

Modelo de equilibrio con oferta en red para la economía española

Iván Rendo Barreiro

March 25, 2020

La presentación tiene cinco puntos. El orden de los mismos es en el que los he ido desarrollando¹.

- ➊ Breve explicación teórica de Acemoglu, Akcigit y Kerr (2015).
- ➋ Modelización de la economía española basada en una red sectorial.
- ➌ Explicación Teórica de Baqaee y Farhi (2019) y su relación con AAK (2015).
- ➍ Análisis estadístico-computacional de la red industrial de la economía española.
- ➎ Discrepancias y puntos a tratar a continuación

¹Probablemente el punto 4 debiera ir tras el 1 o el 2.

I. Teoría Acemoglu, Akcigit y Kerr (2015)

Se parte de una economía estática perfectamente competitiva donde hay n industrias. Cada industria $i = 1, \dots, n$ tiene una **función de producción Cobb-Douglas**:

$$y_i = e^{z_i} l_i^{\alpha_i^l} \prod_{j=1}^n x_{ij}^{a_{ij}} \quad (1)$$

donde x_{ij} es la cantidad de bienes producidos por la industria j que utiliza como input i , l_i es el trabajo y z_i un shock de productividad neutral a la Hicks.

Suponemos que $\forall i, \alpha_i^l > 0$ y $a_{ij} \geq 0 \forall j$ y que la función de producción de cada industria presenta RCE, para lo cual es necesario:

$$\alpha_i^l + \sum_{j=1}^n a_{ij} = 1$$

Así, la **condición de vaciado de mercado** es:

$$y_i = c_i + \sum_{j=1}^n x_{ji} + G_i \quad (2)$$

donde c_i es el consumo final del output de la industria i y G_i las compras por parte del gobierno, que se asume que no se devuelven a los hogares. Las **preferencias** de esta economía se resumen en un agente representativo tal que:

$$u(c_1, c_2, \dots, c_n, l) = \gamma(l) \prod_{i=1}^n c_i^{\beta_i} \quad (3)$$

donde $\beta_i \in (0, 1)$ representa la preferencia por cierto bien i donde $\sum_i \beta_i = 1$ y $\gamma(l)$ es una función diferenciable monóticamente decreciente que captura la desutilidad del trabajo.

El gobierno instauro un impuesto *lump-sum* T para financiar sus compras, por lo que $T = \sum_{i=1} p_i G_i$.

La **restricción presupuestaria de los hogares** queda entonces como:

$$\sum_{i=1}^n p_i c_i = wl - T$$

Asumimos que la economía es de equilibrio competitivo por lo que, para precios dados, las condiciones de vaciado de mercado se satisfacen y los agentes maximizan su utilidad. Desde (1) y la maximización de beneficios llegamos a:

$$\frac{p_j x_{ij}}{p_i y_i} = a_{ij}$$

Definimos, así, la matriz **A** con componentes a_{ij} :

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & \\ a_{21} & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{pmatrix}$$

y definimos también la matriz **inversa de Leontief**:

$$\mathbf{H} = \mathbf{Id} + \mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \dots = (\mathbf{Id} - \mathbf{A})^{-1}$$

Bajo estos supuestos, pueden demostrarse una serie de proposiciones que simplemente enuncio:

Theorem

El impacto de un shock de oferta sectorial en el output del sector i es:

$$d \ln y_i = dz_i + \sum_{j=1}^n (h_{ij} - \delta_j) dz_j \quad (6)$$

donde h_{ij} son los coeficientes de la matriz de Leontief. Esta ecuación implica que **en respuesta a shocks de productividad no hay efectos *upstream*.**

Análogamente, si suponemos $\gamma(l) = (1 - l)^\lambda$ podemos demostrar el siguiente teorema:

Theorem

El impacto sobre el output del sector i de un shock de demanda en el gasto del gobierno es:

$$d \ln y_i = \frac{d \tilde{G}_i}{p_i y_i} + \sum_{j=1}^n (\hat{h}_{ji} - \delta_j) \times \frac{1}{p_j y_j} \times \frac{\beta_j}{1 + \lambda} \times \sum_{k=1}^n d \tilde{G}_k \quad (7)$$

donde $\tilde{G}_j = p_j G_j$ es el gasto nominal del gobierno en el sector j , \hat{h}_{ij} es la componente (i, j) de la matriz $\hat{\mathbf{H}} = (\mathbf{Id} - \hat{\mathbf{A}})^{-1}$ y $\hat{\mathbf{A}}$ es la matriz con componentes $\hat{a}_{ij} = \frac{p_j x_{ij}}{p_j y_j}$, i.e., las ventas de la industria j a la industria i normalizadas por las ventas del sector i .

Con esto se propone una ecuación a estimar basada en las ecuaciones anteriores (6) y (7):

Ecuación Empírica

$$\begin{aligned}\nabla \ln Y_{i,t} = & \zeta_t + \psi \nabla \ln Y_{i,t-1} + \beta^{\text{own}} \text{Shock}_{i,t-1} + \\ & + \beta^{\text{upstream}} \text{Ups}_{i,t-1} + \beta^{\text{downstream}} \text{Down}_{i,t-1} + \varepsilon_{i,t}\end{aligned}$$

Con $\text{Ups}_{i,t}$, efectos *upstream* y $\text{Down}_{i,t}$, *downstream*, a definir con diferentes matices para cada caso, pero siempre basado en las matrices IO y δ_t un conjunto de efectos temporales.

II. Construcción del Modelo Empírico para la Economía Española

Construcción del Modelo Empírico (1/13)

El primer paso de la construcción es obtener todas las variables presentes en la **ecuación empírica** para nuestro caso, que será un modelo donde los shocks serán los producidos por variaciones del gasto del gobierno en los diferentes sectores. Así, las variables serán:

- $Y_{i,t}$: Una medida de la producción del sector i en el instante t .
- ζ_t : Conjunto de efectos temporales en el instante t^2 .
- $Shock_{i,t}$: Variable que mide el *shock* estudiado para el propio sector i (efecto directo) en el instante t .
- $Ups_{i,t}$: Variable que mide los efectos indirectos via *upstream* de distintos shocks en el instante t para la industria i .
- $Down_{i,t}$: Variable análoga para efectos via *downstream*.
- $\varepsilon_{i,t}$: error para una industria i en el instante t .

²En esta versión preliminar, $\zeta_t = 0 \quad \forall t$

Nuestros datos, que tomaremos de distintas fuentes, serán referidos a los años desde 1995 a 2017, anualmente, i.e., $t = 1995, 1996, \dots, 2017$, aunque perderemos algunas observaciones debido a las diferencias que habrá que tomar. Las industrias serán las descritas en la contabilidad nacional, aunque es necesaria una *traducción* que se especifica más adelante.

Construcción del Modelo Empírico (3/13)

La variable $Y_{i,t}$, que está presenten en diferencias de logaritmos es, en nuestro caso, el **valor añadido nominal** para cada industria. Está disponible en ineBase a través del siguiente [enlace]. Los primeros datos -una vez tratados *en forma*³- son los siguientes:

Figure: Valor añadido nominal (en miles de euros)

i/t	2017	2016	2015	2014
1	29925	28928	27138	23797
2	995	966	968	956
3	1633	1580	1370	1426
16	1912	1659	1602	1473

³Los .csv originales traen mucho más detalle y agregados, pero a nosotros nos vale con el índice de la industria y los valores para cada año, por lo que hay que eliminar lo restante.

Construcción del Modelo Empírico (4/13)

La **traducción** necesaria que mencionaba se tiene que llevar a cabo porque las tablas IO vienen con una nomenclatura y las series de valor añadido, con otra (NACE/CPA).

Esta equivalencia viene expresada en las propias TIO descargadas del ineBase en este [enlace].

Por ello, cambiamos el índice original de las series de valor añadido por el de las TIO.

Figure: Muestra de la equivalencia de índices

TIO	NACE/CPA (Vadd)
1	1
2	2
...	...
22	131

Construcción del Modelo Empírico (5/13)

En el mismo documento que descargamos en [enlace] están las tablas input-output para 2016. En este .xls hay tres tablas: **IO a precios básicos**, coeficientes técnicos (matriz **A** de la teoría) y los coeficientes de la inversa de Leontief (matriz **H**). Para ser más consistentes con el trabajo de Acemoglu, utilizamos la primera.

Figure: Muestra Tabla IO a Precios Básicos

indiceTIO	1	1	2	...	64	Σ Emp Inter	Σ Empleos
1	1	871	0.1	...	0	31495.9	55782.7
2	2	10.4	796.7	...	0	1275.5	2093.0
...
64	131	0	0	...	0	0	0

Construcción del Modelo Empírico (6/13)

Desde esta tabla, definimos los coeficientes técnicos de la matriz **A** como y calculamos su correspondiente matriz inversa de Leontief **H**.

Definition

$$a_{ij} = \frac{Sales_{j \rightarrow i}^a}{Sales_i}$$

^a $Sales_i$ hace referencia a la columna de suma de empleos total, y no sólo intermedios, pudiendo ser ésta otra opción para calcular los coeficientes técnicos.

De manera análoga definimos la variable de gasto que nos servirá para crear el shock del gasto público.

Definition

$$GovSales\% = \frac{Sales_{i \rightarrow Gov}}{Sales_i}$$

Construcción del Modelo Empírico (7/13)

Definimos ahora, el shock del gasto público, que será una variable **exógena** del modelo de la siguiente manera:

Definition

$$GovShock_{i,t} = GovSales\%_i^{T^*} \cdot \Delta \ln GovSpending_{t-1}^a$$

^aEl año base T^* es 2016, para ser coherente con las relaciones creadas.

- Así, la variable $GovSales\%$ está presente en las propias TIO descargadas anteriormente y hacemos la transformación a porcentaje.
- La variable $GovSpending_t$ se define como el nivel de gasto en consumo final de las AAPP en millones de euros y está disponible en el siguiente [enlace].

Construcción del Modelo Empírico (8/13)

Una vez que tenemos el shock propio computado, debemos conseguir $Upstream_{i,t}$ y $Downstream_{i,t}$. Estos se definen como:

Definitions

$$Upstream_{i,t} = (\sum_j Input\%_{j \rightarrow i}^{T*} - \delta_i) \cdot Shock_{j,t}$$

$$Downstream_{i,t} = (\sum_j Output\%_{i \rightarrow j}^{T*} - \delta_i) \cdot Shock_{j,t}$$

- Donde la variable $Input\%_{j \rightarrow i}^{T*}$ es el coeficiente h_{ij} de la matriz de Leontief que definimos anteriormente.
- La variable $Output\%_{i \rightarrow j}^{T*}$ es el análogo pero con la matriz $\hat{\mathbf{H}}$ descrita en (7), que parte empíricamente de una $\hat{\mathbf{A}}$ donde los coeficientes \hat{a}_{ij} se definen como $\hat{a}_{ij} = \frac{Sales_{i \rightarrow j}}{Sales_i} = a_{ji} \frac{Sales_j}{Sales_i}$

Debido a una mala construcción en el código de las matrices A, la ecuación del modelo ha cambiado. Además, se ha probado a utilizar más retardos de las variables, y estos son los resultados:

Figure: Resultados de la estimación

Variable	Parameter	Std. Err.	T-stat	P-value
$\nabla \ln y_{t-1}$	0.2000	0.0824	2.4277	0.0153
$Upstream_{t-1}$	0.1508	0.0288	5.2369	0.0000
$Ownshock_{t-1}$	0.6014	0.1419	4.2372	0.0000

Así, la ecuación estimada sería finalmente:

Ecuación Estimada

$$\nabla \ln \hat{Y}_{i,t} = \underset{(0.0824)}{0.2} \nabla \ln \hat{Y}_{i,t-1} + \underset{(0.0288)}{0.1508} \hat{U}_{i,t-1} + \underset{(0.1419)}{0.6014} \hat{S}_{i,t-1} + \hat{\varepsilon}_{i,t}$$

Todos los pasos descritos están implementados en el notebook de Python adjunto. Puede simplemente verse el código en formato *.html* en el siguiente [enlace]⁴.

⁴No se puede visualizar directamente, hay que darle a descargar y luego abrirlo (en cualquier navegador, es como una página web).

Finalmente, se ha creado un **shock** basado en las **exportaciones** de cada sector de una manera análoga a la anterior, con la variable dependiente el valor añadido **real** (a partir de la variable nominal utilizada en el anterior shock y el índice de precios de la industria).

Para ello, debemos crear la variable del shock (que llamamos $XShock_{i,t}$):

- Primero, se toma el porcentaje del total de exportaciones que representan las exportaciones de dicho sector desde las tablas IO en $T^*=2016$:

$$Exports_i^{T^*} \% = \frac{Sales_{i \rightarrow X}^{T^*}}{\sum_i Sales_{i \rightarrow X}^{T^*}}$$

- Con esto, directamente se crea la variable:

$$XShock_{i,t} = Exports_i^{T^*} \% \nabla \ln TotExports_t$$

Siendo $TotExports_t$ el total de exportaciones de la economía española en el año t .

Los resultados de la estimación son, de nuevo, que no existen efectos Downstream. Sin embargo, aquí los efectos propios tampoco aparecen significativos (*está en detalle en el drive en ESP_covid*).

Figure: Resultados de la estimación

Variable	Parameter	Std. Err.	T-stat	P-value
$\nabla \ln y_{t-1}$	0.2735	0.0824	3.3649	0.0008
$Upstream_{t-1}$	0.6169	0.1163	5.3025	0.0000
$Upstream_{t-2}$	0.5992	0.1606	3.7310	0.0002

Finalmente, la ecuación del modelo es la siguiente:

$$\begin{aligned}\nabla \ln \hat{Y}_{i,t} = & \underset{(0.0824)}{0.2735} \nabla \ln \hat{Y}_{i,t-1} + \underset{(0.1163)}{0.6169} \hat{U}_{i,t-1} + \\ & + \underset{(0.1606)}{0.5992} \hat{U}_{i,t-2} + \hat{\varepsilon}_{i,t}\end{aligned}$$

Nota: Al incorporar efectos temporales o de entidad en estas estimaciones (tanto para el shock del gasto público como del gobierno), las estimaciones son lo suficientemente *raras* como para desecharlas.

III. Teoría de Baqaee y Farhi (2019) y su relación con AAK (2015)

III. Teoría de Baqaee y Farhi (2019) y su relación con AAK (2015)

III. Teoría de Baqaee y Farhi (2019) (1/...)

Por hacer...

IV. Análisis Estadístico de la Red Sectorial

(Habría que recalcular los resultados con las nuevas matrices A y, probablemente, considerar también el volumen de compras/ventas entre dos sectores -para transmitir al modelo la idea de importancia- y no los coeficientes técnicos)

En esta sección analizaremos la red formada por los diferentes sectores, la exposición que tienen al resto de, desde un enfoque del análisis de redes; i.e., sin partir de efectos infinitos calculados a través de la inversa de Leontief, cuya interpretación sería prácticamente directa. Así, dividiremos la sección en tres puntos:

- 1 Creación del grafo a analizar.
- 2 Medidas de centralidad.
- 3 Detección de comunidades.

Finalmente habrá un pequeño anexo, un cuarto punto, sobre *predicción de links*.

El análisis completo está en un *Notebook* de *Google Collab* al que se puede acceder en el siguiente [enlace] (se abre directamente en Google Collab y puede correrse online sin problema).

Partiendo de la matriz **A** de coeficientes que calculamos anteriormente, creamos un grafo *dirigido*⁵, que es prácticamente *completo*⁶, de la siguiente manera:

Definición del grafo

Sea $G = (N, E)$ el grafo definido, N el conjunto de nodos $N = \{1, 2, \dots, 62\}$ y $E = \emptyset$ el conjunto de enlaces. Definimos un nuevo enlace $e_{ij} \in E$, con $i, j \in N$, peso $w \in (0, 1] \subset \mathbb{R}$ y siendo a_{ij} coeficiente (i, j) de **A**:

$$e_{ij} : i \xrightarrow{w} j, \text{ si } a_{ij} = w \in (0, 1]$$

Nótese que: $\forall i, j \in N : a_{ij} = 0, a_{ij} \notin (0, 1] \implies \nexists e_{ij}$

⁵Un grafo es *dirigido* si los enlaces tienen un sentido definido. En nuestro caso el origen es el sector vendedor y el destino el comprador.

⁶Un grafo es *completo* si cada nodo está enlazado con cualquier otro.

Una vez creado el grafo, procedemos a analizar la **centralidad** de cada nodo y a analizar posibles estructuras de **comunidad**.

Comenzaremos por la centralidad, que debemos definir.

Definition

La *centralidad* (de un nodo V) se refiere a una medida posible que determina su importancia relativa dentro de éste.

Como puede observarse, la definición es abierta, por lo que depende de la especificación concreta que le queramos dar. Nosotros veremos tres:

- 1 Centralidad de grado
- 2 Centralidad a través de los autovectores
- 3 PageRank

Comenzaremos por la **centralidad de grado**. Es la más directa y se define como el número de enlaces (o el peso de estos, en nuestro caso) que tiene cada nodo como origen (outdegree) o destino (indegree). Sea el grafo $G = (N, E)$, con $v \in N$ y $V(v) \subset N$ vecinos del nodo y $w_{ij} \in (0, 1)$ el peso del enlace $e_{ij} \in E$.

- $InDegree(v) = \sum_{j \in V(v)} w_{jv}$
- $OutDegree(v) = \sum_{j \in V(v)} w_{vj}$
- $DegreeTotal(v) = InDegree(v) + OutDegree(v)$

El degree mide la importancia en términos de **conexión directa**.

Podemos también medir la **centralidad** a través de los **autovectores**. Esta definición parte de la idea de que un nodo tendrá una puntuación más alta si tiene unas conexiones dadas con nodos de alta centralidad que las mismas conexiones con nodos de baja centralidad. Así, sea x_i la puntuación del sector i , a_{ij} los componentes de la matriz y el resto de notación como la dada, para un $\lambda \in \mathbb{R}$ arbitrario que no modifica el resultado.

Definition

$$x_i = \frac{1}{\lambda} \sum_{j \in V(i)} x_j = \frac{1}{\lambda} \sum_{j \in G} a_{ij} x_j$$

Esto es, calculamos directamente $\mathbf{Ax} = \lambda \mathbf{x}$.^a

^aPuede existir solución para varios autovalores λ . Sin embargo, por Perron-Frobenius, sólo el mayor de los autovalores garantiza que \mathbf{x} sea positivo en todas sus componentes.

Finalmente utilizaremos el **PageRank** de L. Page y S. Brin como base para Google. La idea es prácticamente la misma que la anterior, pero su definición varía un poco e incorpora un parámetro $\alpha \in [0, 1]$, que es la probabilidad de que un usuario abra un enlace en una página web (para nosotros, de que un shock se propague) y β .

Definition

El pagerank x_i de un nodo se define como:

$$x_i = \alpha \sum_{j \in V(i)} \frac{a_{ji}}{\text{OutDegree}(j)} x_j + \beta$$

Es importante tener en cuenta que, por su definición, el PageRank da únicamente importancia a los enlaces entrantes, i.e., a mayor puntuación, más expuesto estaría el sector a un shock downstream.

Una pregunta que no podemos hacer es si tiene sentido discernir entre estas tres métricas o si realmente miden lo mismo. Para ello calculamos la matriz de correlación⁷ de Spearman (basado en la posición dentro de la variable de cada observación).

$$\begin{pmatrix} 1 & .56 & .35 \\ & 1 & .86 \\ & & 1 \end{pmatrix}$$

Podemos ver que efectivamente sí tiene sentido separar los conceptos, sobre todo entre el grado y las otras dos métricas, como imaginábamos.

⁷La primera variable es el *degreeTotal*, la segunda el *eigenValue* y la tercera el *pageRank*.

A continuación se presenta una tabla con los sectores *centrales* de la economía según cada métrica:

Figure: Ranking de mayor puntuación por método (I)

R	GradoEntrante	GradoSaliente
1	Alimentación, Bebidas y Tabaco (5)	Artes Gráficas (9)
2	Construcción (27)	Actividades Relacionadas con el Empleo (51)
3	Comercio al por Mayor (29)	Industrias Extractivas (4)
4	Vehículos de Motor (20)	Actividades Postales y de Mensajería (35)
5	Alojamiento, Comidas y Bebidas (36)	Actividades de Seguridad e Investigación (53)

Figure: Ranking de mayor puntuación por método (II)

R	EigenVector	PageRank
1	Alimentación, Bebidas y Tabaco (5)	Construcción (27)
2	Construcción (27)	Alimentación, Bebidas y Tabaco (5)
3	Vehículos de Motor (20)	Vehículos de Motor (20)
4	Alojamiento, Comidas y Bebidas (36)	Agencias de Viajes, Op. Turísticos (52)
5	Comercio al por Mayor (29)	Alojamiento, Comidas y Bebidas (36)

Ahora buscaremos estructura de **comunidad** en la red. Esto es, comprobar si hay grupos de sectores *semicerrados* de tal manera que un shock en uno de ellos se propagara a los demás nodos del grupo, pero no al resto de la red.

Definition

Decimos que una red presenta **estructura de comunidad** si los nodos de la red pueden ser *fácilmente* agrupados en conjuntos de nodos tal que dichos nodos estén conectados densamente en su interior.

Para ello, dado las dificultades *técnicas* de buscar comunidades en un grafo como el nuestro (dirigido y con pesos), probaremos a buscarlas a través de tres algoritmos:

- 1 Label Propagation
- 2 RicciCurvature
- 3 Louvain

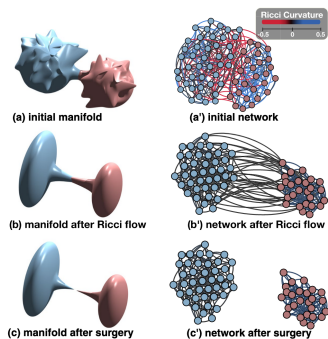
Comenzaremos por el **Label Propagation**, que es un algoritmo sencillo que es simplemente una votación: cada nodo comienza con una etiqueta y en cada iteración se le asigna aquella etiqueta que más presenten sus vecinos (ponderado por los pesos y teniendo en cuenta la dirección) y desempataando aleatoriamente.

En nuestra red el algoritmo devuelve una comunidad central amplísima (43 sectores) y pequeñas comunidades de 1, 3 y 4 sectores aisladas. La de 4, e.g., es una comunidad formada por:

- 46. Servicios técnicos de arquitectura e ingeniería; ensayos y análisis técnicos
- 18. Fabricación de material y equipo eléctrico
- 27. Construcción
- 14. Fabricación de otros productos minerales no metálicos

Análisis Estadístico de la Red Sectorial (12/19)

El algoritmo basado en la Curvatura de Ricci (un tensor aplicado en la teoría de la relatividad) es más complejo y puede verse la teoría en este [enlace] y mejor seguir la detección de comunidades para nuestra red en el notebook adjunto [aquí]. La idea (física y de redes) es la de las siguientes imágenes:



Fuente: Paper Original

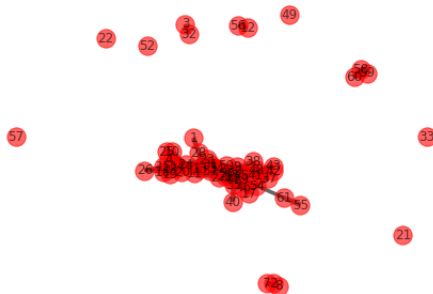
Los resultados podemos entenderlos como las capas de la tierra: en una primera iteración, el algoritmo devuelve el núcleo y el manto unidos e indistinguibles y, de ellos, se desprenden ciertas comunidades pequeñas y aisladas.

En el segundo paso, el manto se desprende del núcleo y forma pequeñas comunidades aisladas mientras la estructura del núcleo no es observable. Finalmente, podemos acabar viendo la estructura del núcleo central de la economía.

Como ejemplo, el núcleo central estaría formado por tres comunidades. En la siguiente diapositiva su imagen.

Análisis Estadístico de la Red Sectorial (14/19)

La primera iteración:

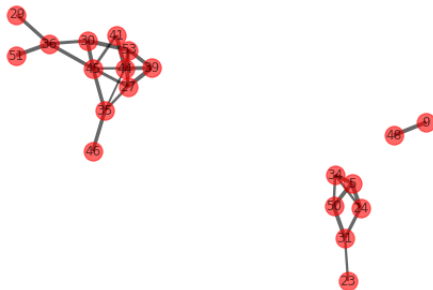


Fuente: Elaboración Propia.

Algunos sectores que se ven: 57, *Actividades de servicios sociales*; 21, *Fabricación de otro material de transporte*; o 33, *Transporte aéreo*.

Análisis Estadístico de la Red Sectorial (15/19)

Aquí, la estructura del núcleo:



Fuente: Elaboración Propia

Como ejemplo, hay una comunidad formada por 9 (*Artes gráficas y reproducción de soportes grabados*) y 48 (*Publicidad y estudios de mercado*).

Otra tiene como elementos el 5(Industria de la alimentación, fabricación de bebidas e industria del tabaco), 34(Almacenamiento y actividades anexas al transporte) o 50(Actividades de alquiler).

Nótese que esto ya es una **partición del núcleo** de la economía.

Finalmente, el algoritmo de **Louvain**, basado en la *modularidad*⁸, nos da comunidades como las siguientes:

① Comunidad 1:

- ① 9. Artes gráficas y reproducción de soportes grabados.
- ② 12. Fabricación de productos farmacéuticos básicos y preparados farmacéuticos.
- ③ 40. Programación, consultoría y otras actividades relacionadas con la informática; servicios de información.
- ④ 56. Actividades sanitarias.

② Comunidad 2:

- ① 35. Actividades postales y de mensajería.
- ② 39. Telecomunicaciones.

Lo bueno de este algoritmo es que nos devuelve comunidades más o menos homogéneas (entre 2 y 11 sectores), pero generalmente de 4 a 6.

⁸La modularidad es la fracción de los enlaces que caen dentro de grupos dados menos el valor esperado que dicha fracción hubiese recibido si los enlaces se hubiesen distribuido al azar.

Como último punto de esta sección IV, incorporaremos la **predicción de links**.

En una red, el problema de predecir un enlace no existente se resuelve buscando aquellos nodos con mayor similitud no conectados, según diferentes nociones de similitud. Nosotros usaremos dos:

- 1 Adamic Adar, recoge la idea de sentido común de que dos elementos son más parecidos si **comparten características -o vecinos- raros**.

Formalmente: $s_{xy} = \sum_{z \in V(x) \cap V(y)} \frac{1}{\ln \#V(z)}$

- 2 Preferential Attachment, no es un gran método y está basado en la idea de *“el rico cada vez se hace más rico”*. Formalmente:

$s_{xy} = \#V(x)\#V(y)$, que sigue una dist. de Pareto.

Para ello, he tenido que modificar un poco la red. Puede verse -de nuevo- [aquí].

Análisis Estadístico de la Red Sectorial (18/19)

Estos son los cinco enlaces más propensos a realizarse bajo el enfoque del **Adamic Adar**:

- ① (5) Industria de la alimentación, fabricación de bebidas e industria del tabaco y (44) Actividades inmobiliarias.
- ② (5) Industria de la alimentación, fabricación de bebidas e industria del tabaco y (53) Actividades de seguridad e investigación.
- ③ (30) Comercio al por menor, excepto de vehículos de motor y motocicletas y (29) Comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas.
- ④ (24) Suministro de energía eléctrica, gas, vapor y aire acondicionado y (31) Transporte terrestre y por tubería.
- ⑤ (9) Artes gráficas y reproducción de soportes grabados y (56) Actividades sanitarias.

Nosotros, desde el punto de vista económico, no debemos pensar que estos sectores vayan a unirse sino que son paralelos en su desarrollo. Como ejemplo muy trivial e irreal: una empresa que fabrique muebles de uso común y otra que fabrique lápices para arquitectos comprarán sus inputs a la misma maderera y venderán a la misma constructora. Esto nos daría alta similaridad pero no parece que vaya a producirse una relación comercial entre ellas.

Discrepancias y puntos a tratar a continuación (1/1)

Algunos puntos a tratar son:

- En las variables ¿cuándo logs, diferencias, etc...?
- Especificación de δ_t .
- Fijarse en utilizar variables reales o nominales con sentido.
- ...

Bibliografía y otros enlaces:

- **Carpeta de Google Drive con todos los Archivos.** [LINK].
- Ni, C., Lin, Y., Luo, F. et al. Community Detection on Networks with Ricci Flow. Sci Rep 9, 9984 (2019). [LINK]
- Baqaee, David, and Emmanuel Farhi. 2018. “Macroeconomics with Heterogeneous Agents and Input-Output Networks.”.[LINK]
- Acemoglu, D., Akicigit, U., and Kerr, W. 2015. “Networks and the Macroeconomy: An Empirical Exploration”. [LINK]
- Una introducción al Análisis de Redes en 5 min. [LINK]

FIN!

FIN!