Clustering gerarchica dei corsi obbligatori ad Economia

Hierarchical clustering for required courses of Economics faculty

Tarsitano Agostino
Dipartimento di Economia Politica - Università della Calabria

Abstract: ordinal data are useful to quantify constructs on a non metric scale. Several indices have been devised to evaluate similarities among judges. Such indices will be considered as distance measures for a hierarchical clustering of required courses in a degree program of Economics faculty

Parole chiave: ordinal measuremetns, hierchical clustering, teaching evaluations.

1. Introduzione

Un'indagine molto diffusa è la seguente: si chiede a dei giudici di disporre secondo una graduatoria priva di *ex equo* un insieme di *item* senza però pronunciarsi sullo scarto tra i vari livelli di giudizio. Si potrebbe anche pensare ad una successione di confronti a coppie in cui il soggetto accetta l'esito del confronto solo se è coerente con la sua graduatoria di valutazione. Qual'è il grado di prossimità tra le valutazioni? Esitono gruppi di di giudici distinti le cui classfiche tendano a somigliarsi?

2. Distanze tra graduatorie

Consideriamo due giudici qualsiasi *i* e *j* tra gli *n* rilevati e siano u_{ik} e u_{jk},k=1,…,m i ranghi da loro attribuite ad *m* item prefissati secondo lo schema:1=massima preferenza, m=minima preferenza. Le due graduatorie sono in perfetto accordo se $u_{ik}=u_{jk}$,k=1,2,…,m ed in completo disaccordo se $u_{ik}=k$ (questa è nota come graduatoria fondamentale) e $u_{jk}=(m-k+1)$ per k=1,2,…,m. Perché d_{ij} quantifichi la vicinanza di giudizio deve soddisfare le condizioni di metricità: 1.$d_{ij}=0$ se e solo se $u_{ik}=u_{jk}$, k=1,…,m; 2. $d_{ij}=d_{ji}\geq 0$ per ogni coppia (i,j); 3.$d_{ij}\leq d_{ir}+d_{jr}$ per ogni terna (i,j,r). Nel presente lavoro ho escluso sia la possibilità di trattare le classifiche come variabili multilivello (le posizioni in progetto erano troppo numerose) che di trattarle come metriche (almeno per la classificazione dei soggetti) perché non agganciabili ad un continuo percettivo univoco. Si sono perciò usati indici di distanza tra classifiche.

Coefficienti di disaccordo

$$ {}_1d_{ij}(a)=\left\{\frac{\sum_{k=1}^{m}\left|u_{ik}-u_{jk}\right|^a}{\underset{u_i,u_j\in G}{Max}\sum_{k=1}^{m}\left|u_{ik}-u_{jk}\right|^a}\right\}^{\frac{1}{a}}; \qquad 0\leq\ {}_1d_{ij}(a)\ \leq 1 $$

dove il massimo è determinato rispetto alle m! permutazione dei primi *m* naturali positivi. La verifica della metricità è immediata. Per a=2 il massimo è $m(m^2-1)/3$ e si ha ${}_1d_{ij}(2)=c_{ij}=[(1+\rho_{ij})/2]^{0.5}$ dove ρ_{ij} è il coefficiente di Spearman. Inoltre:

$$ c_{ij}=\frac{3\sum_{r=1}^{m-1}\sum_{s=r+1}^{m}w_{rs}}{m\left(m^2-1\right)}; \qquad con \quad w_{rs}=\begin{cases}\left(u_{js}-u_{jr}\right)+(s-r) & se\ \ u_{jr}>u_{js}\\ 0 & se\ \ u_{jr}<u_{js}\end{cases} $$

in cui si chiarisce che le coppie discordi contribuiscono in ragione dello spiazzamento dei due ranghi rispetto alla graduatoria fondamentale. Un altro coefficiente di disaccordo si ottiene per a=1 con massimo $[(3-b)m^2-2(1-b)m-2b]/4$ dove b=1 se *m* è dispari e zero altrimenti. In parte, tale misura è collegabile all'indice di cograduazione proposto da Gini.

Distanza di Rizzi. Rizzi (1972) misura la distanza tra due graduatorie con il numero di inversioni che presenta una graduatoria rispetto all'altra. Tale numero è pari al minimo di spostamenti tra coppie di elementi contigui che occorre effettuare per passare da una graduatoria all'altra:

$$ {}_2d_{ij} = \frac{\sum_{k=1}^{m-1}\sum_{l=k+1}^{m} I\left(u_{jk} > l\right)}{0.5m(m-1)}; \quad I(a>b) = \begin{cases} 1 & se \quad a>b \\ 0 & altrimenti \end{cases} $$

In pratica, per ogni elemento u_{jk} della graduatoria fondamentale si conta il numero di elementi della graduatoria abbinata a sinistra di *k* che sono maggiori di *k* (cioè il numero di coppie aventi *k* come 2° componente); tale numero è poi rapportato al massimo per normalizzarlo. Nel lavoro di Rizzi sono anche dimostrate le condizioni di metricità. Il risultato di Moran (1950) permette di ricondurre la distanza di Rizzi al τ di Kendall: ${}_2d_{ij}=1-2\tau_{ij}$.

Distanza di Gordon. Nel tentativo di meglio qualificare la concomitanza di giudizio rispetto al τ di kendall,Gordon (1979) ha suggerito di misurare il divario tra due graduatorie in base al numero di ranghi da rimuovere per arrivare a due graduatorie perfettamente concordanti. Se la graduatoria del soggetto *i* è riscritta come quella fondamentale, l'indice di Gordon è legato al numero di posizioni costituenti la sottosequenza monotona crescente più lunga nei ranghi di *j*:

$$ {}_3d_{ij} = 1 - \frac{\underset{1\le k\le m-1}{Max}\left\{\sum_{l=k+1}^{m} I\left(u_{jk}, u_{jl}\right)\right\}}{m}; \qquad 0 \le {}_3d_{ij} \le 1 $$

Che l'indice di Gordon sia una metrica lo si dimostra con ragionamenti analoghi a (Rizzi, 1972):1) solo per graduatorie coincidenti non occorre effettuare eliminazioni; 2) il minimo di eliminazioni che portano u_j a coincidere con u_i è lo stesso di quello necessario a far coincidere u_i con u_j; 3) ${}_3d_{ir2}+{}_3d_{jr}\geq{}_3d_{ij}$ perché se così non fosse sarebbe possibile arrivare da *i* a *j* con un numero di eliminazioni minore di ${}_3d_{ij}$.

3. Strategia di aggregazione

La procedura di classificazione gerarchica aggregativa delle N entità $\{u_1,u_2,\ldots,u_n\}$ è semplice: partendo da una clustering $L_0=\{C_i=(u_i),\ i=1,2,\ldots,n\}$ costituita da N clusters contenenti ognuno un'entità si perviene, per fusioni successive della coppia di clusters più prossima, a $L_{n-1}=\{u_1,u_2,\ldots,u_n\}$ che le contiene tutte. E' quindi necessario definire una metrica $D(C_i,C_j)$ per i clusters di L_q, q=0,1,…,n-1 che selezioni i due da aggregare; supponiamo che ciò avvenga per una distanza h_q allo stadio q-esimo; a tale livello la clustering L_q include (n-q) clusters di cui almeno uno con più di un elemento se q>0. La sequenza (L_q,h_q), q=1,2,…,n-1 contiene tutti i dettagli. Ogni passaggio è quello più compatibile con la clustering q-esima e la data metrica D con riferimento ad una coppia di cluster, ma non vi è garanzia che sia quello più efficiente rispetto a 3, 4, o più entità né è sicura l'ottimalità del risultato finale. Peraltro, il cluster formatosi al livello q non può più essere diviso, ma solo combinato con altri.

La matrice delle distanze per il livello q=0 coincide con la matrice delle dissimilarità discussa nel paragrafo precedente. Per quelli successivi la matrice si ottiene cancellando la riga e la colonna corrispondenti ad uno dei clusters C_h e C_k interessati da una fusione e sostituendo la riga e la colonna dell'altro con il ricalcolo delle distanze dal nuovo cluster $C_h\cup C_k$ in base alla metrica D; gli altri elementi rimangono invariati e si trasferiscono come tali alla clustering successiva fino a che non siano coinvolti in una aggregazione.

Le tecniche aggregative più note ricadono nella formula di Lance e Williams:

$$LW: \ d\left(C_i \cup C_j, C_k\right) = \alpha_i d\left(C_i, C_k\right) + \alpha_j d\left(C_j, C_k\right) + \beta d\left(C_i, C_j\right) + \gamma \left| d\left(C_i, C_k\right) - d\left(C_j, C_k\right) \right|$$

con $\alpha_i+\alpha_j+\beta\geq 0$ e $\alpha_i+\gamma\geq 0$ per assicurare che la distanza sia non negativa. Ogni combinazione dei parametri $(\alpha_i,\alpha_j,\beta,\gamma)$ esprime un diverso algoritmo aggregativo alcuni corrispondenti ad interpretazioni di entità in uno spazio euclideo ed altri che ne prescindono. Per districarsi tra la miriade di tecniche disponibili (peraltro non tutte inserite nella formula LW) Chen e Van ness (1998) suggeriscono di sceglierere la procedura dopo aver individuato dei criteri di ammissibilità a cui questa deve rispondere. Nel caso dellegraduatorie, in cui la scala di misurazione ha poca rilevanza, sembra appropriato il requisito della ammissibilità monotona e cioè che se la matrice originale delle distanze subisce una trasformazione monotona in qualsiasi stadio della gerarchia la risultante clustering non cambia. Chen e Van ness dimostrano che le uniche combinazioni monotonicamente ammissibili sono il legame singolo ed il completo. Se poi si aggiunge il requisito della ammissibità metrica e cioè le dissimilirà tra clusters debbono costituire una metrica ad ogni livello di aggregazione il cerchio si restringe al solo legame completo che quindi sembra la strategia più indicata per la nostra applicazione.

Si deve però considerare che, se il legame completo gode delle proprietà indicate (ed altre su cui non ci soffermiamo, quali la gradita ultrametricità) non manca di inconvenienti: 1) provoca un effetto di separazione che mantiene lontane entità molto prossime perché incluse in clusters che contengono -anche solo un'entità per cluster- molto distanziate; 2) è fortemente spazio-dilatativo tendendo a formare gruppi poco numerosi; 3)nasconde gli outliers aggregandoli in clusters apparentemente ragionevoli; 4) può generare molte parità nei livelli di aggregazione con difetti di arbitrarietà nella sequenza gerarchica che potrebbe cambiare secondo l'ordine di considerazione delle unità o di non riproducibilità del risultato nel caso di scelte per sorteggio. Insorge perciò l'esigenza di combinare il legame completo con la distanza tra graduatorie più sensibile alle differenza tra i dati e meno incline a generare parità nella matrice delle distanze. Nelle prove preliminari effettuate l'indice più elastico è risultato il coefficiente quadratico di disaccordo.

4. Applicazione

La valutazione dei docenti è prassi comune in molte università. E' però lecita la domanda: " Che cosa esattamente valutano gli studenti?". Una parte della risposta potrebbe risiedere nel carattare perentorio della presenza di un certo insegnamento nel piano di studio (cfr. Kim, Damewood, Hodge, 2000). Per saggiare taler atteggeggiamento ho distribuito dei questionari agli studenti iscritti al corso di laurea in discipline economiche e sociali della facoltà di economia di Arcavata (Cs) in cui, oltre a delle semplici domande, sulla percentuali di esami già superati (±50%), sulla media (±26) , se in corso o fuori corso, maschio o femmina e tipo di diploma della scuola superiore, ho inserito i seguenti quesiti:

1. <u>Disponga i seguenti corsi fondamentali e obbligatori secondo l'utilità per la preparazione a gli altri insegnamenti del suo piano di studio</u>. 1=meno utile .... 14=più utile.
Economia aziendale, Economia politica, Istituzioni di diritto privato, Istituzioni di diritto pubblico, Istituzioni di economia, Lingua inglese, Matematica finanziaria,Matematica generale, Ragioneria generale ed applicata, Scienza politica, Sociologia, Statistica, Storia contemporanea, Storia economica.
2. <u>Disponga i seguenti corsi fondamentali e obbligatori secondo le difficoltà che si incontrano nell'affrotarne l'esame.</u> . 1=più facile .... 14=meno facile

La partecipazione all'indagine a è stata libera, gratuita e non svolta in presenza di docenti (cfr. Braskamp e Ory , 1994). Per attenuare la tendenza ad assegnare i voti più alti o più bassi in ragione dell'ordine di presentazione (cfr. Brook e Upton, 1975), sono stati redatti 20 questionari tipo inserendo -per ogni domandea- una permutazione casuale dei 14 insegnamenti. Ogni tipo è stato replicato 7 volte. Sono rtornati 67 questionari validi.

Il diagramma di dispersione nella figura 1 riporta le posizioni medie assegnate agli insegnamenti per i due aspetti indagati. E' spiccata la presenza di alcuni gruppi di materie: in basso a sinistra le storie, sociologia e scienza politica che pur creando poche difficoltà di superamento sono giudicate meno formative; più difficili, ma più rilevanti per la preparazione risultano: economia aziendale, diritto pubblico e ragioneria con economia istituzionale che si aggrega. Non meraviglia la collocazione di statistica e matematica finanziaria tra gli insegnamenti mediamente utili, ma il cui superamento è faticoso. Al centro sono collocate inglese, economia politica e, con leggera sopresa, matematica generale. Nella figura 2 è riportata la clustering degli studenti per le due graduatorie espresse (Il taglio trasversale mostrato nei grafici è la *stopping rule* di Mojena). Rispetto ai giudizi di importanza (coph.corr. 0.56) si vedono dei gruppi che però risultano privi di una chiave di lettura in termini delle poche variabili dicotome rilevate; la clustering rispetto alle difficoltà (Coph.corr. 0.66) vede almeno due grandi gruppi suddivisi per la prevalenza di medie inferiori a 26 (lato sinistro). Alla percezione degli insegnamenti non è perciò estranea la valutazione media avuta nel complesso degli insegnamenti.

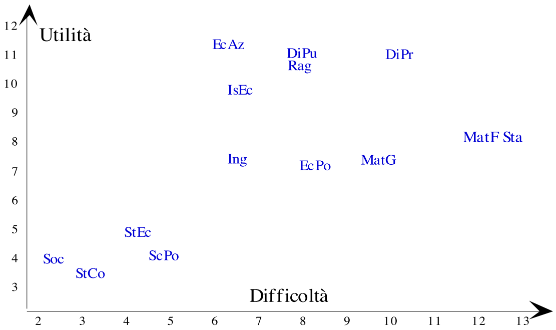

Figura 1: scatterplot dei punteggi medi

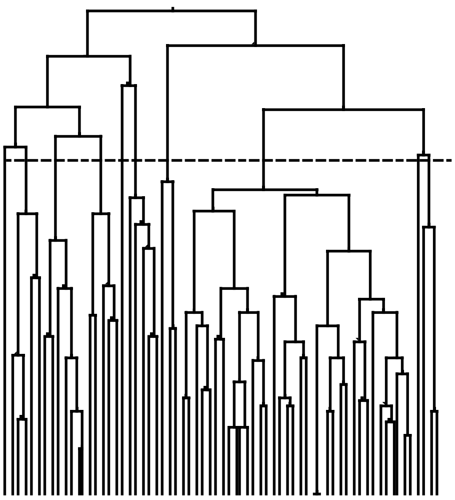
Figura 2a: clustering degli studenti per utilità

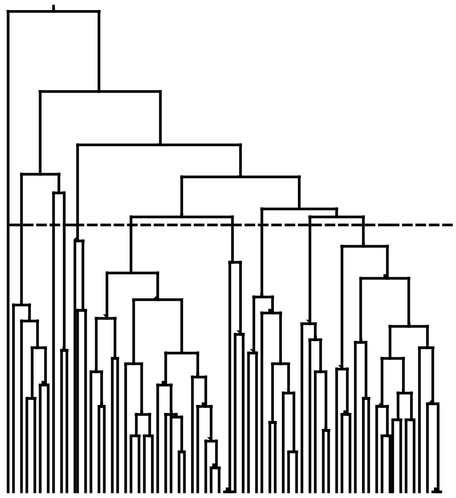
Figura 2a: clustering degli studenti per difficoltà

In conclusione lo studio ha mostrato come la clutering gerarchica fornisce un valido ausilio per chiarire la percezione che gli studenti hanno degli insegnamenti loro offerti nel piano formativo. L'efficacia della tecnica rimane elevata anche con l'uso di meno sofisticate, ma più attendibili graduatorie.

Riferimenti bibliografici

Braskamp L.A. Ory C. (1994). Assessing faculty work. Jossey-Bass Publishing, San Francisco.

Brook D. Upton G.J.G. (1975). Biases in local government elections due to position on the ballot paper. *Applied Statistics*. 23, 414-419.

Chen Z. Van Ness J.(1998). Characterization of nearest and faurthest neighbor algorithms by clustering admissibility conditions". *Pattern Recognition*, 31,1573-1578.

Gordon A.D; (1979). A measure of the agreement between rankings. *Biometrika*, 66, 7-15.

Kim C. Damewood E. Hodge N. (2000). Professor attitude: its effect on teaching evaluations. *Journal of Management Education*, Vol. 24, pp. 458-473.

Moran P.A.P. (1950). Recent development in ranking theory. *J.R.S.S. B*,12, 153-162.

Rizzi A. Badaloni M. (1972) Contributi alla cluster analysis. *Metron*. 30, 154-208.