

Transitività

- Spesso si osserva similarità tra due nodi i e j in termini di caratteristiche individuali x_i e x_j
- Questo fatto può riflettersi sul valore della relazione $y_{i,j}$
- Immaginiamo che x_i sia indicatore che una persona i appartiene a un gruppo
- Allora $x_i x_j$ indica che i e j sono membri dello stesso gruppo ...
- ... potrebbe esserci un effetto su $y_{i,j}$: se positivo parliamo di **omofilia**

Transitività

- La presenza di omofilia può portare a strutture nella rete come **transitività**
- Spesso le reti hanno strutture di dipendenza tra 'tripleste' di nodi → **transitività**
- Formazione di triangoli tra nodi i, j, k
- Molteplicità di triangoli genera 'clusters'
- Misura globale di transitività: $\sum_{i,j,k} y_{i,j}y_{j,k}y_{i,k}$

Additive and multiplicative effects model, AME

- Vogliamo tenere conto della **dipendenza triadica**
- Interazione moltiplicativa di caratteristiche di nodo **non presenti nel dataset**
- Inclusione nel modello di caratteristiche di nodo **latenti**
- **Additive and multiplicative effects model, AME**

$$y_{i,j} = \beta^\top \mathbf{x}_{i,j} + a_i + b_j + \mathbf{u}_i^\top \mathbf{v}_j + \varepsilon_{i,j}$$

- \mathbf{u}_i è un vettore di fattori latenti, non osservati, riferiti al **comportamento di i come mittente**
- \mathbf{v}_j è un vettore di fattori latenti, non osservati, riferiti al **comportamento di j come ricevente**
- L'effetto moltiplicativo $\mathbf{u}_i^\top \mathbf{v}_j$ **quantifica dipendenza triadica**
- ma anche la **possibilità di 'variabili omesse'**

Esempio

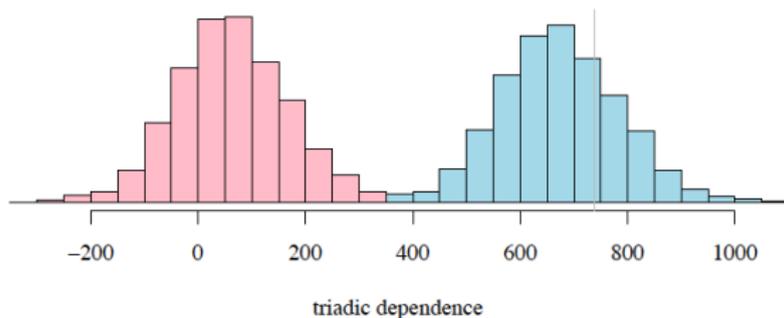
Confronto tra LM, SRRM e AME

regressor	LM			SRRM			AME		
	$\hat{\beta}$	$se(\hat{\beta})$	t -ratio	$\hat{\beta}$	$se(\hat{\beta})$	t -ratio	$\hat{\beta}$	$se(\hat{\beta})$	t -ratio
exp pol	0.015	0.004	4.166	0.015	0.016	0.934	0.012	0.016	0.782
imp pol	0.022	0.004	6.070	0.022	0.016	1.419	0.018	0.015	1.190
exp GDP	0.411	0.021	19.623	0.407	0.095	4.302	0.346	0.103	3.373
imp GDP	0.398	0.020	19.504	0.397	0.094	4.219	0.336	0.103	3.250
dist	-0.057	0.004	-13.360	-0.064	0.005	-11.704	-0.041	0.004	-10.970

- Dal punto di vista delle stime, le conclusioni del modello AME sono coerenti con quelle del SRRM

Esempio

Ma la **dipendenza triadica** viene colta dal modello AME (azzurro) e non dal SRRM (rosa)



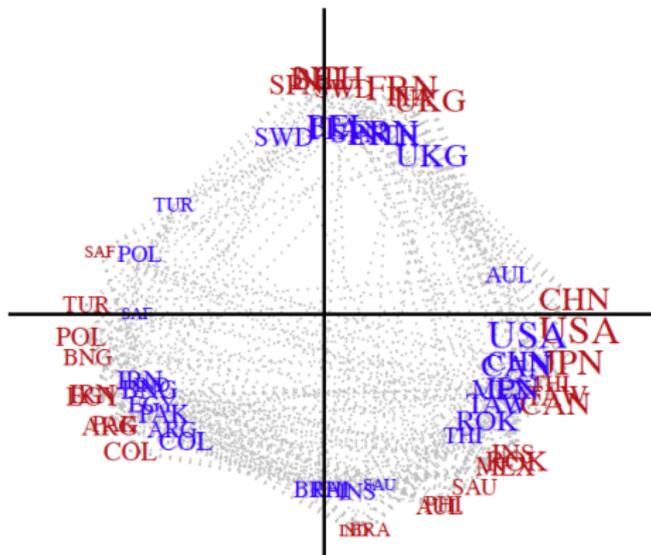
Esempio

Dipendenza triadica può essere studiata anche attraverso il 'circle plot'

Stime di u_i in rosso e di v_i in blu.

I nomi dei paesi indicano la direzione dei vettori e la grandezza del testo la magnitudine. Le linee tratteggiate descrivono coppie di nodi per cui l'import-export è maggiore di quello atteso sulla base degli altri termini del modello.

Le stime mostrano la presenza di gruppi di paesi (clusters).



Dati binari e ordinali

- Spesso i dati di rete sono di tipo **binario** (amici/non amici) o **ordinale** (basso/medio/alto)
- Modello AME può essere esteso a casi in cui la variabile risposta è binaria o ordinale
- Estensioni si basano su rappresentazione *a variabile latente* dei modelli **probit** e **probit per dati ordinali**

Probit AME

Modello **probit AME** per dati binari

$$\begin{aligned}y_{i,j} &= \beta^\top \mathbf{x}_{i,j} + a_i + b_j + \mathbf{u}_i^\top \mathbf{v}_j + \varepsilon_{i,j} \\s_{i,j} &= g(y_{i,j}) = 1(y_{i,j} > 0)\end{aligned}$$

La variabile binaria $s_{i,j}$ assume valore 1 quando la variabile continua latente $y_{i,j}$ è maggiore di 0.

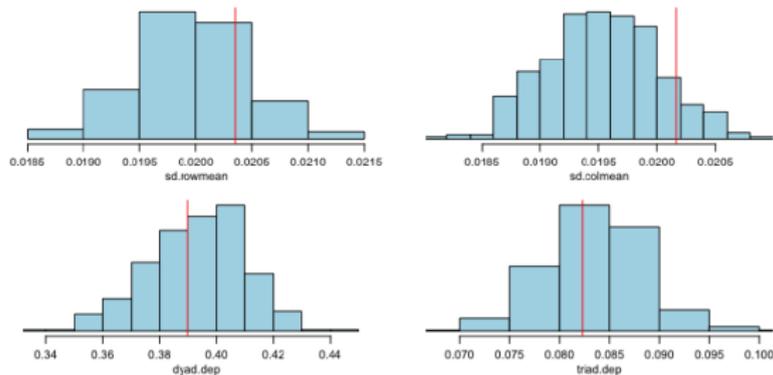
Torniamo al caso di studio ...

Probit AME: variabili di riga `.row`, variabili di colonna `.col`, variabili diadiche `.dyad`, dopo riorganizzazione `.w`

	pmean	psd	z-stat	p-val
intercetta	-1.059	0.119	-8.925	0.000
canale_dep.store.row	-1.161	0.116	-10.012	0.000
canale_outlet.row	-0.938	0.139	-6.741	0.000
canale_monobrand.w.row	-0.229	0.076	-3.031	0.002
canale_outlet.w.row	-0.362	0.127	-2.849	0.004
canale_monobrand.col	0.767	0.127	6.047	0.000
canale_outlet.col	0.957	0.143	6.700	0.000
canale_monobrand.w.col	0.239	0.072	3.334	0.001
Y_lag.dyad	0.731	0.054	13.650	0.000
cross_country.dyad	-3.934	0.291	-13.520	0.000
cross_channel.dyad	-0.749	0.084	-8.883	0.000
cross_country.w.dyad	1.249	0.230	5.434	0.000
cross_channel.w.dyad	0.111	0.067	1.650	0.099

Torniamo al caso di studio ...

Adattamento globale del modello



I risultati di questo modello risultano globalmente soddisfacenti

Torniamo al caso di studio . . .

Valutazioni di business

- La quantità di merce ricevuta e venduta non risultano significative e non sembrano influenzare le scelte degli allocator.
- Presenza di **dipendenza temporale** è significativa.
- Riduzione della quantità di materiale in partenza da e verso i negozi **monobrand**.
- Monobrand e outlet sono meno coinvolti nel processo in quanto mittenti, ma acquisiscono importanza in quanto **destinatari del materiale**.
- Maggiore propensione all'uso di **spedizioni internazionali**.