Lewis Index (TLI)           0.999

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)        -15213.274
  Loglikelihood unrestricted model (H1) -15210.906

  Number of free parameters            17
  Akaike (AIC)                         30460.548
  Bayesian (BIC)                       30542.783
  Sample-size adjusted Bayesian (BIC)   30488.792

Root Mean Square Error of Approximation:

  RMSEA                               0.014
  90 Percent Confidence Interval        0.000 0.053
  P-value RMSEA <= 0.05                0.930

```

Standardized Root Mean Square Residual:

SRMR 0.007

Parameter estimates:

Information	Expected					
Standard Errors	Standard					
Estimate	Std.err	Z-value	P(> z)	Std.lv	Std.all	
Latent variables:						
ali67 =~						
ano67	1.000			2.663	0.774	
pow67	0.979	0.062	15.895	0.000	2.606	0.852
ali71 =~						
ano71	1.000			2.850	0.805	
pow71	0.922	0.059	15.498	0.000	2.628	0.832
ses =~						
edu	1.000			2.607	0.842	
sei	5.219	0.422	12.364	0.000	13.609	0.642
Regressions:						
ali71 ~						
ali67	0.607	0.051	11.898	0.000	0.567	0.567
ses	-0.227	0.052	-4.334	0.000	-0.207	-0.207
ali67 ~						
ses	-0.575	0.056	-10.195	0.000	-0.563	-0.563
Covariances:						
ano67 ~						
ano71	1.623	0.314	5.176	0.000	1.623	0.356
pow67 ~						
pow71	0.339	0.261	1.298	0.194	0.339	0.121
Variances:						
ano67	4.731	0.453	10.441	0.000	4.731	0.400
pow67	2.563	0.403	6.359	0.000	2.563	0.274
ano71	4.399	0.515	8.542	0.000	4.399	0.351
pow71	3.070	0.434	7.070	0.000	3.070	0.308
edu	2.801	0.507	5.525	0.000	2.801	0.292
sei	264.597	18.126	14.597	0.000	264.597	0.588
ali67	4.841	0.467	10.359	0.000	0.683	0.683
ali71	4.083	0.404	10.104	0.000	0.503	0.503
ses	6.798	0.649	10.475	0.000	1.000	1.000

5.3 Funcions d'extracció dels estimadors

Amb la funció `summary` podem tenir informació global sobre la qualitat de l'ajust del model i els valors dels paràmetres. Si volem obtenir un *data frame* amb els estimadors dels paràmetres, podem fer:

```
parameterEstimates(fit.model1)
```

També podem obtenir els estimadors amb els valors estandarditzats fent:

```
standardizedSolution(fit.model1)
```

A tall d'exemple, si fem `parameterEstimates(fit.model1)` obtenim els valors estimats dels paràmetres de l'anàlisi factorial confirmatòria definit a la secció 4.1:

```
> parameterEstimates(fit.model1)
  lhs op  rhs  est  se      z est.std est.std.all
1 visual =~  x1 1.000 0.000  NA  0.900  0.772
2 visual =~  x2 0.554 0.100 5.554 0.498  0.424
3 visual =~  x3 0.729 0.109 6.685 0.656  0.581
4 textual =~ x4 1.000 0.000  NA  0.990  0.852
5 textual =~ x5 1.113 0.065 17.014 1.102  0.855
6 textual =~ x6 0.926 0.055 16.703 0.917  0.838
```



```

7   veloc == x7 1.000 0.000 NA 0.619 0.570
10  8   veloc == x8 1.180 0.165 7.152 0.731 0.723
9   veloc == x9 1.082 0.151 7.155 0.670 0.665
10  10  x1 == x1 0.549 0.114 4.833 0.549 0.404
11  11  x2 == x2 1.134 0.102 11.146 1.134 0.821
12  12  x3 == x3 0.844 0.091 9.317 0.844 0.662
15  13  x4 == x4 0.371 0.048 7.779 0.371 0.275
14  14  x5 == x5 0.446 0.058 7.642 0.446 0.269
15  15  x6 == x6 0.356 0.043 8.277 0.356 0.298
16  16  x7 == x7 0.799 0.081 9.823 0.799 0.676
17  17  x8 == x8 0.488 0.074 6.573 0.488 0.477
20  18  x9 == x9 0.566 0.071 8.003 0.566 0.558
19  19  visual == visual 0.809 0.145 5.564 1.000 1.000
20  20  textual == textual 0.979 0.112 8.737 1.000 1.000
21  21  veloc == veloc 0.384 0.086 4.451 1.000 1.000
22  22  visual == textual 0.408 0.074 5.552 0.459 0.459
25  23  visual == veloc 0.262 0.056 4.660 0.471 0.471
24  24  textual == veloc 0.173 0.049 3.518 0.283 0.283

```

Per a obtenir un llistat dels indicadors de qualitat d'estimació d'un model podem fer:

```

#llistat general
fitMeasures(fit.model1)

#per a un indicador concret
5 fitMeasures(fit.model1, "rmsea")

```

El resultat d'aplicar `fitMeasures(fit.model1)` per al model definit a 4.1:

```

> fitMeasures(fit.model1)
      chisq          df          pvalue
      85.306         24.000           0.000
  baseline.chisq  baseline.df  baseline.pvalue
      918.852         36.000           0.000
      cfi          tli          logl
      0.931          0.896      -3737.745
unrestricted.logl          npar          aic
      -3695.092         21.000       7517.490
      bic          ntotal          bic2
      7595.339         301.000       7528.739
      rmsea  rmsea.ci.lower  rmsea.ci.upper
      0.092          0.071          0.114
      rmsea.pvalue          srmr
      0.001          0.065

```

`lavaan` dona per defecte una gamma relativament determinada d'estimadors. Tanmateix, proporciona els paràmetres necessaris per calcular-ne d'altres que no hi figuren, com ara el *Goodness of Fit Index* (GFI).

6 Índexs de modificació

Com d'altres programes de SEM, `lavaan` proporciona els *índexs de modificació* dels paràmetres fixes. L'índex de modificació ens indica com disminueix la χ^2 (és a dir, com millora l'ajust del model) si deixem lliure un paràmetre fixe. Per exemple, podem avaluar els índexs de modificació de l'anàlisi factorial confirmatòria fent:

```
mi.model1 <- modindices(fit.model1)
```

Si, per exemple, volem obtenir els índexs per les correlacions entre errors i variables latents, podem fer:

```
mi.model1[mi.model1$op == "~~", ]
```

El resultat de fer aquesta operació es mostra a continuació (s'han eliminat algunes files):

	lhs	op	rhs	mi	epc	sepc.lv	sepc.all
1	x1	~	x1	0.000	0.000	0.000	0.000
2	x1	~	x2	3.606	-0.184	-0.184	-0.134
5	...						
8	x1	~	x8	0.634	-0.041	-0.041	-0.035
9	x1	~	x9	7.335	0.138	0.138	0.117
10	x2	~	x2	0.000	0.000	0.000	0.000
11	x2	~	x3	8.532	0.218	0.218	0.164
	...						
27	x4	~	x6	6.221	-0.235	-0.235	-0.185
28	x4	~	x7	5.920	0.098	0.098	0.078
29	x4	~	x8	3.805	-0.069	-0.069	-0.059
	...						
40	x7	~	x7	0.000	0.000	0.000	0.000
41	x7	~	x8	34.145	0.536	0.536	0.488
42	x7	~	x9	5.183	-0.187	-0.187	-0.170
43	x8	~	x8	0.000	0.000	0.000	0.000
44	x8	~	x9	14.946	-0.423	-0.423	-0.415
45	x9	~	x9	0.000	0.000	0.000	0.000
46	visual	~	visual	0.000	0.000	0.000	0.000
	...						

Dels resultats obtinguts podem observar que es pot millorar força l'ajust del model si deixem lliures algunes correlacions entre variables del mateix factor:

- Per a la variable latent v , la correlació entre x_1 i x_2
- Per a la variable l , la correlació entre x_4 i x_6
- Per a la variable t , les correlacions entre x_7 i x_8 , i entre x_8 i x_9

Podem incorporar al model aquestes modificacions:

```

modell1.mod <- 'visual =~ x1 + x2 + x3
textual =~ x4 + x5 + x6
veloc =~ x7 + x8 + x9

x2 =~ x3
x4 =~ x6
x7 =~ x8
x8 =~ x9'

```

Els paràmetres d'ajustament d'aquest model seran:

```

> fitMeasures(fit.modell1.mod)
      chisq          df          pvalue
      35.490         20.000          0.018
baseline.chisq  baseline.df  baseline.pvalue
      918.852         36.000          0.000
      cfi          tli          logl
      0.982          0.968        -3712.837
unrestricted.logl  npar          aic
      -3695.092        25.000        7475.674
      bic          ntotal          bic2
      7568.352        301.000        7489.066
      rmsea          rmsea.ci.lower  rmsea.ci.upper
      0.051          0.021          0.077
      rmsea.pvalue          srmsr
      0.448          0.047

```

Si comparem els resultats d'aquest model amb l'original, veiem que la χ^2 passa de 85.306 a 35.490, de manera que la resta de paràmetres d'ajustament també milloren substancialment.

7 Conclusions

En els últims anys, R s'està convertint en una alternativa vàlida als paquets estadístics usats en diverses àrees científiques, entre elles les ciències socials [2]. R permet aplicar la majoria de tècniques estadístiques usant un únic programa informàtic, i en ser programari en codi obert permet que els investigadors puguin accedir i modificar el codi, i beneficiar-se de les aportacions de la comunitat. El desenvolupament de `lavaan` és un exemple de com poden resoldre dos dels majors problemes del programari en codi obert: la mala documentació i els costos d'aprenentatge. Tot i que està en les fases inicials del seu desenvolupament, `lavaan` compta ja amb una bona documentació [8, 7], i la forma d'especificar el model redueix, al nostre entendre, els costos d'aprenentatge per a aquells usuaris familiaritzats amb els models d'equacions estructurals. Cal esperar que, en el futur pròxim, disposem d'un paquet complet i estable en codi obert per a l'aplicació d'una de les tècniques més útils en ciències socials.

Referències

- [1] Bates T, Mehta P, Fox J (2011) OpenMx: an Open Source Extended Structural Equation Modeling Framework. *Psychometrika*, Vol. 76, pp. 306–317.
- [2] Culpepper S A, Aguinis H, (2010) R is for revolution: a cutting-edge, free, open source statistical package. *Organizational research methods*, Vol. 14, No. 4, pp. 735-740.
- [3] Fox J (2006) Structural Equation Modeling with the sem Package in R. *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 13, pp. 465–486.
- [4] Fox J (2010) sem: Structural Equation Models. R package version 0.9-21, disponible en <http://CRAN.R-project.org/package=sem>
- [5] Neale M C, Boker S M, Xie G, Maes H H (2003) Mx: Statistical Modeling. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry. 6th Edition.
- [6] Rosseel Y (2012a) lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, vol 48, No. 2.
- [7] Rosseel Y (2012b) Package 'lavaan'. Disponible en: <http://cran.r-project.org/web/packages/lavaan/lavaan.pdf>
- [8] Rosseel Y (2011) lavaan: an R package for structural equation modelling and more. Version 0.4-9 (BETA). Disponible en <http://users.ugent.be/yrosseel/lavaan/lavaanIntroduction.pdf>
- [9] R Development Core Team (2011) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, disponible en <http://www.R-project.org>.
- [10] Schumacker R E, Lomax R G (2004) A beginner's guide to structural equation modeling. Lawrence Erlbaum
- [11] Wheaton B, Muthén B, Alwin D F, Summers G F (1977) Assessing reliability and stability in panel models. In D. R. Heise (Ed) *Sociological methodology 1977* (pp. 84–136). San Francisco: Jossey-Bass.



This work is licensed under *Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Licence*.