

<http://www.mauroennas.eu>

Elementi di scaling multidimensionale per la classificazione e il posizionamento nelle ricerche di marketing

Mauro Ennas

Allegati
Simulazioni SPSS

- 1 MDS_SPSS_esempio_adl.sav
 - 2 MDS_SPSS_città_italiane_proxcal.sav
 - 3 MDS_PROXSCAL_input_25e_pscalare.sav
 - 4 MDS_PROXCAL_input_25e_distanze.sav
-

Fogli di calcolo Excel

- 1 003_PW_PARMA_ENNAS_MDS_distanze_italiana.xls
-

© 2010

Indice degli argomenti

Indice degli argomenti	4
Indice delle figure	5
Indice delle tabelle	5
Indice syntax e script	5
Multidimensional scaling	7
Metodi di scaling multidimensionale di tipo metrico	8
Metodi di scaling multidimensionale di tipo non-metrico	9
Cenni sui metodi di scaling multidimensionale classico in SPSS.....	10
Esempio numerico: analisi MDS dei dati di input dell'analisi ADL.....	11
Multidimensional Scaling con SPSS ALSCAL: distanze tra città	14
Multidimensional Scaling con SPSS PROXSCAL: distanza tra città.....	20
Glossario	25
Bibliografia	29
Indice dei nomi	32

Indice delle figure

Figura 2 - MDS PROXCAL: i dati in ingresso sono dati di prossimità (distanze).....	12
Figura 3 - Analisi MDS PROXCAL (i dati in ingresso non sono dati di prossimità, SPSS provvederà a trasformarli in distanze prima dell'analisi).....	13
Figura 4 - Analisi multidimensionale MDS ALSCAL (i dati di input sono dati di prossimità).....	13
Figura 5 - Scree plot.....	15
Figura 6 - Mappa delle percezioni.....	17
Figura 7 - Diagramma di Shepard (caso metrico).....	18
Figura 8 - Diagramma di Shepard 1 (caso non-metrico).....	19
Figura 9 - Diagramma di Shepard 2 (caso non-metrico).....	19
Figura 10 - Rappresentazione in tre dimensioni.....	22
Figura 11 - Grafici dei residui (distanze vs. dissimilarità trasformate ricavate dal modello).....	23

Indice delle tabelle

Tabella 1 - Dati di ingresso della procedura di Multidimensional Scaling (MDS) PROXSCAL con SPSS (tali dati non sono dati di prossimità e dovranno essere preventivamente trasformati da SPSS in dati di prossimità prima dell'analisi di posizionamento metrico): la matrice è di dimensione 25×25, simmetrica (sono stati omessi i dati della parte triangolare).....	11
Tabella 2 - Matrice delle distanze (matrice di prossimità), costruita a partire dalla matrice dei prodotti scalari. Questa matrice viene utilizzata come input del calcolo MDS PROXCAL con dati di prossimità e al calcolo ALSCAL per il posizionamento multidimensionale.....	12
Tabella 3 - Distanze tra le città italiane (matrice simmetrica).....	14
Tabella 4 - Valori dello stress e del miglioramento nel calcolo iterativo dell'S-Stress.....	16
Tabella 5 - Correlazione quadratica e Stress.....	16
Tabella 7 - Cronologia delle iterazioni.....	20
Tabella 8 - Misure di stress e adattamento.....	21
Tabella 9 - Decomposizione dello stress non trasformato normalizzato.....	21
Tabella 10 - Coordinate finali.....	22

Indice syntax e script

Syntax 1 - SPSS GEMSCAL (disponibile in SPSS tramite SYNTAX).....	10
Syntax 2 - SPSS MDS ALSCAL.....	14
Syntax 3 - SPSS MDS PROXCAL.....	20

Multidimensional scaling

Il *multidimensional scaling* (MDS) permette di rappresentare il profilo di un oggetto in più di una dimensione allo stesso tempo. Si assume che l'osservatore dell'oggetto percepisca le differenze con altri oggetti, anche non correlati con esso, per similitudine e dissimilitudine. Il pregio maggiore di questo metodo sta proprio nel riuscire a scalare le distanze psicologiche tra gli oggetti in termini di dissimilarità e rappresentarle su mappe multidimensionali. Nel *marketing* si suppone che il consumatore abbia delle sue specifiche mappe mentali che lo guidano nell'associare gli oggetti tra loro per similarità o ad allontanarli e mantenerli separati tra loro per dissimilarità. Queste mappe mentali non sono esplicite e il MDS permette di renderle esplicite.

Il grado di similarità o di dissimilarità è legato alla distanza tra due punti in uno spazio di rappresentazione, tipicamente bidimensionale o tridimensionale.

Esistono due sottogruppi principali di metodi che utilizzano il *multidimensional scaling*, il tipo metrico e quello non-metrico. Nel primo gruppo sono incluse le metodologie MDS che partono da dati rappresentati da una scala graduata di riferimento (dati quantitativi), in quello non metrico i dati iniziali sono espressi tramite punteggi (*rank*) e sono dei giudizi qualitativi. Entrambe le tecniche pongono la loro attenzione sulle relazioni ed hanno entrambe come scopo di trovare, allo stesso tempo, le dimensioni e le configurazioni (*pattern*) dei punti rappresentabili in quelle dimensioni, in modo tale che le distanze intermedie rispetto ai punti calcolati risultino significative. Entrambe le tecniche sono in grado di fornire soluzioni accurate ed applicate agli stessi casi riescono a fornire risultati comparabili per efficacia di rappresentazione. Nelle ricerche di *marketing* hanno ampia applicazione per confrontare prodotti e marchi, ma possono anche essere utilizzate per comparare dati sia qualitativi che quantitativi.

Le limitazioni del metodo MDS sono legate al fatto che in alcuni casi particolari i dati d'ingresso possono produrre risultati poco significativi dal punto di vista statistico tanto da non potere essere validabili. Le metodologie metriche sono più robuste di quelle non metriche e producono delle mappe molto dettagliate ove applicate. Le metodologie non-metriche soffrono della soggettività delle indicazioni e della alta variabilità temporale delle stesse, tanto che dati sugli stessi oggetti misurati da differenti osservatori forniscono risultati a volte contrastanti. In base ai dati di input possiamo distinguere lo scaling multidimensionale in quattro tipologie differenti:

- **Multidimensional scaling classico**

Introdotta da *Torgerson* e noto come *Torgerson Scaling*, prende in ingresso una matrice dati che rappresenta le dissimilarità tra coppie di oggetti e produce in *output* una matrice di coordinate la cui configurazione minimizza una funzione di perdita (*loss function*) chiamata *strain function*.

- **Multidimensional scaling metrico**

Una classe di metodi che include anche il metodo classico e generalizza la procedura di ottimizzazione con una varietà di funzioni di perdita e matrici di *input* di distanze note con o senza pesi. In questo contesto una funzione di perdita è nota come *stress* e viene minimizzata utilizzando una procedura apposita.

- **Multidimensional scaling non-metrico**

L'approccio MDS non-metrico trova sia la relazione monotona non-parametrica tra le dissimilarità derivanti dal confronto elemento per elemento nella matrice di *input*, la distanza *euclidea* tra oggetti, e la posizione di ciascun oggetto in uno spazio bidimensionale o tridimensionale. La relazione viene rilevata tipicamente utilizzando metodi di regressione.

- **Multidimensional scaling generalizzato**

Un'estensione dello scaling metrico multidimensionale che ha come risultato uno spazio arbitrario non-euclideo. Nel caso in cui le dissimilarità siano delle distanze su una superficie e il risultato cada su un'altra superficie, il MDSG permette di trovare la minima distorsione delle due superfici compenstrate.

Sotto il nome di scaling multidimensionale vi sono le tecniche statistiche che partendo da una matrice di distanza (dissimilarità metrica) o prossimità (dissimilarità non metrica) tra n unità, consentono di ottenere una rappresentazione geometrica (quindi una soluzione metrica) delle n unità analizzate in un numero limitato di dimensioni (tipicamente due o tre). L'obiettivo di queste tecniche è quello di ridurre la dimensione dello spazio delle osservazioni. Si parte da una matrice di n osservazioni ciascuna di p caratteristiche e si cerca una nuova matrice $n \times s$ con $s < p$. Il punto di partenza di queste tecniche è sempre la matrice delle distanze o similarità \hat{D} : questa matrice può essere calcolata a partire da una matrice \hat{X} di dati osservati, costituiti da n oggetti di dimensione K .

Metodi di scaling multidimensionale di tipo metrico

Quando si suppone che la matrice \hat{D} sia composta da elementi di tipo metrico, ossia espressi da scale di misura costituite da rapporti o da intervalli, allora lo spazio s -dimensionale di dimensioni ridotte rispetto al numero di caratteristiche (componenti) p sarà uno spazio delle proiezioni, o meglio delle distanze ottenute per proiezione nel sottospazio di dimensione inferiore. Una volta note le coordinate di n punti, di p dimensioni ciascuno, è immediato il calcolo della distanza euclidea tra ogni copia di oggetti. Possiamo partire dalla matrice delle osservazioni \hat{X} o dalla matrice dei prodotti scalari $\hat{P}_{mn} = \hat{X}\hat{X}^T$ il cui elemento generico

$$p_{ij} = \sum_{k=1}^p x_{ik} x_{jk}, \forall i \neq j$$

Il generico elemento delle matrici \hat{D} sarà espresso come segue

$$d_{ij} = \sqrt{\left(\sum_{k=1}^p x_{ik} - x_{jk} \right)^2} = \sqrt{(p_{ii} + p_{jj} - 2 \cdot p_{ij})}$$

Il che equivale a scrivere¹

$$d_{ij}^2 = (p_{ii} + p_{jj} - 2 \cdot p_{ij})$$

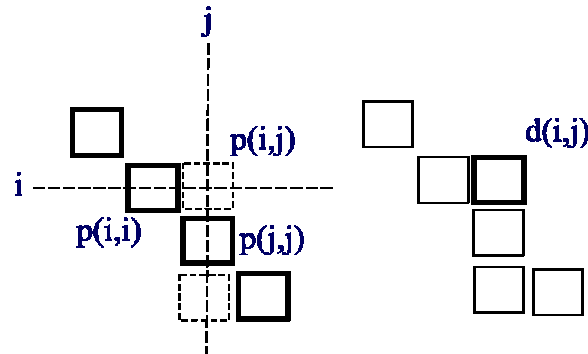


Figura 1 - Dalla matrice dei prodotti scalari alla matrice delle distanze.

Dalla \hat{D}_{nm} si passa alla matrice di rappresentazione \hat{R}_{ns} mediante proiezione di tutti i valori rispetto alle s dimensioni (due o tre) dello spazio ridotto di rappresentazione.

Metodi di scaling multidimensionale di tipo non-metrico

Se i dati della matrice \hat{D} sono ordinali o se si vuole considerare tali dati qualitativi, si utilizza una metodica di scaling non-metrico. L'idea di base di questi metodi è che sia possibile, a partire da \hat{D} , di dimensione $n \times n$, ricavare una matrice trasformata $\hat{\Delta}$ a partire dai dati di arrivo, ovvero dalla matrice di rappresentazione \hat{R}_{ns} , col vincolo che tali dati siano quanto più prossimi a quelli della matrice di partenza \hat{D} . In pratica si parte da una configurazione casuale dei dati nello spazio di arrivo di dimensione s \hat{R}_{ns} e con una procedura iterativa si ricavano i dati trasformati $\hat{\Delta}_{nm}$ che approssimano in qualche modo i dati di partenza \hat{D}_{nm} . Utilizzando una metrica euclidea, dalla matrice di rappresentazione \hat{R}_{ns} sarà sempre possibile calcolare una matrice di distanze $\hat{\Delta}_{nm}$ il cui generico elemento sia, ad esempio,:

$$\delta_{ij} = \sqrt{\sum_{K=1}^s (y_{ij} - y_{jk})^2}$$

Per determinare la configurazione finale si impone il vincolo di debole monotonicità, ossia che l'ordinamento delle distanze generate (δ_{ij}) sia quanto più possibile quello delle distanze delle matrici delle distanze di partenza (d_{ij}):

$$d_{ij} < d_{lm} \rightarrow \delta_{ij} \leq \delta_{lm}, \quad \forall i \neq j, l \neq m \in i, j, l, m \in [1, n].$$

In genere la procedura iterativa si fermerà quando un certo coefficiente di adattamento (che chiameremo *stress*) avrà raggiunto dei valori accettabili. Esistono più procedure di questo tipo, tra le quali quelle proposte da Shepard (1962) e da Kruskal (1964) sono tra le più note. L'indice di stress può essere calcolato come segue:

¹ Se si volesse ricavare la matrice \hat{P} dalla matrice delle distanze \hat{D} si veda [2].

$$Stress = \sqrt{\frac{\sum_{i < j} \sum_{j=2}^n (\delta_{ij} - f_{ij})^2}{\sum_{i < j} \sum_{j=2}^n \delta_{ij}^2}}$$

dove, δ_{ij} sono gli elementi di $\hat{\Delta}_{nm}$ e f_{ij} sono i valori monotoni trasformati a partire da d_{ij} . Se esiste una trasformazione strettamente monotona tra d_{ij} e δ_{ij} il valore dello Stress risulterà nullo, valori inferiori o uguali a 0,05 sono considerati di “buon adattamento”, quelli superiori a 0,2 sono considerati di “adattamento scadente”. Se l’adattamento viene considerato insoddisfacente si procede al lancio di una nuova procedura iterativa aumentando il numero di dimensioni s dello spazio di proiezione.

Cenni sui metodi di scaling multidimensionale classico in SPSS

L’analisi classica viene eseguita su dati posti in una sola matrice, tramite calcolo delle distanze euclidea tra gli elementi e mediante una tecnica di calcolo delle componenti principali. Può essere sia metrico che non metrico (*non-metric CMDS*) nel qual caso le misure sono ordinali. Con più matrici si parla di *Replicated CMDS* (RCMDS), supponendo lo stimolo (*perceptual map*) uguale per ciascuna matrice (ciascun oggetto dell’analisi). Gli output sono simili ma nel caso non metrico i calcoli sono maggiori in quanto si utilizzano algoritmi iterativi che ricalcolano le distanze tra gli oggetti preservando il rapporto tra le distanze (scala). Un’altra estensione del metodo classico CMDS è il metodo INDSCAL (*Individual differences Euclidean distance*) noto anche come MDS pesato (WMDS, *Weighted MDS*), applicabile su matrici multiple ma senza considerare lo stimolo uguale per tutte le matrici. Ciascun oggetto contribuisce in modo differente alla matrice in base ai pesi. E’ presente anche il modello GEMSCAL (*Generalized Euclidean Metric individual differences model*) disponibile tramite *script*:

```
ALSCAL VARIABLES = V1 TO Vn
/SHAPE = ASYMMETRIC
/CONDITION = ROW
/MODEL = GEMSCAL
/CRITERIA = DIM(4) DIRECTIONS(4)
```

Syntax 1 - SPSS GEMSCAL (disponibile in SPSS tramite SYNTAX).

- **ASCAL (Asymmetric Euclidean distance model)**
Il modello ASCAL (*Asymmetric Euclidean distance model*) utilizza una matrice asimmetrica con più di due dimensioni. Il modello AINDS (*Asymmetric Individual Differences Euclidean distance model*) più matrici asimmetriche con più dimensioni. Tutti questi modelli sono quelli disponibili nel modulo di analisi MDS ALSCAL di SPSS [23].
- **PROXSCAL (Proximity multidimensional scaling model)**
Questa procedura permette di realizzare l’analisi multidimensionale di dati di prossimità derivando una rappresentazione tramite I minimi quadrati delle distanze degli oggetti in uno spazio di dimensioni ridotte rispetto alle dimensioni degli oggetti in input. Un algoritmo di ottimizzazione garantisce la convergenza monotona di dati metrici e non-metrici sotto svariati modelli e vincoli².

² Una trattazione matematica dettagliata di tutti i modelli la si può trovare nel sito di SPSS:
<http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/index.html>.

Esempio numerico: analisi MDS dei dati di input dell'analisi ADL

Nel paragrafo seguente, utilizzando i dati metrici di input dell'analisi discriminante lineare, e proveremo a dimostrare che la classificazione realizzata nel Capitolo 2 coincide con il posizionamento realizzato attraverso l'analisi multidimensionale degli input trasformati in distanze per un utilizzo diretto nell'analisi delle dissimilarità. In pratica utilizzeremo n=25 vettori tridimensionali (K gruppi, con K=3) trasformati in una matrice simmetrica n×n, con diagonale non nulla, i cui elementi sono i prodotti scalari di ogni vettore tridimensionale moltiplicato per ogni altro dell'insieme completo. Questa matrice può essere trasformata in una matrice delle distanze (n×n, con gli elementi della diagonale uguali a zero), come indicato nel paragrafo "Metodi di scaling multidimensionale di tipo metrico" ed utilizzata come input per l'analisi MDS delle dissimilarità con SPSS.

In pratica, si sono usati i dati (25 elementi tridimensionali) :

6,00	1,16	0,11	1,48	0,63	0,89	0,28	0,12	-0,39	0,03	-1,08	0,20	-0,74	-1,76	-1,08	-0,91	-0,74	-0,22	-1,68	-1,76	1,14	0,45	0,71	1,05	0,62
0,27	0,12	1,49	0,05	0,91	1,27	0,99	0,34	0,41	0,41	0,27	0,05	0,12	0,70	0,34	0,63	-0,38	0,27	0,20	0,41	-1,31	-1,72	-2,32	-1,96	-1,53
1,00	1,00	1,00	1,00	1,00	1,00	1,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	3,00	3,00	3,00	3,00	3,00

e si è calcolata la matrice simmetrica dei prodotti scalari che è stata poi utilizzata per un'analisi PROXSCAL con SPSS: con dati di input da trasformare in dati di prossimità a cura della procedura SPSS. A partire da questa matrice dei prodotti scalari si è calcolata la matrice delle distanze che è stata utilizzata per l'analisi multidimensionale MDS ALSCAL e per l'analisi PROXSCAL, questa volta con matrici delle distanze in input³. Si sono confrontati i risultati così ottenuti tra loro e con i risultati dell'analisi discriminante lineare del Capitolo 2, verificando la coincidenza del posizionamento metrico MDS con la classificazione metrica ADL.

6,00	2,75	1,56	3,20	2,17	2,66	1,68	2,26	1,52	2,16	0,47	2,31	0,94	-0,43	0,49	0,82	0,80	1,74	-0,44	-0,50	4,34	3,21	3,43	4,04	3,52	
	2,35	1,31	2,71	1,84	2,18	1,45	2,18	1,59	2,09	0,79	2,24	1,16	0,05	0,79	1,03	1,10	1,77	0,09	0,01	4,16	3,32	3,54	3,98	3,54	
		3,22	1,24	2,43	2,99	2,50	2,52	2,57	2,61	2,28	2,10	2,10	2,84	2,38	2,83	1,35	2,37	2,10	2,41	1,18	0,49	-0,37	0,20	0,79	
			3,18	1,97	2,37	1,47	2,19	1,44	2,07	0,42	2,30	0,92	-0,57	0,42	0,69	0,89	1,68	-0,47	-0,58	4,61	3,59	3,94	4,46	3,85	
				2,22	2,72	2,08	2,38	2,13	2,39	1,57	2,17	1,65	1,53	1,63	2,00	1,19	2,10	1,13	1,27	2,52	1,71	1,33	1,87	1,99	
					3,41	2,51	2,54	2,17	2,55	1,39	2,24	1,50	1,33	1,48	2,00	0,86	2,14	0,76	0,96	2,34	1,20	0,67	1,43	1,60	
						2,06	2,37	2,29	2,42	1,96	2,11	1,91	2,19	2,03	2,36	1,41	2,20	1,72	1,91	2,02	1,42	0,90	1,36	1,66	
							4,13	4,09	4,14	3,96	4,04	3,96	4,03	3,99	4,11	3,78	4,06	3,87	3,93	5,69	5,47	5,30	5,46	5,55	
								4,32	4,16	4,54	3,94	4,34	4,98	4,56	4,62	4,13	4,20	4,74	4,86	5,01	5,11	4,76	4,78	5,12	
									4,17	5,23	4,03	4,03	4,23	4,10	4,23	3,82	4,10	4,03	4,11	5,50	5,31	5,07	5,23	5,39	
										5,23	3,80	4,83	6,09	5,25	5,15	4,69	4,31	5,86	6,01	4,42	5,05	4,61	4,34	4,92	
											4,04	3,86	3,68	3,80	3,85	3,83	3,97	3,67	3,66	6,16	6,00	6,02	6,11	6,05	
												4,56	5,38	4,84	4,75	4,50	4,20	5,26	5,35	5,00	5,45	5,19	4,98	5,35	
													7,59	6,14	6,04	5,03	4,58	7,09	7,39	3,08	3,99	3,12	2,77	3,83	
														5,28	5,19	4,67	4,33	5,87	6,04	4,33	4,93	4,45	4,20	4,81	
															5,22	4,43	4,37	5,64	5,86	4,15	4,51	3,90	3,81	4,47	
																4,69	4,06	5,16	5,14	5,66	6,32	6,36	5,97	6,12	
																	4,12	4,43	4,50	5,39	5,44	5,22	5,24	5,45	
																		6,85	7,03	3,84	4,90	4,36	3,85	4,65	
																			7,27	3,46	4,49	3,79	3,34	4,27	
																					12,01	11,78	12,86	12,77	11,72
																						12,18	13,33	12,87	11,93
																							14,91	14,31	13,01
																								13,97	12,67
																									11,74

Tabella 1 - Dati di ingresso della procedura di Multidimensional Scaling (MDS) PROXSCAL con SPSS (tali dati non sono dati di prossimità e dovranno essere preventivamente trasformati da SPSS in dati di prossimità prima dell'analisi di posizionamento metrico): la matrice è di dimensione 25×25, simmetrica (sono stati omessi i dati della parte triangolare).

³ Si veda il file PW_PARMA_ENNAS_MDS_distanze_italiana.xls, parte integrante di questo report.

0,00	0,13	3,37	0,05	1,15	1,37	1,96	2,87	4,55	3,13	7,56	2,69	5,95	11,72	7,57	6,85	6,35	3,92	10,99	11,55	6,61	9,03	11,31	9,17	7,98
	0,00	2,95	0,11	0,91	1,40	1,52	2,13	3,49	2,35	6,01	1,91	4,58	9,85	6,04	5,52	4,84	2,93	9,03	9,60	6,05	7,90	10,18	8,36	7,02
		0,00	3,93	0,60	0,64	0,28	2,32	2,41	2,16	3,91	3,07	3,58	5,13	3,74	2,78	5,21	2,60	5,87	5,67	12,88	14,43	18,88	16,79	13,38
			0,00	1,47	1,85	2,31	2,93	4,63	3,22	7,57	2,62	5,90	11,91	7,61	7,02	6,08	3,94	10,96	11,61	5,97	8,19	10,22	8,23	7,23
				0,00	0,44	1,57	2,51	1,24	5,01	0,94	3,79	6,43	3,65	2,70	4,96	0,70	7,46	6,11	8,31	12,11	14,78	11,78	9,97	9,08
					0,00	1,45	2,10	1,63	3,73	1,79	2,84	5,12	3,41	2,52	4,22	1,13	5,61	5,18	9,87	11,36	14,52	12,16	9,77	6,09
						0,00	0,27	-0,02	1,45	-0,07	0,69	3,49	1,33	0,99	1,23	-0,22	3,17	3,28	4,57	6,93	9,78	8,35	6,09	6,27
							0,00	0,18	0,32	0,70	-0,18	1,97	1,13	0,55	1,04	-0,14	1,56	2,29	6,85	6,97	10,49	9,17	6,27	6,19
								0,00	-1,06	1,22	0,53	3,16	1,30	0,86	1,28	-0,26	2,90	3,07	5,13	7,07	10,08	8,58	6,19	6,25
									0,00	1,69	-1,30	0,25	0,86	0,18	0,45	0,19	-0,56	1,11	9,17	6,50	10,73	9,90	6,25	5,76
										0,00	0,89	4,09	1,36	1,32	0,87	0,02	3,47	3,61	3,35	6,34	8,86	7,77	5,76	6,02
											0,00	1,38	0,99	0,56	0,44	0,22	0,53	1,83	7,36	6,28	9,98	9,19	6,02	6,86
												0,00	0,60	-0,72	0,66	-0,01	-2,75	-0,42	13,24	7,27	12,65	11,54	6,86	6,33
													0,00	0,11	0,55	0,12	-0,57	1,07	9,39	6,66	10,94	10,01	6,33	6,61
														0,00	1,05	-0,19	-0,07	1,20	9,58	7,31	11,62	10,24	6,61	5,47
															0,00	0,68	0,59	2,15	5,83	5,20	8,50	8,38	5,47	6,08
																0,00	2,11	2,69	5,73	6,70	9,90	8,71	6,08	6,29
																	0,00	0,05	11,37	6,21	11,10	10,61	6,29	6,54
																		0,00	12,37	6,66	11,81	11,08	6,54	1,07
																			0,00	0,64	0,97	1,28	1,07	0,75
																				0,00	0,43	1,56	0,75	0,04
																					0,00	0,25	0,04	-0,31
																						0,00	-0,31	0,00

Tabella 2 - Matrice delle distanze (matrice di prossimità), costruita a partire dalla matrice dei prodotti scalari. Questa matrice viene utilizzata come input del calcolo MDS PROXCAL con dati di prossimità e al calcolo ALSCAL per il posizionamento multidimensionale.

In pratica in Tabella 1 sono quelli della matrice \hat{P} dei prodotti scalari e quelli della Tabella 2 sono quelli delle distanze ricavate dai prodotti scalari utilizzando la $d_{ij}^2 = (p_{ii} + p_{jj} - 2 \cdot p_{ij})$.

Il risultato ottenuto per ognuna delle analisi di dissimilarità effettuata è una mappa di posizionamento dei campioni che riflette la classificazione ottenuta con la discriminazione lineare ottenuta col metodo bayesiano nel Capitolo 2 del presente lavoro [5].

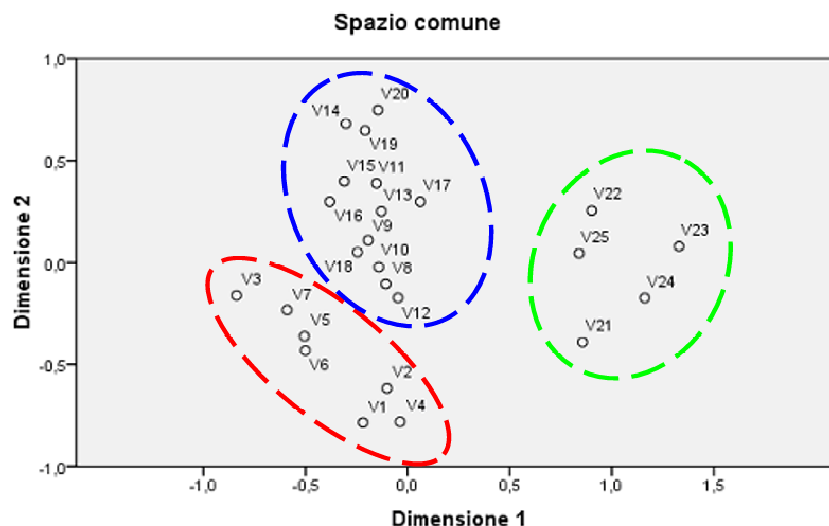


Figura 2 - MDS PROXCAL: i dati in ingresso sono dati di prossimità (distanze).

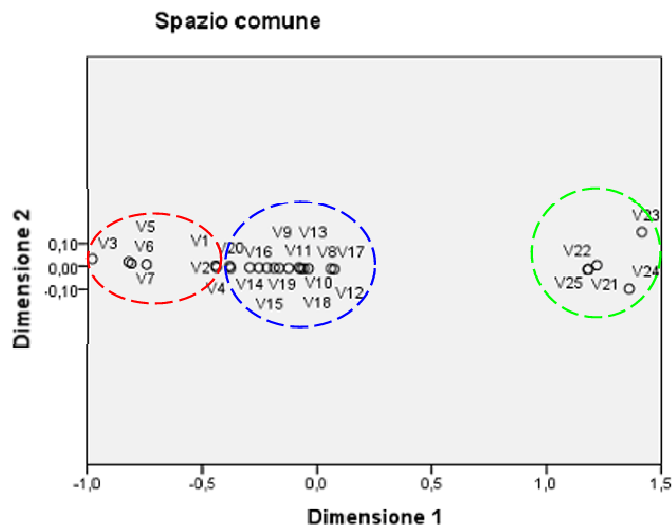


Figura 3 - Analisi MDS PROXCAL (i dati in ingresso non sono dati di prossimità, SPSS provvederà a trasformarli in distanze prima dell'analisi).

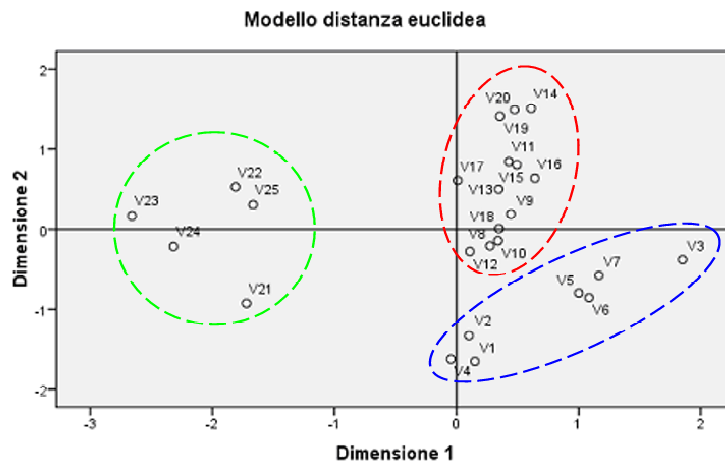


Figura 4 - Analisi multidimensionale MDS ALSCAL (i dati di input sono dati di prossimità).

Dall'analisi comparativa delle mappe di posizionamento possiamo notare che la Figura 3, ricavata ponendo in input alla procedura PROXSCAL di SPSS dei prodotti scalari non trasformati e richiedendo alla procedura stessa di trasformarli in distanze, utilizza un algoritmo di trasformazione differente rispetto alla semplice trasformazione indicata nel paragrafo "Metodi di scaling multidimensionale di tipo metrico".

Multidimensional Scaling con SPSS ALSCAL: distanze tra città

In questo paragrafo riformuleremo un esempio metrico classico del multidimensional scaling, la mappa delle distanze delle città (nel nostro caso italiane). Le distanze utilizzate sono distanze in linea d'aria.

	AN	AO	BA	BO	CB	FI	GEN	AQ	MI	NA	PG	PZ	RC	RM	TOR	UD	VE
AN	0																
AO	609	0															
BA	466	1063	0														
BO	572	536	1026	0													
CB	764	1292	348	923	0												
FI	260	471	662	101	472	0											
GEN	188	785	407	393	702	230	0										
AQ	423	783	877	206	736	315	555	0									
MI	257	354	711	206	570	295	253	433	0								
NA	401	940	255	570	137	468	707	238	764	0							
PG	143	622	567	253	377	151	389	175	447	373	0						
PZ	464	1060	140	668	201	607	846	377	875	158	512	0					
RC	855	1383	439	1014	542	912	1150	681	1208	462	817	390	0				
RM	308	740	412	370	222	268	507	122	564	219	173	358	662	0			
TOR	544	113	998	327	857	400	170	720	138	869	551	995	1312	669	0		
UD	412	556	866	265	725	365	509	588	383	834	439	863	1278	634	512	0	
VE	300	449	754	158	613	258	402	476	276	727	327	752	1171	527	405	127	0

Tabella 3 - Distanze tra le città italiane (matrice simmetrica).

GET

FILE='C:\Documents and Settings\user\Desktop\MDS\MDS_città_italiane_alscal.sav'.

DATASET NAME DataSet6 WINDOW=FRONT.

ALSCAL

VARIABLES=AN AO BA BO CB FI GEN AQ MI NA PG PZ RC RM TOR UD VE

/SHAPE=SYMMETRIC

/LEVEL=ORDINAL

/CONDITION=MATRIX

/MODEL=EUCLID

/CRITERIA=CONVERGE(0.001) STRESSMIN(0.005) ITER(30) CUTOFF(0) DIMENS(2,2)

/PLOT=DEFAULT ALL

/PRINT=DATA HEADER.

Syntax 2 - SPSS MDS ALSCAL.

Vediamo alcune definizioni preliminari utilizzate nel seguito.

S-Stress

SPSS utilizza la minimizzazione dell'S-Stress di Young come criterio per fermare le iterazioni nel processo di ricerca della soluzione. Poiché è noto che questo procedimento porta a soluzioni sub-ottime⁴, attualmente alla procedura ALSCAL sono preferite le procedure PROXSCAL e PREFSCAL (Multidimensional Unfolding).

⁴ L'S-Stress è una "funzione di perdita" che fornisce un peso tanto maggiore quanto maggiori sono le dissimilarità: ciò è associato ad un elevato errore (Ramsey, 1988) che porta a soluzioni sub-ottime.

Stress

E' una misura della bontà dell'adattamento (*fit*) per il modello di Multidimensional Scaling (MDS) [1].

Più piccolo è lo stress migliore l'accordo tra dati e modello. Tale stress misura le differenze delle distanze tra i punti nello spazio di calcolo di MDS e le corrispondenti distanze tra i dati in input e i risultati elaborati dall' algoritmo. Lo stress misura l'errore ma da anche indicazioni sull'utilizzo scorretto del numero di dimensioni dello spazio di calcolo (da 1 a 6) e rappresentazione. Lo stress di *Young S-Stress* dipende dalle distanze al quadrato, mentre lo stress di *Kruskal* dalle distanze. SPSS li calcola entrambi ma usa l'*S-Stress* per fermare le iterazioni. Si indica col termine "Stress non-normalizzato" la somma dei quadrati dei residui da una regressione monotona (scarti delle distanze dai valori di una regressione monotona di riferimento):

$$SNN = \sum_{i,j} (d_{ij} - d_{ij}^m)^2 ,$$

possiamo definire fattori di normalizzazione FN le somme:

$$FN1 = \sum_{i,j} (d_{ij})^2 , \text{ somma dei quadrati delle distanze}$$

$$FN2 = \sum_{i,j} (d_{ij} - \bar{d})^2 , \text{ some del quadrato delle deviazioni dalla distanza media}$$

e i relativi stress:

$$S_1 = \frac{\sum_{i,j} (d_{ij} - d_{ij}^m)^2}{\sum_{i,j} (d_{ij})^2} = \sqrt{\frac{SNN}{FN1}} \quad \text{e} \quad S_2 = \frac{\sum_{i,j} (d_{ij} - d_{ij}^m)^2}{\sum_{i,j} (d_{ij} - \bar{d})^2} = \sqrt{\frac{SNN}{FN2}}$$

Le due formule riportate sono le espressioni di *Kruskal* (1 e 2) dello *stress*.

Scree Plot

E' un modo grafico alternativo per determinare il numero ottimale di dimensioni, utilizzandolo nello stesso modo col quale viene utilizzato per l'analisi fattoriale e l'analisi delle componenti principali per determinare i fattori da estrarre e l'entità delle componenti principali rispettivamente tramite la rappresentazione grafica dei moduli degli autovalori della matrice (Figura 5).

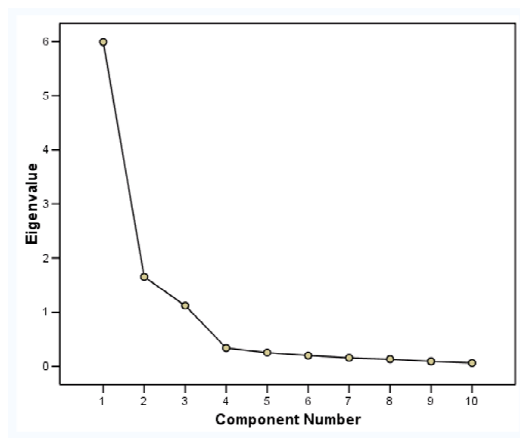


Figura 5 - Scree plot.

Iteration history for the 2 dimensional solution (in squared distances)

Young's S-stress formula 1 is used.

Iteration	S-stress	Improvement
1	,18845	
2	,15639	,03205
3	,14940	,00699
4	,14702	,00238
5	,14602	,00099

Iterations stopped because
S-stress improvement is less than ,001000

Tabella 4 - Valori dello stress e del miglioramento nel calcolo iterativo dell'S-Stress.

Il valore dell'S-Stress viene utilizzato per fermare l'iterazione dell'algoritmo MDS. L'indice di correlazione (RSQ) è una misura di adattamento di uso comune. Con $RSQ \geq 0.6$ l'adattamento si considera accettabile [1]. RQS è la correlazione al quadrato delle distanze in input, con lo spazio di dimensione p scalato utilizzando le coordinate calcolate tramite MDS. RSQ riflette la porzione di varianza della matrice delle distanze in input che viene spiegata dai dati scalati su $s \leq p$ dimensioni. Ciò significa che aumentando il numero di dimensioni di rappresentazione dobbiamo aspettarci un approssimarsi di RSQ all'unità.

Stress and squared correlation (RSQ) in distances

RSQ values are the proportion of variance of the scaled data (disparities) in the partition (row, matrix, or entire data) which is accounted for by their corresponding distances.
Stress values are Kruskal's stress formula 1.

For matrix
Stress = ,14422 RSQ = ,89920

Configuration derived in 2 dimensions

Tabella 5 - Correlazione quadratica e Stress.

Le coordinate degli stimoli (ossia di ogni oggetto osservato) indicano le coordinate di tali oggetti nel nuovo spazio scalato e quindi mettono in relazione ogni stimolo con le dimensioni dello spazio di rappresentazione. Il posizionamento nella mappa bidimensionale è la rappresentazione grafica di questi dati. Nella mappa MDS, nota come mappa delle percezioni (configurazione degli stimoli derivati, Figura 6) l'orientamento è arbitrario. Ciò su cui ci si deve concentrare in questa mappa è la possibile formazione di *cluster* di punti omogenei in qualche senso, nel caso specifico relazioni geografiche di vicinanza (ad esempio Nord, Centro e Sud).

Stimulus Coordinates			
Dimension			
Stimulus Number	Stimulus Name	1	2
1	AN	,0367	-,6053
2	AO	2,1310	,5451
3	BA	-1,4953	-1,0207
4	BO	,9479	,8689
5	CB	-1,4197	-,5911
6	FI	,3835	,1721
7	GEN	,6416	-,9509
8	AQ	-,0836	1,1198
9	MI	1,0292	-,4441
10	NA	-1,2149	,3137
11	PG	-,0736	,3200
12	PZ	-1,5586	,3424
13	RC	-2,5838	,7320
14	RM	-,4694	,4565
15	TOR	1,6995	-,2597
16	UD	1,1465	-,6867
17	VE	,8830	-,312

Tabella 6 - Coordinate degli stimoli dalle quali si ricava la mappa delle percezioni.

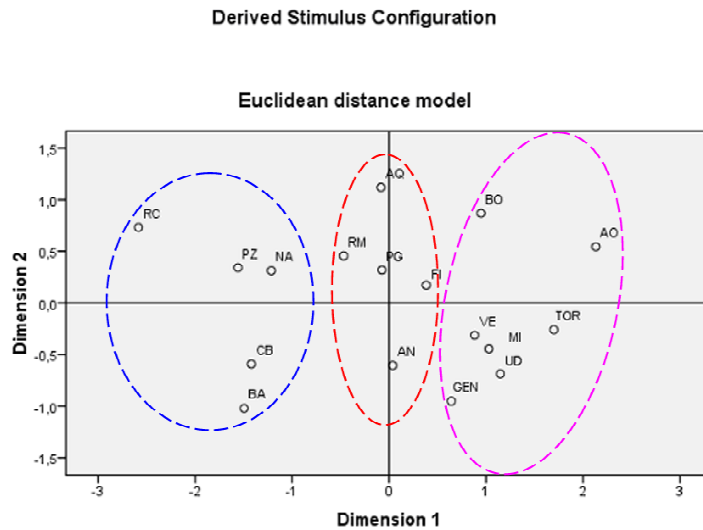


Figura 6 - Mappa delle percezioni.

In Figura 7 è rappresentato il *diagramma di Shepard* che mostra le cosiddette disparità (ossia le distanze in ingresso, trasformate nello spazio di dimensione p dal MDS). In pratica i dati d'ingresso (distanze \hat{D}) vengono trasformati conservando le differenze tra due punti nello spazio trasformato di dimensione $p < n$. Se il modello potesse mappare perfettamente i dati (massimo adattamento) allora le distanze e le disparità sarebbero coincidenti e il *diagramma di Shepard* giacerebbe sulla retta bisettrice del I quadrante ($\theta=45^\circ$) nel caso metrico in esame [1].

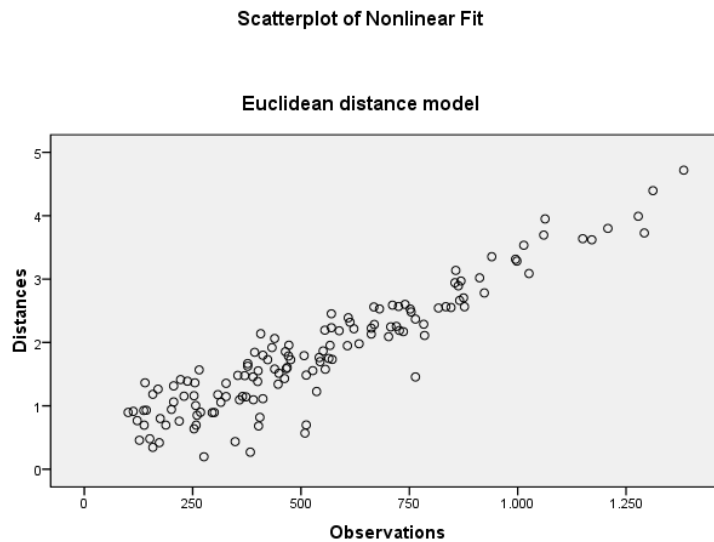


Figura 7 - Diagramma di Shepard (caso metrico).

Nel caso non-metrico l'adattabilità migliore del modello ai dati produrrebbe una funzione lineare a gradino (*step-line*) corrispondente alla funzione monotona utilizzata per realizzare la trasformazione: deviazioni da questo andamento verrebbero considerate come errori nell'adattamento (*lack of fit*). Nel caso in esame si nota che il modello si adatta bene per valori distanza/disparità > 2 , ma non tanto bene al di sotto di questo valore. Se il modello è non-metrico si fa riferimento alla Figura 8 e le misure sono ordinali. Il confronto sarà tra dati osservati (dai più piccoli ai più grandi) e distanze. Un buon adattamento del modello ai dati è *omoschedastico*⁵, inoltre i punti saranno vicini alla retta per valori alti e bassi delle due dimensioni considerate.

⁵ Si dice di un processo numerico la cui varianza non cambia tra diverse osservazioni campionarie.

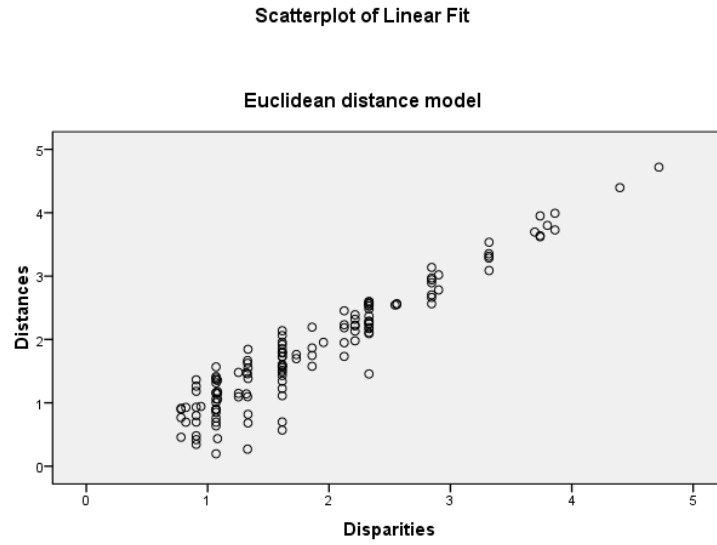


Figura 8 - Diagramma di Shepard 1 (caso non-metrico).

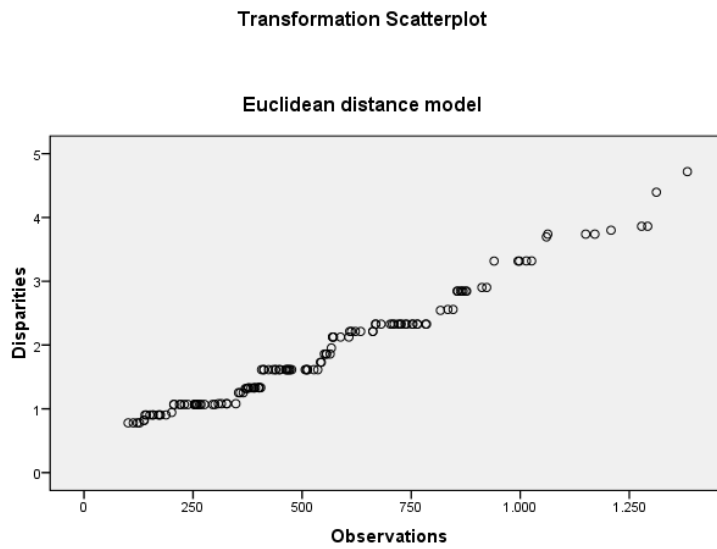


Figura 9 - Diagramma di Shepard 2 (caso non-metrico).

Multidimensional Scaling con SPSS PROXSCAL: distanza tra città

```

GET DATA
  /TYPE=XLS
  /FILE='C:\Documents and Settings\user\MDS\MDS_distanze_italiana.xls'
  /SHEET=name 'distanze'
  /CELLRANGE=full
  /READNAMES=on
  /ASSUMEDSTRWIDTH=32767.
DATASET NAME DataSet3 WINDOW=FRONT.
SAVE OUTFILE='C:\Documents and '+Settings\user\Desktop\MDS_città_italiane_proxcal.sav'
  /COMPRESSED.
PROXSCAL VARIABLES=AN AO BA BO CB FI GEN AQ MI NA PG PZ RC RM TOR UD VE
  /SHAPE=LOWER
  /INITIAL=SIMPLEX
  /TRANSFORMATION=INTERVAL
  /PROXIMITIES=DISSIMILARITIES
  /ACCELERATION=NONE
  /CRITERIA=DIMENSIONS(2,3) MAXITER(100) DIFFSTRESS(.0001) MINSTRESS(.0001)
  /PRINT=COMMON DISTANCES TRANSFORMATIONS INPUT HISTORY STRESS DECOMPOSITION
  /PLOT=STRESS COMMON TRANSFORMATIONS(ALL) RESIDUALS(ALL).
    
```

Syntax 3 - SPSS MDS PROXCAL.

Dimensionalità:2

Iterazione	Stress non trasformato normalizzato	Miglioramento
0	,12383 ^a	
1	,05971	,06412
2	,04241	,01730
3	,03412	,00829
4	,03093	,00319
5	,02943	,00150
6	,02856	,00087
7	,02796	,00060
8	,02751	,00045
9	,02715	,00035
10	,02688	,00027
11	,02667	,00021
12	,02650	,00017
13	,02636	,00014
14	,02625	,00011
15	,02615	,00010 ^b

a. Stress di configurazione iniziale: inizio semplice.

b. Il processo di iterazione è stato interrotto perché il miglioramento è inferiore al criterio di convergenza.

Tabella 7 - Cronologia delle iterazioni.

La cronologia delle iterazioni viene utilizzata per verificare la convergenza dell'algoritmo e per verificare il valore di start delle iterazioni.

Dimensionalità:2

Stress non trasformato normalizzato	,02615
Stress-I	,16171 ^a
Stress-II	,35385 ^a
S-stress	,04726 ^b
Dispersione spiegata	,97385
Coefficiente di convergenza di Tucker	,98684

PROXSCAL minimizza lo stress non trasformato normalizzato.

a. Livello di scala ottimale = 1,027.

b. Livello di scala ottimale = ,990.

Tabella 8 - Misure di stress e adattamento.

Dimensionalità:2		Sorgente	
		SRC_1	Media
Oggetto	AN	,0312	,0312
	AO	,0340	,0340
	BA	,0529	,0529
	BO	,0414	,0414
	CB	,0538	,0538
	FI	,0066	,0066
	GEN	,0714	,0714
	AQ	,0258	,0258
	MI	,0203	,0203
	NA	,0086	,0086
	PG	,0085	,0085
	PZ	,0080	,0080
	RC	,0035	,0035
	RM	,0073	,0073
	TOR	,0088	,0088
	UD	,0313	,0313
	VE	,0313	,0313
	Media	,0262	,0262

Tabella 9 - Decomposizione dello stress non trasformato normalizzato.

La Tabella 8 indica gli stress utilizzati per la verifica della bontà dell'adattamento. La procedura PROXSCAL utilizza un numero ampio di indicatori della bontà dell'adattamento (maggiori che in ALSICAL). Infatti utilizzano anche la dispersione spiegata (DAF, *Dispersion Accounted Factor*) ed il coefficiente di congruenza di Tucker (TCC, *Tucker Congruence Coefficient*) [2] entrambi coefficienti di bontà dell'adattamento: risultano prossimi all'unità nei casi di elevato adattamento. La Tabella 9 di decomposizione dello stress indica, stimolo per stimolo, quali oggetti contribuiscono a in che misura alla composizione dello stress.

	Dimensione	
	1	2
AN	-,066	-,257
AO	-,950	,458
BA	,790	-,298
BO	-,372	,303
CB	,615	-,419
FI	-,196	,091
GEN	-,428	-,183
AQ	,115	,436
MI	-,577	,034
NA	,582	,061
PG	,027	,001
PZ	,770	,016
RC	1,219	,416
RM	,267	,132
TOR	-,834	,134
UD	-,534	-,558
VE	-,427	-,368

Tabella 10 - Coordinate finali.

Le coordinate finali sono la soluzione calcolata nel caso bidimensionale. Esse differiscono notevolmente da quelle calcolate con ALSICAL, ma una comparazione non è possibile perché l'orientamento della mappa risultante non è lo stesso.

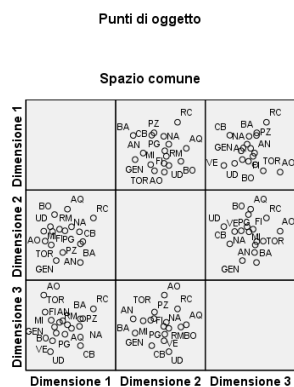


Figura 10 - Rappresentazione in tre dimensioni.

Grafico dei residui

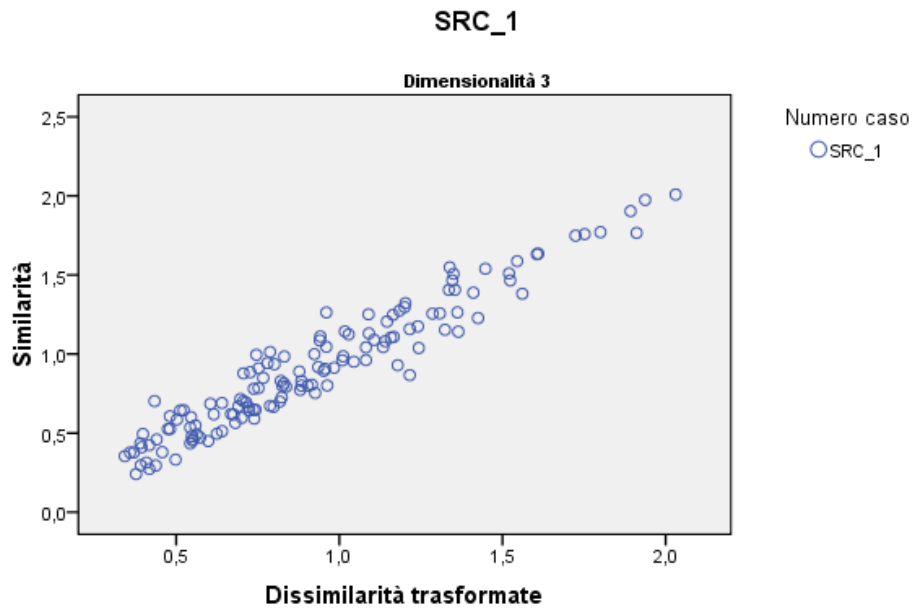


Grafico dei residui

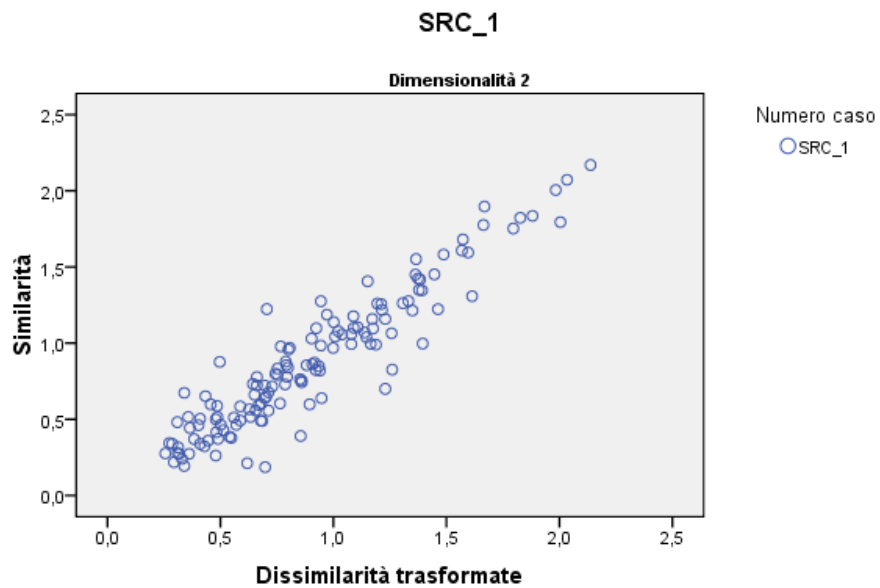


Figura 11 - Grafici dei residui (distanze vs. dissimilarità trasformate ricavate dal modello).

Glossario

Analisi statistica multivariata

Con statistica multivariata s'intende quella parte della statistica in cui l'oggetto dell'analisi è per sua natura formato da almeno due componenti, il che è spesso il caso nell'ambito di scienze quali la medicina, psicologia, sociologia, ecologia e biologia. Fanno parte della statistica multivariata metodi quali:

- analisi della correlazione canonica e analisi delle componenti principali
- analisi fattoriale
- analisi delle corrispondenze
- analisi dei cluster
- analisi discriminante
- analisi di regressione multidimensionale

Correlazione

Per correlazione si intende una relazione tra due variabili casuali tale che a ciascun valore della prima variabile corrisponda con una certa regolarità un valore della seconda. Non si tratta necessariamente di un rapporto di causa ed effetto ma semplicemente della tendenza di una variabile a variare in funzione di un'altra. Talvolta le variazioni di una variabile dipendono dalle variazioni dell'altra (relazione tra la statura dei padri e quella dei figlio ad esempio), talvolta sono comuni (relazioni tra la statura e il peso di un individuo); talvolta sono reciprocamente dipendenti (relazione tra prezzo e domanda di una merce: il prezzo influisce sulla domanda e la domanda influisce sul prezzo). Il grado di correlazione fra due variabili viene espresso mediante i cosiddetti indici di correlazione. Questi assumono valori compresi tra meno uno (quando le variabili considerate sono inversamente correlate) e l'unità (quando vi sia correlazione assoluta cioè quando alla variazione di una variabile corrisponde una variazione rigidamente dipendente dall'altra), ovviamente un indice di correlazione pari a zero indica un'assenza di correlazione e quindi le variabili sono indipendenti l'una dall'altra. I coefficienti di correlazione sono derivati dagli indici di correlazione tenendo presenti le grandezze degli scostamenti dalla media. In particolare, il coefficiente di correlazione di Pearson è calcolato come rapporto tra la covarianza delle due variabili ed il prodotto delle loro deviazioni standard [9].

Covarianza/Varianza

È un indice che misura la "contemporaneità" della variazione (in termini lineari) di due variabili casuali. Essa può assumere sia valori positivi che negativi. Nel caso di valori positivi indica che al crescere di una caratteristica statisticamente cresce anche l'altra, nel caso di valori negativi accade il contrario. Nella statistica inferenziale, quando due variabili sono tra di loro indipendenti, allora la loro covarianza è nulla (l'inverso non è necessariamente verificato). Si utilizza spesso la notazione:

$$\text{cov}(x, y) = \sigma_{xy}$$

$$\sigma_{xy} = \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$$

essendo μ_x e μ_y rispettivamente la media aritmetica di x e y .

In caso di ponderazione,

$$\sigma_{xy} = \sum_{j=1}^k f_j (x_j - \mu_x)(y_j - \mu_y)$$

È un operatore simmetrico, cioè

$$\text{cov}(x, y) = \text{cov}(y, x)$$

La covarianza può essere scomposta in due termini, diventando

$$\sigma_{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \mu_x \mu_y$$

ovvero la media dei prodotti meno il prodotto delle medie.

Quando $y=x$, allora la covarianza si trasforma in varianza:

$$\sigma_{xx} = \text{cov}(x, x) = \text{var}(x) = \sigma_x^2.$$

Deviazione standard (*Standard Deviation*)

Indice di dispersione della popolazione o del campione. Detto anche *Scarto Quadratico Medio*.

Mappa/Mappatura delle percezioni (*Perceptual mapping*)

E' una tecnica grafica utilizzata nel marketing per tentare di visualizzare la percezione che hanno i potenziali clienti di un determinato prodotto in relazione ad altri prodotti di riferimento.

Marketing

E' un ramo dell'economia che si occupa dello studio descrittivo del mercato e dell'analisi dell'interazione del mercato, degli utilizzatori con l'impresa. Il termine prende origine dall'inglese market, cui viene aggiunta la desinenza del gerundio per indicare la partecipazione attiva, cioè l'azione sul mercato stesso. Marketing significa letteralmente "piazzare sul mercato" e comprende quindi tutte le azioni aziendali riferibili al mercato destinate al piazzamento di prodotti, considerando come finalità il maggiore profitto e come causalità la possibilità di avere prodotti capaci di realizzare tale operazione [14].

Marketing analitico

Consiste in un insieme di tecniche e metodologie volte ad analizzare con metodi quantitativi, il mercato nella sua accezione più larga (dei clienti finali, o degli intermediari, ecc.) per mappare i desideri del cliente, oppure i suoi comportamenti (segmentazione), e per conoscere gli ambiti di mercato già eventualmente occupati dai rivali diretti e indiretti (posizionamento).

Marketing strategico

Si basa sull'analisi dei bisogni degli individui e delle organizzazioni. Questo primo aspetto del processo di marketing riguarda anzitutto l'individuazione, all'interno del mercato di riferimento, dei prodotti-mercato e dei segmenti già esistenti o potenziali. Di questi il marketing strategico misura l'attrattività in termini quantitativi, qualitativi (con riferimento all'accessibilità al mercato) e dinamici (con riferimento alla durata economica che è rappresentata dal ciclo di vita del prodotto). Tali operazioni consentono di scegliere una strategia di sviluppo che colga le opportunità esistenti sul mercato (rappresentate sostanzialmente da bisogni insoddisfatti) e che, tenendo conto delle risorse e competenze dell'impresa, offrano alla stessa un potenziale di crescita e di redditività attraverso l'acquisizione ed il mantenimento di un vantaggio competitivo [13][14].

Marketing operativo

E' la parte applicativa dell'intero processo di marketing, a monte del quale ci sono le fasi di marketing analitico e marketing strategico. La componente operativa (o tattica) del marketing ha il compito di realizzare concretamente le strategie definite nelle fasi precedenti. le caratteristiche: orientamento all'azione, opportunità esistenti, ambiente stabile, comportamento reattivo, orizzonte a breve termine, responsabilità della funzione di marketing [13][14].

Marketing Mix

Indica la combinazione (*mix*) di variabili controllabili (leve decisionali) di marketing che le imprese impiegano per raggiungere i propri obiettivi. Le variabili che tradizionalmente si includono nel *marketing mix* sono le 4P teorizzate da Jerome McCarthy e riprese in seguito da molti altri: *Product* (Prodotto), *Price* (Prezzo), *Place* (Distribuzione), *Promotion* (Comunicazione) [13].

Percezione

Il complesso processo elettrochimico che connette i livelli sensoriali di un organismo attraverso il sistema nervoso e che opera la sintesi dei dati sensoriali in forme dotate di significato.

Posizionamento

Il posizionamento di un prodotto può essere visto come una decisione strettamente connesso a quella della selezione dei segmenti di mercato in cui l'impresa decide di competere. Il posizionamento consiste nella misura della percezione che hanno i clienti di un prodotto o di una merce, relativamente alla posizione dei prodotti o delle marche concorrenti.

Prossimità

E' una relazione di distanza di tipo qualitativo (non quantitativo). Essa implica le locuzioni "più vicino a rispetto a ...", oppure "più lontano da rispetto a ..." ed è in grado di costruire relazioni di distanza ordinate e non metriche (ossia non basate su una misura numerica della distanza effettiva).

R di Pearson (indice di correlazione)

Dividendo la covarianza con il prodotto delle deviazioni standard delle due variabili, si ottiene l'indice di correlazione di Pearson:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.$$

Nella conjoint analysis viene utilizzato, insieme al Tau di Kendall, per valutare le correlazioni tra le preferenze osservate e quelle stimate

Scaling multidimensionale

Le tecniche che, partendo da una matrice di distanza o di prossimità tra n unità, consentono di ottenere una soluzione metrica, ossia una rappresentazione geometrica delle n unità in un numero limitato di dimensioni.

Scree-plot

La rappresentazione grafica degli autovalori (in ordinata) e dell'ordine di estrazione dei fattori (in ascissa) dà un'immagine dell'importanza relativa dei primi autovalori nella sequenza ricavata. Si escludono quei i fattori i cui valori appartengono alla spezzata che corre quasi parallela all'asse delle ascisse, in quanto hanno tasso di variazione prossima allo zero (derivata nulla). La pendenza dei segmenti che congiungono due punti in corrispondenza di due generici autovalori danno indicazioni sulla varianza spiegata, perciò i gomiti sono più significativi delle zone piatte.

Segmentazione

Col termine "segmentazione" del mercato s'intende l'attività di identificazione di "gruppi di clienti" cui è indirizzato un determinato prodotto o servizio. La segmentazione è "il processo attraverso il quale le imprese suddividono la domanda in un insieme di clienti potenziali, in modo che gli individui che appartengono allo stesso insieme siano caratterizzati da funzioni della domanda il più possibile simili tra loro e, contemporaneamente, il più possibile diverse da quelle degli altri insiemi" [14].

Stress

E' una misura della bontà dell'adattamento (fit) per il modello di Multidimensional Scaling (MDS). Più piccolo è lo stress migliore l'accordo tra dati e modello. Tale stress misura le differenze delle distanze tra i punti nello spazio di calcolo di MDS e le corrispondenti distanze tra i dati in input e i risultati elaborati dall'algoritmo.

Stimolo

Non solo nella conjoint analysis ma anche più in generale si indica come stimolo una configurazione, un profilo, un'istanza di un oggetto che deve essere sottoposto ad una valutazione da parte di un rispondente ad

un questionario o intervista. In particolare, nella conjoint analysis gli stimoli corrispondono ai prodotti differenziati del piano fattoriale completo o frazionato, oggetto dell'analisi, ad ogni stimolo verranno associate opportune preferenze.

Tau di Kendall

E' un test di una statistica non-parametrica utilizzato per misurare il grado di corrispondenza tra due punteggi ed assegnare una significatività alla stessa misura, ciò significa che misura la forza dell'associazione tra le valutazioni del processo di classificazione in atto. Il test è stato sviluppato ed introdotto da Maurice Kendall nel 1938.

Bibliografia

1. AA.VV. - Multidimensional scaling Statnotes – North Carolina State University - <http://faculty.chass.ncsu.edu/garson/PA765/mds.htm>
2. Abdi Hervvé, RV Coefficient And Congruence Coefficient, Encyclopedia of Measurement and Statistics (2007) - <http://www.utdallas.edu/~herve/Abdi-RV2007-pretty.pdf>
3. Abe Shigeo – Pattern Classification – Springer (2001)- http://www.amazon.com/Pattern-Classification-Shigeo-Abe/dp/1852333529/ref=sr_1_1?ie=UTF8&s=books&qid=1262625717&sr=1-1
4. Barbaranelli, C. - Analisi dei dati con SPSS Vol. II. Milano: LED (2003) - <http://www.ibs.it/code/9788879163156/barbaranelli-claudio/analisi-dei-dati-con.html>
5. Bracalente, Mulas, Cossignani - Statistica aziendale - McGraw Hill (2009) - <http://www.ibs.it/code/9788838664960/bracalente-mulas-cossignani/statistica-aziendale.html>
6. Brasini Sergio, Freo Marzia, Tassinari Franco, Tassinari Giorgio - Statistica aziendale e analisi di mercato - Il Mulino, Bologna (2002) - <http://www.ibs.it/code/9788815088765/zzz1k1456/statistica-aziendale-e-analisi.html>
7. De Finetti B. – Sul significato soggettivo della probabilità (1931) <http://www.brunodefinetti.it/Opere/Sul%20significato%20soggettivo%20della%20probabilit%20E0.pdf>
8. Fabbri G. e Orsini R – Reti Neurali per le scienze economiche – Franco Muzio Editore (1993) - <http://www.libreriauniversitaria.it/reti-neurali-scienze-economiche-fabbri/libro/9788870216561>
9. Field Andy – Discovering Statistics using SPSS for Windows – SAGE Publication (2000) - <http://www.ibs.it/book/9781412977524/field-andy/discovering-statistics-using.html>
10. Gower J. C. and Legendre P. - Metric and Euclidian properties of dissimilarity coefficient – Journal of Classification (1986) - http://www.bio.umontreal.ca/legendre/reprints/Gower&Leg_JClass86.pdf
11. Howard Martin, Sappiamo cosa vuoi, Minimum Fax 2005, <http://www.ibs.it/code/9788875210687/howard-martin/sappiamo-cosa-vuoi-chi.html>
12. Jiawei Han – Data Mining, concepts and techniques – Morgan Kaufmann (2001).- http://www.amazon.com/Data-Mining-Concepts-Techniques-Management/dp/1558609016/ref=ntt_at_ep_dpt_1
13. Kotler Philip, Marketing management (2007) - <http://www.ibs.it/code/9788871922935/kotler-philip/marketing-management.html>
14. Molteni Luca, Gabriele Triolo - Ricerche di *marketing* - McGraw Hill (2003). - <http://www.ibs.it/code/9788838663925/molteni-luca/ricerche-di-marketing.html>
15. Molteni L., Gnechi M. – Le reti neurali nel marketing: il problema della segmentazione per obiettivi – LIUC Papers n. 45, Serie Metodi quantitativi (1997) - <http://www.biblio.liuc.it/liucpap/pdf/45.pdf>
16. Piccolo Domenico - Statistica per le decisioni - Il Mulino, Bologna (2004) - <http://www.ibs.it/code/9788815097705/piccolo-domenico/statistica-per-le-decisioni.html>
17. R (The) Development Core Team - R: A Language and Environment for Statistical Computing - Copyright (©) 1999–2009 R Foundation for Statistical Computing, ISBN 3-900051-07-0 - <http://cran.r-project.org/doc/manuals/refman.pdf>
18. Raghavan Vijay V. Tandhe M.Y.L. Ip – Techniques for measuring stability of clustering: a comparative study (1982) - <http://elvis.slis.indiana.edu/irpub/SIGIR/1982/pdf16.pdf>
19. Randy Julian – Lecture slides – Lilly Research Laboratories <http://miner.chem.purdue.edu/Lectures/>
20. Ramsey F. P. - Truth and probability (1926) - <http://fitelson.org/probability/ramsey.pdf>
21. Ruminati Rino, Psicologia economica, a cura di Ruminati Rino Enrico Rubatelli e Maurizio Mistri, Carocci 2008, <http://www.ibs.it/code/9788843044290/zzz1k1456/psicologia-economica.html>
22. Russel S. J., Norvig P. – Intelligenza artificiale – Pearson Education Italia (2005) - <http://www.ibs.it/code/9788871922287/russell-stuart-j/intelligenza-artificiale-approccio.html>
23. SPSS for Windows Documentation <http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/index.html>
http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/SPSS_16.0_Algorithms.pdf
http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/SPSS_Conjoint_16.0.pdf

24. Takane Yoshio - Applications of Multidimensional Scaling in Psychometrics - Department of Psychology McGill University - <http://takane.brinkster.net/Yoshio/c039.pdf>
25. Tammo H. A. Bijmolt and Michel Wedde - A Comparison of Multidimensional Scaling Methods for Perceptual Mapping - Journal Of Marketing Research, MAY 1999 – <http://www.springerlink.com/content/kgp57070538664w1/>
26. Technical Papers Library Sawtooth Software - <http://www.sawtoothsoftware.com/education/techpap.shtml>
27. Universität Hamburg – SPSS Algorithms – <http://www1.uni-hamburg.de/RRZ/Software/SPSS/Algorith.120/>
<http://www1.uni-hamburg.de/RRZ/Software/SPSS/Algorith.115/proxscal.pdf>
<http://www1.uni-hamburg.de/RRZ/Software/SPSS/Algorith.115/alscal.pdf>
28. Ulrich K. T., Eppinger S. D., Filippini R. – Progettazione e sviluppo del prodotto _ McGraw-Hill (1995) <http://www.ibs.it/code/9788838663970/ulrich-karl-t-eppinger/progettazione-e-sviluppo-di.html>
29. Written Ian H., Eibe Frank – Data Mining, practical machine learning tools – Morgan Kaufmann (2000) http://www.amazon.com/Data-Mining-Techniques-Implementations-Management/dp/1558605525/ref=sr_1_1?ie=UTF8&s=books&qid=1262625852&sr=1-1

Indice dei nomi

A	
algoritmo	10; 13; 15; 16; 21; 27
analisi discriminante	11
Analisi statistica multivariata	25
ASCAL	10
autovalori.....	15; 27
C	
Chi-quadro ...	1; 8; 9; 10; 12; 15; 18; 19; 22; 23; 25; 26; 27
classificazione	1; 11; 12
clienti	27
<i>cluster</i>	16
coefficiente	9; 22
coefficienti	22
componenti principali.....	10; 15
concorrenti	27
confronto	8; 18
consumatore	7
Correlazione.....	16; 25
Covarianza.....	25
D	
decisioni	29
<i>diagramma di Shepard</i>	18
dissimilarità	7; 8; 11; 12; 14; 23
distanza	7; 8; 15; 18; 20
domanda	27
I	
impresa	27
M	
mappa delle percezioni	16
Mappatura delle percezioni	26
<i>marketing</i>	1; 7; 29
Marketing	26; 29; 30
mercato	27; 29
metodo	7; 8; 10; 12
misura	8; 15; 16; 22; 27
modello.....	10; 15; 18; 23; 27
monotona	8; 10; 15; 18
multidimensional scaling.....	7; 10; 14
O	
<i>omoschedastico</i>	18
P	
Pearson	25; 27; 29
percezione.....	27
Percezione	27
posizionamento	1; 11; 12; 13; 16; 27
Posizionamento.....	27
probabilità.....	29
processo.....	14; 18; 20; 27
prodotto	27
prossimità	8; 10; 11; 12
Prossimità.....	27
PROXSCAL.....	10; 11; 13; 14; 20; 21; 22
R	
ricerca	14
S	
Scaling multidimensionale	27
Scree-plot.....	27
segmentazione	27
Segmentazione	27
servizio	27
similarità	7; 8
soggettivo	29
statistica	29
Stimolo	27
stress.....	8; 9; 15; 16; 21; 22; 27
Stress	10; 14; 15; 16; 20; 21; 27
T	
Tau di Kendall	27; 28
tecnica	10
V	
varianza.....	16; 18

