Guía Paso a Paso para construir un modelo PLSPM con fStats

Utilizaremos un ejemplo de la literatura (Gefen y Straub, 2005) para explicar, paso a paso, cómo emplear *f*Stats para estimar un modelo PLSPM. El diagrama del modelo es el que se muestra en la Figura 1.

Gefen, D., and Straub, D. (2005). A Practical Guide to Factorial Validity Using PLS-Graph: Tutorial and Annotated Example, *Communications of the AIS* (16), pp. 91-109.



Figura 1. Diagrama del modelo que queremos estimar.

1. Preparación de los datos

Los datos los tendremos preparados en una hoja de cálculo (en nuestro caso supondremos que dicha hoja se llama *Data*), empleando una fila para cada individuo y una columna para cada indicador.

Dispondremos de una fila para el nombre de cada indicador y una o más filas para especificar a qué constructo queremos asignar cada indicador. En nuestro ejemplo, según se muestra en la Figura 2, tenemos 100 individuos y 10 indicadores. Los cuatro primeros indicadores (eo1, ..., eo4) corresponden al constructo 1 (según se muestra en la Fila 2 de la Figura 2), los cuatro siguientes se asignan al constructo 2 y los dos últimos al constructo 3.

	Α	В	C	D	E	F	G	н	I	J	К	L	М	Ν	0	P C	Q R	S
1												_						
2		1	1	1	1	2	2	2	2	3	3	_	(Constructs Names			Lin	ks
3	Id	eou1	eou2	eou3	eou4	pu1	pu2	pu3	pu4	use1	use2		I	Long Name	Short Name	Туре	From	То
4	1	3	2	2	2	3	4	3	4	5	5		1 I	Facilidad de Uso	EOU	1	1	2
5	2	1	2	2	2	2	2	2	5	6	6		2 t	Jtilidad Percibida	PU	1	1	3
6	3	1	1	1	1	2	4	2	2	1	1	_	3 I	ntención de Uso	USE	1	2	3
7	4	3	4	3	3	3	4	3	3	3	3							
8	5	2	3	2	2	2	2	2	2	2	3							
9	6	2	2	2	2	2	2	2	2	4	4							
10	7	2	1	2	2	1	1	1	1	2	2							
11	8	2	2	2	2	2	3	3	3	4	4							

Figura 2. Organización sugerida de los datos, incluyendo los nombres y la filiación de los indicadores.

2. Llamar al diálogo PLS Path Modeling

Estando en la hoja en la que tenemos los datos, pulsamos la combinación **CTRL+SHIFT+A**, con lo que aparecerá el diálogo **fStats Applications**, que se muestra en la Figura 3. En este diálogo pulsamos el el botón **PLS Path Modeling** para llamar al diálogo correspondiente, que se muestra en la Figura 4.

Elementos del diálogo PLS Path Modeling

Name of the model: si lo desea puede introducir un nombre para identificar el modelo a estimar.

Input type: puede estimar un modelo PLSPM a partir de los datos (Raw Data) o a partir de la matriz de covarianzas de los indicadores (Covariance Matrix). Si elige hacerlo a partir de la matriz de covarianzas tiene que especificar un número de observaciones y la aplicación generará automáticamente un conjunto de datos del tamaño elegido, siguiendo el modelo normal multivariante, con la estructura de correlaciones especificada en la matriz de covarianzas.

Data Origin: seleccione con el ratón (o indique manualmente) el rango de Excel en el que está la matriz de datos (una fila por cada observación y una columna por cada indicador).

Name of the items: seleccione con el ratón, o indique manualmente, el rango fila en el que están escritos los nombres de los indicadores.



Figura 3. Dialogo fStats Applications

Items Filiation: seleccione con el ratón, o indique manualmente, un rango con una o más filas y con tantas columnas como indicadores hay en la matriz de datos. Se permite emplear varias filas para especificar modelos con constructos de orden superior. En la posición correspondiente a cada indicador se pondrá el número que identifica a su constructo (se puede dejar en blanco o poner un cero cuando queremos excluir del modelo algún indicador).

Name of the constructs: seleccione con el ratón, o indique manualmente, un rango con dos columnas y una fila por cada constructo a incluir en el modelo. En la primera columna a puede indicar el nombre completo de cada constructo y en la segunda un nombre abreviado para ser empleado en las tablas.

fStats: PLS Path Modeling	X
Name of the model:	
Input Type: N: Raw Data	Data <u>O</u> rigin:
Name of the items:	Items <u>Fi</u> liation:
Name of the <u>c</u> onstructs:	Type of constructs:
Structural <u>m</u> odel:	Estimation Scheme:
Bootstrap Samples: 0	
Specify a Filter Save i	n a Sheet
Close Sho	w <u>H</u> elp <u>E</u> stimate

Figura 4. Diálogo PLS Path Modeling

fStats: PLS Path Modeling	×
Name of the model: Ejemplo Ge	fen & Straub 2005
Input Type: N: Da	ata <u>O</u> rigin:
Raw Data 🔹	Data!\$B\$4:\$K\$103
Name of the items: Ite	ems <u>F</u> iliation:
Data!\$B\$3:\$K\$3	Data!\$B\$2:\$K\$2
Name of the <u>c</u> onstructs: <u>Ty</u>	pe of constructs:
Data!\$N\$4:\$O\$6	Data!\$P\$4:\$P\$6
Structural <u>m</u> odel: Es	timation Scheme:
Data!\$R\$4:\$S\$6	Path Weighting
Bootstrap Samples: 1000	
Specify a Filter Save in a S	heet
Close Show He	elp <u>E</u> stimate

Figura 5. Diálogo PLS Path Modeling cumplimentado con la información del modelo a estimar.

Type of constructs: seleccione con el ratón, o indique manualmente, un rango columna con tantas filas como constructos en el modelo, indicando, para cada constructo, si este es reflectivo, poniendo 1 en la celda correspondiente, o formativo, poniendo 2 en la celda correspondiente.

Structural model: seleccione con el ratón, o indique manualmente, un rango con dos columnas y una fila por cada enlace del modelo estructural. Para cada enlace indique en la primera columna el número que identifica al constructo origen y en la segunda columna el número que identifica al constructo destino.

Estimation Scheme: elija en la lista desplegable el esquema para la estimación interna del modelo. Las opciones son: **Centroid**, **Factorial**, **Path Weighting** (por defecto) y **Simple PCA**.

Bootstrap Samples: indique el número de muestras bootstrap con reemplazamiento para realizar inferencias con los diferentes estadísticos resultantes de la estimación del modelo.

En la Figura 5 se muestra el diálogo **PLS Path Modeling** cumplimentado con la especificación del modelo de la Figura 1, siguiendo la nomenclatura de celdas que se muestra en la Figura 2.

Ya tenemos especificado el modelo y podemos estimarlo, pulsando el botón **Estimate**, o podemos guardar la especificación en una hoja de Excel, pulsando el botón **Save in a Sheet**.

Si elige guardar la especificación del modelo, la aplicación crea una nueva hoja de cálculo con todos los elementos que definen el modelo, incluyendo los datos, según se muestra en la Figura 6.



Figura 6. Especificación del modelo guardada en una nueva hoja de Excel, obtenida al pulsar el botón Save in a Sheet del diálogo PLS Path Modeling, cumplimentado según se muestra en la Figura 5.

El modelo puede estimarse desde el diálogo **PLS Path Modeling**, pulsando el botón **Estimate**, o puede estimarlo desde la hoja de Excel en que se ha guardado su especificación, pulsando la combinación: **CTRL+SHIFT+R**, en cualquiera de los dos casos se obtiene una nueva hoja de cálculo con los resultados de la estimación, según se muestra en la Figura 7. En algunos casos la estimación puede demorarse más de un minuto, por lo que se recomienda tener paciencia.

Es interesante notar que, al tener la especificación del modelo guardada en una hoja de Excel, puede hacer cambios en la misma y volver a estimar el modelo, sin necesidad de volver a invocar el diálogo **PLS Path Modeling**, simplemente pulsando la combinación de teclas **CTRL+SHIFT+R**.

	A	В	С	D	E	F	G	Н	I	J	К	L	
1	8 marzo 201	6					Fast jump to						
2	Model:	Ejemplo Gefe	n & Straub 20	Observations	100		Stru	ctural Paths:	\$P\$9		Bootstrap:	\$Q\$19	
3				Items:	10		Constructs	correlations:	\$W\$9	Items	correlations:	\$Y\$20	
4	Iterations:	4		Constructs:	3		Partial	correlations:	\$AB\$9		Warnings:	\$A\$36	
5	Warnings:	0		Links:	3		eigenvalue	s and alphas:	\$A\$20		Scores:	\$G\$35	
6	Filter:	None		Scheme:	Structural		Loadings	and weights:	\$G\$20	Indiv. Part	correlations:	\$G\$140	
7				Bootstrap:	1000		Select the destination an push CTRL + SHIFT + G to Go						

Figura 7. Parte superior de la hoja de Excel en la que se muestra el modelo estimado.

3. Análisis de la salida que se obtiene al estimar el modelo

dirigirnos

Ilustraremos las salidas que se obtienen con el ejemplo visto en los apartados anteriores.

En la parte superior derecha de la hoja de resultados tenemos un resumen de las características del modelo estimado. En la Figura 8 se muestra dicho resumen para nuestro ejemplo.

para

A la derecha del resumen tenemos un listado de los apartados que componen la salida de la estimación. Este listado

rápidamente a cada uno de los apartados referidos. En la Figura 9 se muestra el listado que obtenemos en

también

nuestro ejemplo.

sirve

	A	В	С	D	E
1	8 marzo 201	6			
2	Model:	Ejemplo Gefe	n & Straub 20	Observations	100
3				Items:	10
4	Iterations:	4		Constructs:	3
5	Warnings:	0		Links:	3
6	Filter:	None		Scheme:	Structural
7				Bootstrap:	1000

Figura 8. Resumen de las características del modelo estimado.

	G	Н	I	J	К	L					
1	Fast jump to										
2	Stru	uctural Paths:	\$P\$9		Bootstrap:	\$Q\$19					
3	Constructs correlations: \$W\$9 Items correlations: \$Y\$20										
4	Partial correlations: \$AB\$9 Warnings: \$A\$3										
5	eigenvalue	es and alphas:	\$A\$20		Scores:	\$G\$35					
6	Loadings	and weights:	\$G\$20	Indiv. Part.	correlations:	\$G\$140					
7		Select the dea	stination an p	ush CTRL + SH	HFT + G to Go						
F	ioura 9 List	tado de los a	inartados di	le compone	n la salida du	el modelo					

Figura 9. Listado de los apartados que componen la salida del modelo estimado.

Cada apartado viene acompañado de una referencia a su ubicación en la hoja de cálculo, de manera que si nos colocamos sobre ella, y pulsamos **CTRL+SHIFT+G**, llegamos directamente al apartado deseado. Por ejemplo, si nos colocamos en la celda I6, en la que tenemos la referencia a la ubicación de

los Loadings y los Weights, y pulsamos la combinación **CTRL+SHIFT+G**, la aplicación nos lleva directamente a la posición dónde se relacionan dichos resultados, según se muestra en la Figura 10.

	G	Н	Ι	J	К	L	М	Ν	0
20	Item t	o scale co	orrelations	s (loading	gs and c	ross loadi	ings)		
21		EOU	PU	USE		Loadings	Weights	VIF	$R_{i \cdot \text{Other}}^2$
22	eou1	0,809	0,387	0,240	eou1	0,809	0,292	1,870	46,52%
23	eou2	0,817	0,402	0,237	eou2	0,817	0,299	1,859	46,22%
24	eou3	0,868	0,362	0,199	eou3	0,868	0,265	2,488	59,80%
25	eou4	0,846	0,490	0,211	eou4	0,846	0,341	1,925	48,06%
26	pu1	0,371	0,814	0,330	pu1	0,814	0,285	1,927	48,12%
27	pu2	0,297	0,760	0,178	pu2	0,760	0,196	1,789	44,09%
28	pu3	0,447	0,878	0,394	pu3	0,878	0,342	2,277	56,07%
29	pu4	0,488	0,864	0,418	pu4	0,864	0,369	2,048	51,17%
30	use1	0,181	0,370	0,895	use1	0,895	0,528	1,673	40,23%
31	use2	0,295	0,386	0,913	use2	0,913	0,578	1,673	40,23%

Figura 10. Posición de los loadings, los cross-loadings, los weights

Para volver al principio de la hoja de cálculo basta con pulsar CTRL+Inicio.

4. Detalle de los resultados obtenidos al estimar nuestro modelo

Fiabilidad de los constructos y relación entre los mismos

En esta tabla, que se muestra en la Figura 11, se lista, para cada constructo, el alpha de Cronbach, la fiabilidad compuesta (CR), el AVE, la raíz cuadrada del AVE (los valores en negrita en la diagonal). También se muestra, para cada par de constructos, la correlación simple entre ellos (debajo de la diagonal) y el Heterotrait-Monotrait ratio (sobre la diagonal).

	В	С	D	E	F	G	Н	L I
10		alpha	CR	AVE		EOU	PU	USE
11	Facilidad Percibida de Uso	0,856	0,902	0,698	EOU	0,835	0,557	0,322
12	Utilidad Percibida	0,853	0,899	0,690	PU	0,497	0,830	0,487
13	Intención de Compra Online	0,776	0,899	0,817	USE	0,266	0,418	0,904

Figura 1	1. Alpha de	Cronbach,	Fiabilidad	compuesta,	AVE,	correlaciones	/ HT	/MT
				· · · · · · · · · · · · · · · · · · ·	,		/	/

En la Figura 12 se muestra la parte derecha de la tabla anterior. En esta parte se muestra, para cada constructo, el porcentaje de varianza del mismo que se podría explicar mediante una regresión a partir de los demás constructos. Para los constructos endógenos se muestra, además, el coeficiente de determinación lineal R^2 (junto a su intervalo de confianza construido mediante Bootstrap) y la versión predictiva de dicho coeficiente: el coeficiente Q^2 .

	F	G	Н	-	J	К	L	М	Ν
9								<i>R</i> ² 95% CI	
10		EOU	PU	USE	$R_{i \cdot \text{Other}}^{2}$	R^2	Q^2	Lower	Upper
11	EOU	0,835	0,557	0,322	25,09%				
12	PU	0,497	0,830	0,487	33,49%	24,68%	20,14%	4,87%	43,17%
13	USE	0,266	0,418	0,904	17,93%	17,93%	12,07%	1,50%	30,58%
14	Bold figures i	n diagonal are	e the square ro	oots of the AV	E				
15	C:1				1	_			

15 Simple correlations between pairs of constructs in the lower triangle

16 Heterotrait-Multitrait (HT/MT) ratios in the upper triangle

Figura 12. Coeficientes de determinación lineal: $R_{i\cdot 0ther}^2$, R^2 y Q^2

Coeficientes del modelo estructural

En la Figura 13 se muestra la estimación de los coeficientes del modelo (los *betas*). Si hemos especificado que deseamos emplear Bootstrap, al rellenar la plantilla del modelo, dispondremos del *p-value* y el intervalo de confianza de cada

	Р	Q	R	S	Т	U
9					Coeff 95% CI	
10	From	То	Coefficient	p-value	Lower	Upper
11	EOU	PU	0,497	0,000	0,280	0,714
12	EOU	USE	0,077	0,505	-0,129	0,283
13	PU	USE	0,380	0,001	0,162	0,598

Figura 13. Tabla de coeficientes del modelo estructural

coeficiente estimado, para estudiar su significatividad.

Correlaciones simples y parciales entre pares de constructos

	W	Х	Y	Z	AA	AB	AC	AD	AE
9	Correlation n	natrix for the o	constructs			Partial Correl	lation matrix f	or the constru	icts
10	R	EOU	PU	USE		Rp	EOU	PU	USE
11	EOU	1,000	0,497	0,266		EOU	1,000	0,440	0,074
12	PU	0,497	1,000	0,418		PU	0,440	1,000	0,342
13	USE	0,266	0,418	1,000		USE	0,074	0,342	1,000

En la Figura 14 se muestran las matrices de correlaciones simples y parciales para los constructos.

Figura 14. Matrices de correlaciones (simples y parciales) para los constructos

Estudio de la unidimensionalidad latente de los constructos

En la Figura 15 se muestra una tabla con los valores propios de la matriz de correlaciones de los indicadores de cada constructo, el alpha de Cronbach, y el que se obtendría en cada caso si eliminamos cada uno de los indicadores.

Este estudio sólo es pertinente para los constructos reflectivos.

	A	В	С	D	E
21	Construct	Eigenvalues	alpha	alpha i.e.	Item
22	EOU	2,795	0,856	0,831	eou1
23		0,490		0,828	eou2
24		0,416		0,785	eou3
25		0,300		0,820	eou4
26	PU	2,775	0,853	0,816	pu1
27		0,500		0,827	pu2
28		0,424		0,795	pu3
29		0,301		0,812	pu4
30	USE	1,634	0,776	1,000	use1
31		0,366		1,000	use2

Figura 15. Unidimensionalidad latente de los constructos.

Correlación de los indicadores con su constructo (loadings), con otros constructos (crossloadings) y pesos

En la Figura 16 se muestran dos tablas. En la primera se muestra, para cada indicador (en filas) su correlación con cada constructo (en columnas). Se resaltan, en negrita, las correlaciones de cada indicador con su propio constructo (loadings). El resto son las correlaciones de cada indicador con cada uno de los otros constructos (cross-loadings). En la segunda tabla se vuelve a mostrar los loadings y se muestran los weights.

	G	Н	I	J	Κ	L	М			
20	Item to scale	Item to scale correlations (loadings and cross loadings)								
21		EOU	PU	USE		Loadings	Weights			
22	eou1	0,809	0,387	0,240		0,809	0,292			
23	eou2	0,817	0,402	0,237		0,817	0,299			
24	eou3	0,868	0,362	0,199		0,868	0,265			
25	eou4	0,846	0,490	0,211		0,846	0,341			
26	pu1	0,371	0,814	0,330		0,814	0,285			
27	pu2	0,297	0,760	0,178		0,760	0,196			
28	pu3	0,447	0,878	0,394		0,878	0,342			
29	pu4	0,488	0,864	0,418		0,864	0,369			
30	use1	0,181	0,370	0,895		0,895	0,528			
31	use2	0,295	0,386	0,913		0,913	0,578			
							-			

Figura 16. Loadings, cross-loadings y weights, para cada indicador.

Estudio de la posible multicolinealidad en constructos formativos

En los constructos formativos interesa comprobar el grado de multicolinealidad entre sus indicadores, lo cual se puede hacer, para cada indicador, mediante el Factor de Inflación de la Varianza (VIF) o con el coeficiente de determinación lineal de la regresión que se construye para tratar de predecir cada indicador a partir del resto de los indicadores de su mismo constructo ($R_{i\cdot Other}^2$). Estos valores se muestran en la Figura 17. El *VIF* y el $R_{i\cdot Other}^2$ están relacionados entre sí, ya que:

$$R_{i \cdot Other}^2 = \frac{VIF - 1}{VIF}$$

	K	L	М	Ν	0
21		Loadings	Weights	VIF	$R_{i \cdot 0 \text{ther}}^2$
22	eou1	0,809	0,292	1,870	46,52%
23	eou2	0,817	0,299	1,859	46,22%
24	eou3	0,868	0,265	2,488	59,80%
25	eou4	0,846	0,341	1,925	48,06%
26	pu1	0,814	0,285	1,927	48,12%
27	pu2	0,760	0,196	1,789	44,09%
28	pu3	0,878	0,342	2,277	56,07%
29	pu4	0,864	0,369	2,048	51,17%
30	use1	0,895	0,528	1,673	40,23%
31	use2	0,913	0,578	1,673	40,23%

Figura 17. VIF y $R_{i.other}^2$ para cada indicador

Variables latentes estimadas para cada observación

En la Figura 18 vemos, para cada observación, el valor estimado de la variable latente (constructo).

	G	Н	-	J
35	Construct val	ues for each ii	ndividual	
36		EOU	PU	USE
37	1	0,128	0,977	1,073
38	2	-0,495	0,460	1,715
39	3	-1,314	-0,110	-1,494
40	4	1,214	0,677	-0,210
41	5	0,083	-0,439	-0,508
42	6	-0,184	-0,439	0,431
43	7	-0,451	-1,391	-0,852
44				
45	:	:	:	:
46				

Figura 18. Variables latentes estimadas para cada individuo

Resultados del análisis Bootstrap

En la Figura 19 se		0	Р	Q	R	S	Т	U	V
nuede ver nara cada	19	M = 1000 boo	tstrap combin	nations of the	coefficients a	nd the R2	for the endog	enous constru	cts
pucue ver, para caua	20	From:	EOU	EOU	PU				
muestra bootstrap, la	21	To:	PU	USE	USE	_	PU	USE	
estimación obtenida	22		0,404	-0,239	0,298	-	16,32%	8,84%	
nara los coeficientes	23		0,501	0,102	0,266		25,06%	10,86%	
para los coenciences	24		0,459	0,017	0,434		21,09%	19,57%	
del modelo estructural	25		0,526	0,220	0,338		27,62%	24,06%	
v el coeficiente de	26		0,463	0,053	0,367		21,40%	15,55%	
determinación lineal	27								
	28		:	:	:		:	:	
<i>R</i> ² .	29]							

Figura 19. Coeficientes estimados y R^2 de los constructos endógenos para cada muestra bootstap.

Matriz de correlaciones de los indicadores

En la Figura 20 se presenta la matriz de correlaciones de los indicadores.

	W	Х	Y	Z	AA	AB	AC	AD	AE	AF	AG
21		eou1	eou2	eou3	eou4	pu1	pu2	pu3	pu4	use1	use2
22	eou1	1,000	0,516	0,656	0,550	0,315	0,259	0,324	0,369	0,180	0,251
23	eou2	0,516	1,000	0,638	0,579	0,281	0,249	0,367	0,400	0,188	0,237
24	eou3	0,656	0,638	1,000	0,644	0,194	0,201	0,326	0,423	0,105	0,249
25	eou4	0,550	0,579	0,644	1,000	0,421	0,275	0,457	0,433	0,129	0,247
26	pu1	0,315	0,281	0,194	0,421	1,000	0,591	0,623	0,543	0,272	0,323
27	pu2	0,259	0,249	0,201	0,275	0,591	1,000	0,556	0,557	0,185	0,139
28	pu3	0,324	0,367	0,326	0,457	0,623	0,556	1,000	0,677	0,367	0,347
29	pu4	0,369	0,400	0,423	0,433	0,543	0,557	0,677	1,000	0,354	0,401
30	use1	0,180	0,188	0,105	0,129	0,272	0,185	0,367	0,354	1,000	0,634
31	use2	0,251	0,237	0,249	0,247	0,323	0,139	0,347	0,401	0,634	1,000

Figura 20. Matriz de correlaciones de los indicadores.

Correlación simple entre pares de constructos y correlación parcial bloqueando el efecto de cada constructo

En la Figura 21 tenemos, en cada fila, la comparación de la correlación lineal simple entre un par de constructos con la correlación parcial que se obtiene si bloqueamos el efecto de cada uno de los demás constructos. En este ejemplo podemos ver que la correlación simple entre EOU y USE se difumina (se reduce desde 0.266 hasta 0.074) cuando bloqueamos el efecto de PU.

	G	Н	I	J	К	L
139				r _{ij∙k} : partia	l correlation b	locking for
140				eac	h other consti	ruct
141	Xi	Xj	rij	EOU	PU	USE
142	EOU	PU	0,497			0,440
143	EOU	USE	0,266		0,074	
144	PU	USE	0,418	0,342		
	E	<u></u>	(d. l			

Figura 21. Comparación de la correlación simple entre pares de constructos y la correlación parcial cuando bloqueamos el efecto de cada uno de los otros constructos.

Esto nos sugiere que la relación apreciada entre EOU y USE podría deberse, en gran medida, a la relación que ambas variables tienen con PU.

Advertencias

Los modelos de ecuaciones estructurales son muy complejos y su estimación produce salidas muy prolijas. Para ayudar al analista se ha implementado un conjunto de alarmas que centrarán su atención en aspectos que puedan requerirla. La forma de hacerlo es un listado descriptivo de advertencias. Además, los números que presentan un comportamiento "indebido" aparecerán en rojo en la tabla en la que se muestran. En la Figura 22 se muestra la tabla de advertencias que, en este caso, está vacía.



Figura 22. Tabla de advertencias

Los eventos que provocan advertencias son los siguientes:

- Constructos con alpha de Cronbach por debajo del umbral 0.7.
- Indicadores que reducen el alpha de Cronbach de su constructo.
- Constructos cuya matriz de correlaciones tienen más de un eigenvalue mayor que 1.
- Pares de constructos cuya correlación es mayor que la raíz cuadrada del AVE de alguno de los dos constructos implicados.
- Pares de constructos cuyo HT/MT ratio es mayor que 0.85.
- Indicadores con algún cross-loading mayor que el loading.
- Constructos cuyo AVE es menor que 0.5.

5. Cambios en la definición del modelo

Una vez estimado el modelo podemos decidir eliminar algún os indicadores, eliminar o añadir algún enlace entre un par de constructos, eliminar o añadir constructos.

Eliminar algún indicador

Para eliminar algún indicador del modelo basta con quitar su filiación. Para hacerlo podemos poner 0 en el lugar correspondiente o, simplemente dejarlo en blanco. En la Figura 23 vemos lo que habría que hacer para ignorar los indicadores eou2, pu3 y pu4.

Filiation:	1	0	1	1	2	2	0	0	3	3
Id	eou1	eou2	eou3	eou4	pu1	pu2	pu3	pu4	use1	use2
1	3	2	2	2	3	4	3	4	5	5
2	1	2	2	2	2	2	2	5	6	6
3	1	1	1	1	2	4	2	2	1	1

Figura 23. Procedimiento para eliminar indicadores del modelo.

Eliminar una conexión en el modelo

Para eliminar un *path* hay que quitarlo en la plantilla, como se muestra en la Figura 24, y hay que advertirlo en la celda C9, indicando que el rango en el que se especifican los paths del modelo se ha reducido (antes era E18:F20 y ahora es E18:F19). Una vez hechas las modificaciones estimamos el nuevo modelo pulsando la combinación de teclas: **CTRL+SHIFT+R**.



Figura 24. Eliminar un path del modelo

6. Filtrar por alguna variable de segmentación

Podemos decidir estimar el modelo para una submuestra: sólo mujeres, individuos en un segmento de edad, individuos que no pertenezcan a determinado segmento, Imaginemos, por ejemplo, que nuestros datos vienen acompañados de dos variables de segmentación: Sexo (H y M) y Edad (I, II y III). La tabla de datos podría tener ahora el aspecto que se muestra en la Figura 25.

_	Filiation:	1	0	1	1	2	2	0	0	3	3	_	
_	Id	eou1	eou2	eou3	eou4	pu1	pu2	pu3	pu4	use1	use2	Sexo	Edad
	1	3	2	2	2	3	4	3	4	5	5	Н	Ι
	2	1	2	2	2	2	2	2	5	6	6	Н	Ι
	3	1	1	1	1	2	4	2	2	1	1	М	II
	4	3	4	3	3	3	4	3	3	3	3	Н	II
	:	:	:	:	:	:	:	:	:	:	:	:	:

Figura 25. Tabla de datos con dos variables de segmentación (Sexo y Edad).

R

В

Filter Name: Hombres

Filter Name: Hombres

Filter Origin: 'Data'!\$L\$4:\$L\$103

11

11

Para estimar el modelo sólo sobre los Hombres diríamos en la plantilla que incluya sólo los 12 individuos con H en la columna L4:L103, según se muestra:

El mismo	o efe	cto	se	consi	gue	e si	
pedimos	que	se	ex	cluya	а	las	
mujeres:							

Para estimar el modelo para los individuos que en Edad no tengan el valor III, se especifica:

12	Filter Origin:	'Data'!\$L\$4:\$L\$103	Exclude	М
	-			
		1		
	В	C	D	E
11	Filter Name:	Jóvenes	Туре	Value
12	Filtor Origin:	'Data'I\$M\$4.\$M\$103	Evclude	Ш

C

С

D

Туре

Include

D

Туре

F

Value

Н

E

Value

7. Datos Faltantes

Las bases de datos a partir de las cuales queremos estimar modelos PLS tienen, a menudo, datos faltantes, de manera que, antes de estimar el modelo, se recomienda imputar valores en las posiciones vacías. Para hacer esto se recomienda un método que respete la estructura de covarianzas entre las variables, ya que esta determina totalmente la estimación del modelo.

*f*Stats dispone de varias funciones para imputar datos faltantes, pero para facilitar su uso, emplearemos el diálogo Missing Data, que se muestra en la Figura 26, y al que podemos acceder desde el cuadro de aplicaciones al que invocamos con la combinación de teclas CTRL+SHIFT+A.

fStats Missing Data Inputation	fStats Missing Data Inputation								
Select the <u>m</u> atrix with missing values:									
Imputation method	Result place								
C EM Algorithm	• Overlapped								
	○ In a ne <u>w</u> sheet								
Format for inputed positions									
Original Gris Negr	3.1416								
Reference Close	Impute								

Figura 26. Diálogo Missing Data, para imputar datos faltantes.

Para emplear este diálogo tenemos que seguir los siguientes pasos:

- 1. Seleccionar el rango que contiene la matriz de datos.
- 2. Especificar el código que hemos empleado para identificar los daros faltantes. Por defecto se sugiere emplear el código "NaN".
- Elegir el método de imputación entre el algoritmo EM y el método TSR (Trimmed Scores 3. Regression). Si elegimos el método TSR tenemos que especificar el número de componentes (se recomienda emplear tantas componentes como constructos tenemos, o una más). Para más información sobre el método TSR y sobre la implementación del método EM, puede consultarlo en: Abel Folch-Fortuny, Francisco Arteaga, Alberto Ferrer (2015). PCA model building with missing data: new proposals and a comparative study. Chemometrics and Intelligent Laboratory Systems 146, pp. 77-88.

- **4**. Elegir dónde queremos colocar la matriz con los datos faltantes imputados. La opción por defecto es Overlapped, es decir, las imputaciones sustituirán a los datos faltantes en la misma hoja en la que están. La alternativa es presentar la nueva matriz completa en una nueva hoja de cálculo.
- 5. Pulsamos el botón Impute para imputar los datos faltantes.

8. Rotaciones Varimax y Procrustes confStats

Estimar un modelo PCA y rotar los loadings obtenidos, para ver si podemos agrupar los indicadores según las hipótesis que hemos construido, constituyen una alternativa al CFA (Análisis Factorial Confirmatorio). La aplicación *f*Stats nos permite hacerlo a partir de los datos o a partir de la matriz de correlaciones de los mismos.

Rotaciones de los loadings a partir de los datos

Recordamos la Figura 2, en la que se presentan los datos del ejemplo y llamamos al diálogo **fStats Applications**, que se mostró en la Figura 3, en el que seleccionamos el botón **Confirmatory PCA with Rotation**. En la figura 27 se muestra dicho diálogo.

VARIMAX and Procrustes Rotations	VARIMAX and Procrustes Rotations $\qquad imes$
Data origin:	Data origin: Data!\$B\$4:\$K\$103
<u>N</u> ame of the items:	Name of the items:
_	Data!\$B\$3:\$K\$3
Input type	Input type
• Ra <u>w</u> data • Correlation matri <u>x</u>	
Name of the model:	Name of the model:
	Ejemplo Rotación desde Raw data
Cancel <u>R</u> un	Cancel <u>R</u> un

Figura 27. Diálogo VARIMAX and Procrustes Rotations.

Para emplear este diálogo tenemos que seguir los siguientes pasos:

- 1. Seleccionar el rango que contiene la matriz de datos.
- 2. Seleccionar el rango que contiene el nombre de las variables.
- 3. Marcar la opción Raw Data.
- 4. Dar un nombre al análisis.

En la parte derecha de la Figura 27 se muestra el diálogo relleno con la información del ejemplo.

Rotaciones de los loadings a partir de la matriz de correlaciones

Podemos construir la matriz de correlaciones de los datos empleando la función **CorrelationMatrix** de *f*Stats. Para ello seleccionaremos un rango de 10 filas y 10 columnas y escribimos la función: **=CorrelationMatrix(B4:K103)** y, como es una función matricial, en lugar de pulsar **Enter** tenemos que pulsar **CTRL + SHIFT + Enter**.

	М	Ν	0	Р	Q	R	S	Т	U	V
4	1	0,516	0,656	0,55	0,315	0,259	0,324	0,369	0,18	0,251
5	0,516	1	0,638	0,579	0,281	0,249	0,367	0,4	0,188	0,237
6	0,656	0,638	1	0,644	0,194	0,201	0,326	0,423	0,105	0,249
7	0,55	0,579	0,644	1	0,421	0,275	0,457	0,433	0,129	0,247
8	0,315	0,281	0,194	0,421	1	0,591	0,623	0,543	0,272	0,323
9	0,259	0,249	0,201	0,275	0,591	1	0,556	0,557	0,185	0,139
10	0,324	0,367	0,326	0,457	0,623	0,556	1	0,677	0,367	0,347
11	0,369	0,4	0,423	0,433	0,543	0,557	0,677	1	0,354	0,401
12	0,18	0,188	0,105	0,129	0,272	0,185	0,367	0,354	1	0,634
13	0,251	0,237	0,249	0,247	0,323	0,139	0,347	0,401	0,634	1
		0.0			1 .		1 1 .	1 1	1	

Figura 28. Matriz de correlaciones para los datos del ejemplo.

Por ejemplo, podemos hacer esto en el rango M4:V13 y obtenemos el resultado que se muestra en la Figura 28.

Ahora llamamos al diálogo Confirmatory PCA with Rotation

Para emplear este diálogo tenemos que seguir los siguientes pasos:

- 1. Seleccionar el rango que contiene la matriz de correlaciones.
- 2. Seleccionar el rango que contiene el nombre de las variables.
- **3**. Marcar la opción **Correlation Matrix**.
- **4**. Dar un nombre al análisis.

En la parte derecha de la Figura 27 se muestra el diálogo relleno con la información correspondiente al ejemplo.

Cuando tenemos relleno el diálogo pulsamos el botón **Run** y la aplicación nos da la opción de elegir el número de componentes a extraer. Por defecto nos ofrece extraer tantas componentes como valores propios mayores que uno tiene la matriz de correlaciones de la matriz de datos, como se ilustra en la Figura 29.

Number of components	×
Select the number of components to extract	Aceptar Cancelar
E	componente

igura 29. Selección del número de componentes principales a extraer.

Detalle de las salidas de la rotación de los loadings con fStats

En la Figura 30 vemos la parte superior de la hoja de cálculo en la que tenemos un índice de contenidos a partir del cual podemos llegar a los diferentes apartados que componen la salida.

	C	D	[[G	Н	I	J	К	L			
1	1 7 febrero 2016				Fast jump to .	Fast jump to							
2	Study	Ejemplo		2]	Eigenvectors: \$C\$19 Procrustes target: \$C\$							
3	n Comp	3		3	Unrota	Unrotated Loadings: \$C\$34 Procrustes rotated:							
4	n Obs (n)	0		4	VARI	MAX rotated:	\$C\$47	Procru	\$C\$76				
5	n Var (p)	10		5		Select the destination an push CTRL + SHIFT + G to Go							

Figura 30. Parte superior de la hoja de resultados de la rotación de los loadings.

La primera salida que tenemos es la matriz de correlaciones de los ítems. Después tenemos los valores propios y los vectores propios correspondientes al número de componentes principales extraídas, según se muestra en la Figura 31.

e-values	l_1	l_2	l 3				
	4,479	1,608	1,231				
e-vectors	\mathbf{v}_1	\mathbf{v}_2	v ₃	loadings	\mathbf{p}_1	\mathbf{p}_2	p ₃
eou1	0,317	0,346	0,089	eou1	0,671	0,439	0,098
eou2	0,321	0,344	0,084	eou2	0,679	0,436	0,093
eou3	0,322	0,456	0,112	eou3	0,681	0,578	0,124
eou4	0,346	0,311	-0,036	eou4	0,732	0,395	-0,040
pu1	0,327	-0,277	-0,291	pu1	0,693	-0,351	-0,323
pu2	0,288	-0,256	-0,465	pu2	0,609	-0,325	-0,516
pu3	0,364	-0,240	-0,190	pu3	0,771	-0,305	-0,211
pu4	0,372	-0,181	-0,118	pu4	0,788	-0,230	-0,131
use1	0,220	-0,386	0,541	use1	0,465	-0,490	0,600
use2	0,253	-0,275	0,572	use2	0,536	-0,349	0,635

Figura 31. Valores y vectores propios

Figura 32. Loadings antes de la rotación

Si examinamos los loadings en la Figura 32 vemos que casi todos los ítems cargan sobre la primera componente. Esto se debe a que las componentes se extraen de manera secuencial y, cuando se extrae cada una de las sucesivas componentes, estas tratan de explicar la mayor cantidad posible de la variabilidad no explicada por las anteriores componentes. Las rotaciones permiten repartir la variabilidad extraída por el número de componentes elegido, entre estas componentes, satisfaciendo algún criterio explicitado por el analista.

El primer criterio elegido es el denominado VARIMAX, en el que se maximiza la suma de las varianzas de los cuadrados de los loadings (el cuadrado de las correlaciones entre variables y componentes), con lo que se consigue una estructura más simple. Idealmente, una estructura se dice simple cuando:

- 1. Cualquier variable dada tiene una alta carga de una sola componente y cargas cercanas a cero en el resto de las componentes.
- 2. Cualquier componente dada está constituida por sólo unas pocas variables con cargas elevadas, mientras que el resto de variables tienen cargas cercanas a cero en esta componente.

La variabilidad total explicada por las componentes extraídas no se modifica, lo que se modifica es su reparto entre las componentes. En la Figura 33 se muestran los loadings después de la rotación VARIMAX. Para facilitar la interpretación se ha difuminado los loadings cuyo valor absoluto está por debajo de un umbral, que puede ser especificado por el usuario (en la Figura 33 el umbral empleado es 0.4). El ejemplo que nos ocupa es un caso ideal, ya que vemos que las 10 variables se agrupan en tres factores claros, que son los que deseamos obtener, con lo que este análisis puede considerarse una alternativa al CFA. En la Figura 34 se muestra la matriz ortonormal que define la rotación que VARIMAX que transforma los loadings de la Figura 32 en los de la Figura 33.

1 III esitotu	0,400				
r-loadings	\mathbf{p}_1	\mathbf{p}_2	\mathbf{p}_3	Comm	Spec Var
eou1	0,167	0,782	0,114	0,652	0,348
eou2	0,178	0,784	0,115	0,660	0,340
eou3	0,092	0,894	0,072	0,812	0,188
eou4	0,310	0,771	0,047	0,693	0,307
pu1	0,810	0,159	0,164	0,708	0,292
pu2	0,856	0,097	-0,034	0,743	0,257
pu3	0,772	0,261	0,260	0,732	0,268
pu4	0,700	0,337	0,294	0,690	0,310
use1	0,186	0,030	0,883	0,816	0,184
use2	0,144	0,186	0,870	0,813	0,187
		Ex	pl Var:	73,19%	

Thursday

0 400

En la Figura 33 también se muestran las comunalidades y las varianzas específicas, para cada ítem, así como el porcentaje de varianza explicado por las tres componentes.

Varima	Varimax Rotation Matrix									
0,65	1 0,665	5 0,366								
-0,48	2 0,734	4 -0,478								
-0,58	7 0,135	5 0,798								

Figura 33. Loadings, después de la rotación VARIMAX.

Figura 34. Matriz ortonormal que define la rotación VARIMAX.

Si modificamos el umbral el formato de las celdas conteniendo los loadings cambia automáticamente.

Después de aplicar la rotación VARIMAX puede suceder que no consigamos la estructura deseada. En ese caso podemos emplear una rotación "guiada", es decir una rotación que tiene una estructura objetivo (denominada Target en la aplicación), de manera que se busca la rotación que, partiendo de los loadings originales, nos lleve a una estructura lo más parecida posible a la especificada en la matriz Target. Esto se muestra en la Figura 35.

Target	\mathbf{p}_1	\mathbf{p}_2	p ₃	Procrustes	\mathbf{p}_1	\mathbf{p}_2	\mathbf{p}_3
eou1	0,000	1,000	0,000	eou1	0,186	0,781	0,089
eou2	0,000	1,000	0,000	eou2	0,196	0,783	0,089
eou3	0,000	1,000	0,000	eou3	0,109	0,893	0,050
eou4	0,000	1,000	0,000	eou4	0,323	0,767	0,013
pu1	1,000	0,000	0,000	pu1	0,821	0,149	0,105
pu2	1,000	0,000	0,000	pu2	0,853	0,083	-0,094
pu3	1,000	0,000	0,000	pu3	0,792	0,253	0,202
pu4	1,000	0,000	0,000	pu4	0,723	0,331	0,240
use1	0,000	0,000	1,000	use1	0,247	0,042	0,868
use2	0,000	0,000	1,000	use2	0,206	0,198	0,855

Figura 35. Matriz Target y Loadings rotados para aproximarse a la misma.

El procedimiento aplicado es la solución al denominado Orthogonal Procrustes Problem, que da lugar a la matriz de rotación **R** que resuelve el siguiente problema de optimización:

$\mathbf{R} = arg min \|\mathbf{L}\mathbf{R} - \mathbf{T}\|_F^2$ **R** orthogonal

En la expresión anterior **L** es la matriz de loadings antes de la rotación, **T** es la matriz Target y el operador $\|\cdot\|_F^2$ es la norma de Fröbenius, por lo que la expresión a optimizar puede escribirse: $tr[(\mathbf{LR} - \mathbf{T})(\mathbf{LR} - \mathbf{T})^T]$.

Tanto en la matriz Target como en la matriz de loadings rotados con la rotación Procrustes se verifica el mismo formato condicionado al umbral que hemos descrito en la matriz de loadings sin rotar de la Figura 33.

Si modificamos libremente los valores de la matriz Target podemos ver como la matriz de loadings rotados con la rotación Procrustes se actualiza instantáneamente, por lo que podemos guiar la rotación de manera interactiva para tratar de conseguir la estructura deseada.

En la Figura 36 se muestran las dos matrices de rotaciones empleadas: la primera es la rotación VARIMAX y la segunda es la rotación Procrustes. Se debe notar que ambas son ortonormales.

Varimax R	otation M	atrix	Procrust	es Rotatio	n Matrix
0,651	0,665	0,366	0,684	0,661	0,309
-0,482	0,734	-0,478	-0,503	0,734	-0,457
-0,587	0,135	0,798	-0,529	0,157	0,834

Figura 36. Matriz de rotación VARIMAX (izquierda) y Procrustes (derecha).

Finalmente, en la Figura 37, se muestra la matriz de los residuos de las correlaciones después de extraer las tres componentes.

Residuals	eou1	eou2	eou3	eou4	pu1	pu2	pu3	pu4	use1	use2
eou1	0,348	-0,140	-0,066	-0,110	0,037	0,044	-0,039	-0,046	0,024	-0,018
eou2	-0,140	0,340	-0,088	-0,086	-0,006	0,025	-0,004	-0,022	0,030	-0,034
eou3	-0,066	-0,088	0,188	-0,077	-0,035	0,039	0,003	0,036	-0,003	0,007
eou4	-0,110	-0,086	-0,077	0,307	0,040	-0,063	0,005	-0,058	0,006	0,018
pu1	0,037	-0,006	-0,035	0,040	0,292	-0,112	-0,086	-0,126	-0,028	0,035
pu2	0,044	0,025	0,039	-0,063	-0,112	0,257	-0,121	-0,065	0,053	0,027
pu3	-0,039	-0,004	0,003	0,005	-0,086	-0,121	0,268	-0,028	-0,014	-0,040
pu4	-0,046	-0,022	0,036	-0,058	-0,126	-0,065	-0,028	0,310	-0,046	-0,019
use1	0,024	0,030	-0,003	0,006	-0,028	0,053	-0,014	-0,046	0,184	-0,167
use2	-0,018	-0,034	0,007	0,018	0,035	0,027	-0,040	-0,019	-0,167	0,187

Figura 37. Matriz de los residuos de las correlaciones.