

Desarrollo de una librería open source para análisis multivariante con modelos Disjoint-PARAFAC y Disjoint-Tucker3

PROBLEMA

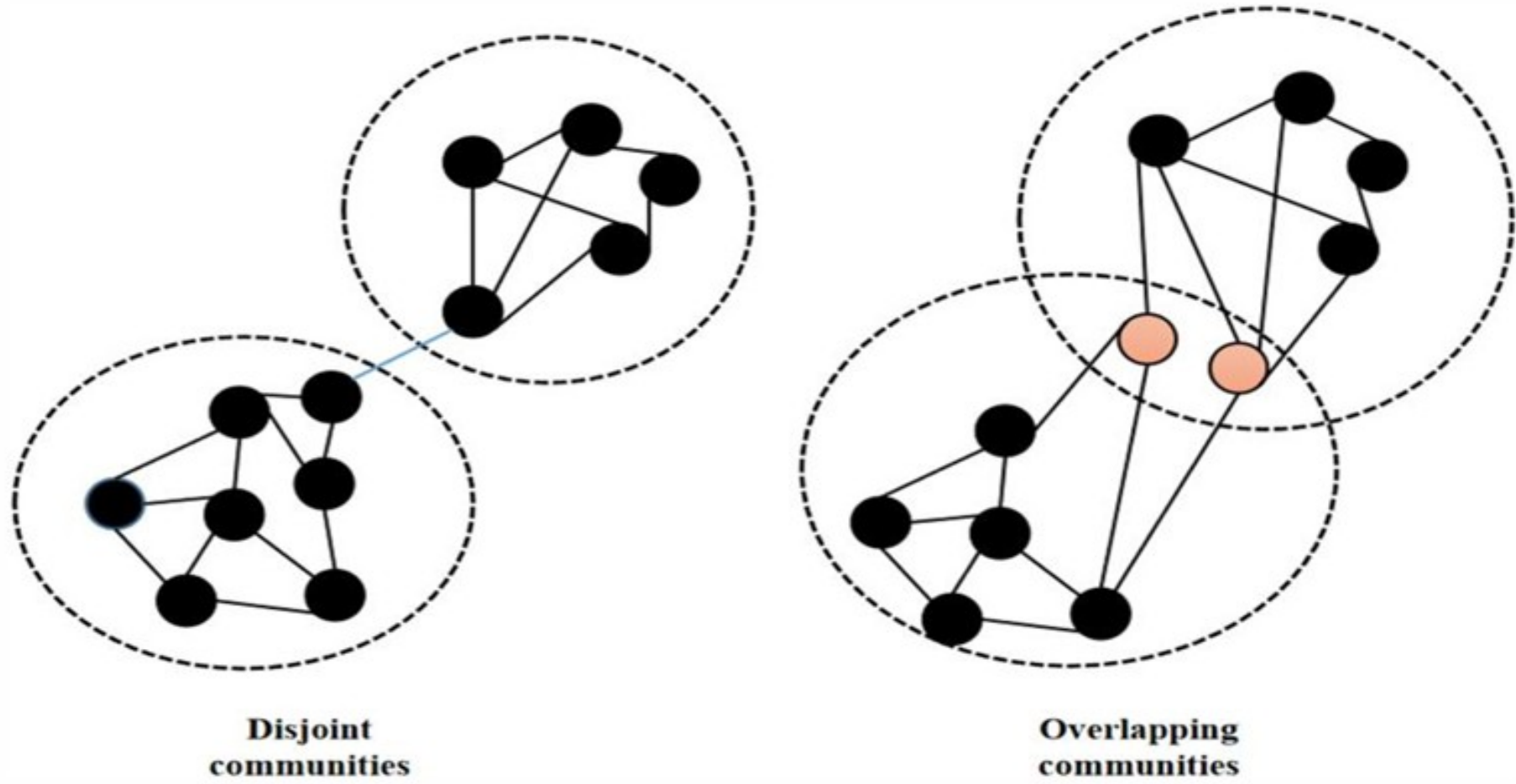
PARAFAC y Tucker3 son modelos matemáticos fundamentales para el análisis de datos de tres vías. Si bien sus variantes SPARSE mejoran la interpretabilidad de los resultados, actualmente no existe una librería de software de código abierto que facilite su implementación y uso práctico en la comunidad científica, particularmente en los lenguajes R y Python.

OBJETIVO GENERAL

Desarrollar una librería de código abierto en R y Python que facilite el análisis e interpretación de datos de tres vías utilizando técnicas SPARSE en los modelos matemáticos PARAFAC y Tucker3.

PROPUESTA

El proyecto se desarrolló mediante el uso de descomposición tensorial lineal aplicada a datos de tres vías estructurados como tensores de tercer orden. Se implementaron técnicas DPCA sobre los modelos PARAFAC y Tucker3, incorporando restricciones SPARSE para reducir el solapamiento entre componentes. La implementación se realizó en los lenguajes R y Python, permitiendo la exportación de resultados a hojas de cálculo para su análisis comparativo y visualización.

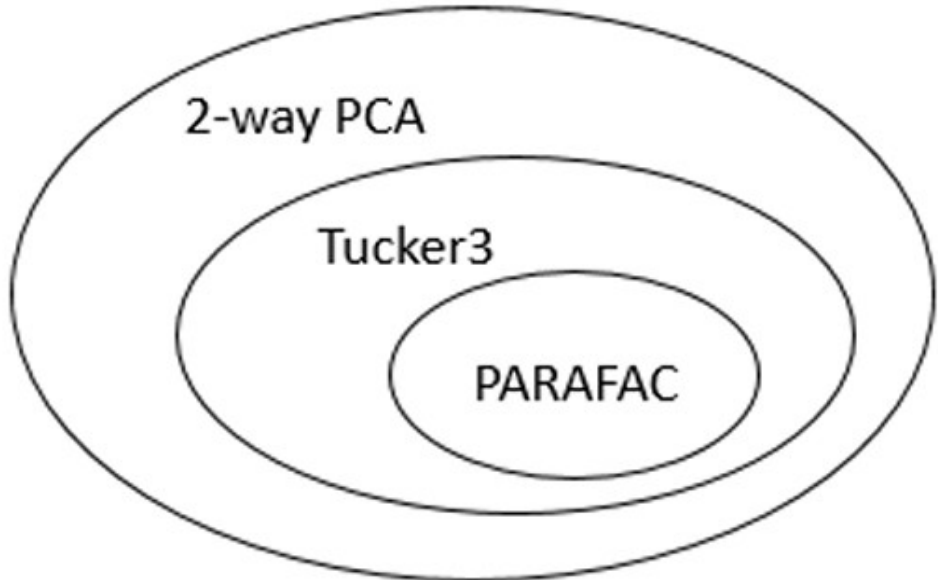


$$T \approx \begin{matrix} c_1 \\ | \\ a_1 \end{matrix} b_1 + \begin{matrix} c_2 \\ | \\ a_2 \end{matrix} b_2 + \dots + \begin{matrix} c_k \\ | \\ a_k \end{matrix} b_k = \begin{matrix} C \\ | \\ A \end{matrix} B$$

(a) PARAFAC decomposition

$$T \approx \begin{matrix} C \\ | \\ A \end{matrix} \begin{matrix} G \\ | \\ B \end{matrix}$$

(b) Tucker decomposition



RESULTADOS

Los modelos disjuntos generaron estructuras más claras y estables, con una reducción significativa del solapamiento entre componentes, lo que facilitó la interpretación de los resultados frente a los métodos clásicos.

Estos efectos se evidencian en las tablas comparativas donde las versiones SPARSE de PARAFAC y Tucker3 presentan una mejor separación de patrones latentes.

TuckerALS (Tradicional)

País	C1	C2	C3
Argentina	-0.32	0.28	-0.15
Brazil	-0.35	0.22	-0.18
Canada	0.28	0.45	0.32
Chile	-0.30	-0.18	0.25
Colombia	-0.38	-0.22	0.12
Ecuador	-0.33	-0.15	0.18
USA	0.35	0.52	0.28
Mexico	-0.32	-0.20	0.15
Peru	-0.28	0.35	-0.22
Uruguay	-0.25	-0.12	0.20

¿A qué grupo pertenece cada país?

DisjointTuckerALS (Nuestra Librería)

País	C1	C2	C3
Argentina	0	-0.52	0
Brazil	0	-0.51	0
Canada	0	0	-0.66
Chile	-0.44	0	0
Colombia	-0.52	0	0
Ecuador	-0.48	0	0
USA	0	0	-0.75
Mexico	-0.46	0	0
Peru	0	-0.69	0
Uruguay	-0.31	0	0

Agrupación clara por colores

CONCLUSIONES

La implementación de técnicas SPARSE en modelos PARAFAC y Tucker3 mejora significativamente la interpretabilidad de los datos de tres vías, permitiendo una estructura más definida y coherente de los componentes obtenidos.

Como trabajo futuro, se propone extender la librería a conjuntos de datos de mayor dimensionalidad y evaluar métricas adicionales de estabilidad y robustez para fortalecer su aplicación en distintos contextos científicos.