# Flussi elettorali e teorema di Perron-Frobenius

*Luca Ricolfi*

*Il contributo illustra un importante teorema matematico-statistico, il teorema di Perron-Frobenius, proponendone l'applicazione alle matrici di flussi elettorali. Questa innovativa proposta consente di rispondere ad un altrettanto innovativo quesito per l'analisi elettorale: quale distribuzione dei consensi si avrebbe oggi in Italia se la struttura dei flussi elettorali fosse priva di cambiamenti, ovvero fosse quella del passaggio 2006-08?*

Mi sono sempre chiesto perché, in analisi dei dati, sia così raro l'uso del teorema di Perron-Frobenius, nonostante le strutture dei dati che consentirebbero di sfruttarlo utilmente siano tutt'altro che rare o marginali. In questo breve contributo:

1. espongo il teorema nel caso di matrici quadrate non negative irriducibili;
2. introduco alcune ulteriori definizioni di matrici, più restrittive di quelle considerate da Perron e Frobenius;
3. richiamo il concetto di catena di Markov omogenea;
4. ridiscuto il teorema di Perron-Frobenius nel caso speciale di una matrice di transizione associata ad una catena di Markov regolare;
5. richiamo alcune strutture dei dati che consentono di sfruttare il teorema;
6. mostro un'applicazione del teorema in ambito elettorale.

Per gli studiosi di analisi elettorale, probabilmente l'applicazione più naturale ed utile del teorema riguarda le matrici di flusso. Una matrice di flusso è una tabella di frequenza, non necessariamente quadrata, che descrive in modo completo come il corpo elettorale di una determinata unità territoriale (Stato, regione, città, quartiere) ha votato in due elezioni successive.

Le matrici di flusso possono essere costruite sia a partire da dati di sondaggio (grazie al ricordo di voto), sia a partire da dati panel (intervistando in due occasioni elettorali successive uno stesso insieme di soggetti), sia a partire da dati ufficiali – normalmente a livello di sezione elettorale – ricostruendo la tabella completa dei flussi elettorali mediante tecniche matematico-statistiche, come il modello di Goodman.

Per applicare il teorema di Perron-Frobenius è indispensabile che le categorie (partiti o liste elettorali) delle elezioni al tempo t e al tempo t + 1 coincidano o siano rese omogenee mediante opportune aggregazioni.

## 1. Il teorema di Perron-Frobenius

Dimostrato all'inizio del Novecento, il teorema riguarda le matrici quadrate non negative, ossia le matrici che contengono solo elementi reali positivi o nulli.

Ad esempio la matrice 2 x 2:

$$\begin{matrix} 3 & 0 \\ 1 & 2 \end{matrix}$$

Si noti che la matrice non è simmetrica e inoltre contiene uno zero.

Il teorema si presenta in varie versioni a seconda del tipo di matrici quadrate non negative che esso ha per oggetto: matrici non negative qualsiasi, matrici positive, matrici non negative irriducibili, matrici non negative primitive, matrici stocastiche. Il contributo originario (1907) di Oskar Perron si riferiva alle matrici positive (con tutti gli elementi reali e maggiori di zero), quello successivo (1912) di Georg Frobenius si riferiva alle matrici non negative irriducibili, ossia matrici che, pur ammettendo alcuni zeri, risultano globalmente connesse[1].

Esso stabiliva, per tale classe di matrici:

– l'esistenza di un autovalore dominante semplice[2], reale e strettamente positivo, da allora denominato "radice di Perron" o "autovalore di Perron-Frobenius";

– l'esistenza e l'unicità di una coppia di autovettori (uno sinistro e uno destro), entrambi strettamente positivi, associati all'autovalore dominante.

In termini formali ciò significa che, data una matrice quadrata non negativa e irriducibile A, è sempre possibile trovare una coppia di vettori u e v strettamente positivi tali che:

$$u' A = \lambda u' \qquad [1]$$

$$A v = \lambda v \qquad [2]$$

dove $\lambda$ è l'autovettore dominante di A, mentre i due vettori[3] u' e v costituiscono rispettivamente l'autovettore sinistro (talora denominato autovettore di Frobenius) e l'autovettore destro associati a $\lambda$.

È forse il caso di aggiungere, per chi fosse abituato ad usare le comuni procedure di *Eigen Decomposition* (ED) e *Singular Value Decomposition* (SVD), che né l'una né l'altra si applicano al caso qui considerato. La ED lavora solo su matrici quadrate reali e simmetriche e ciò rende coincidenti autovettore sinistro e destro. La SVD si applica a matrici rettangolari (quindi anche quadrate non simmetriche, il nostro caso) ma i vettori sinistri e destri che restituisce non sono autovettori bensì vettori singolari, che obbediscono ad una definizione differente:

$$u' A = \sigma v' \qquad [3]$$

dove il vettore singolare sinistro u' ha tanti elementi quante sono le righe di A, mentre il vettore singolare destro v' ha tanti elementi quante sono le colonne di A.

## 2. Matrici speciali

Successivamente, i risultati di Perron e Frobenius sono stati sviluppati per speciali classi di matrici non negative, fra le quali particolare importanza assumono le matrici di transizione stocastiche (che contengono le probabilità di transizione fra K stati), in particolare nel caso delle cosiddette catene di Markov.

Per matrice *primitiva* si intende una matrice quadrata non negativa, irriducibile e aperiodica. Una matrice quadrata non negativa irriducibile è anche aperiodica se, moltiplicata per se stessa un numero sufficiente di volte, converge ad una matrice positiva. In buona sostanza, questa condizione equivale a richiedere che, per t che tende a infinito, esista il limite della matrice $A^t$ e tale limite sia una matrice positiva; o, equivalentemente, che – se si fa trascorrere un tempo $\tau$ sufficientemente lungo – qualsiasi stato iniziale abbia una probabilità non nulla di risultare trasformato in qualsiasi altro. Le matrici primitive sono importantissime, perché godono di tutte le proprietà che, nel 1907, Oskar Perron aveva dimostrato per le matrici positive.

Per matrice *stocastica* si intende una matrice non negativa, non necessariamente quadrata, in cui tutti i valori di riga (o tutti i valori di colonna) hanno somma 1. Spesso una tale matrice può essere interpretata come un sistema di probabilità condizionali, che descrive con quali probabilità da un insieme di I condizioni di partenza si passa ad un insieme di J condizioni di arrivo. Per esempio, come i membri di un insieme di I classi sociali di origine si distribuiscono fra un insieme di J comportamenti elettorali. O come un sistema con K stati possibili (ad esempio: occupato, disoccupato, inoccupato) evolve fra il tempo t e il tempo t + 1.

Per matrice *di transizione* si intende una matrice stocastica $A_t$ che specifica completamente le probabilità che fra il tempo t e il tempo t + 1 un dato sistema dinamico passi da uno qualsiasi dei suoi stati ad un qualsiasi altro[4]. In una matrice di transizione, quindi, l'insieme degli stati di partenza (di norma posti sulle righe) e l'insieme degli stati di arrivo (colonne) coincidono, e quindi la matrice non può che essere quadrata.

*Tab. 1. Esempio di matrice di transizione*

| | | t + 1 | | |
|---|---|---|---|---|
| | | A | B | C |
| t | A | 0,80 | 0,05 | 0,15 |
| | B | 0,20 | 0,30 | 0,50 |
| | C | 0,25 | 0,35 | 0,60 |

La diagonale principale della matrice contiene le probabilità di permanenza nel medesimo stato, mentre gli elementi *extra diagonal* contengono le probabilità di cambiare stato. Nelle applicazioni di analisi elettorale, in cui il sistema dinamico non è altro che un insieme di elettori o votanti, le probabilità di permanenza si riferiscono ai cosiddetti *stayers* (persone che non cambiano il loro comportamento di voto, o elettori fedeli), mentre le probabilità di cambiamento si riferiscono ai cosiddetti *movers* (persone che cambiano il loro comportamento di voto, o elettori infedeli).

## 3. Catene di Markov finite

Si chiama *catena di Markov* finita (o discreta) una matrice di transizione nella quale per ogni stato (riga), le relative probabilità di passaggio ai vari stati possibili (colonne) dipendano solo dallo stato di partenza e non dal comportamento passato del sistema (per questo si dice che i processi markoviani sono "senza memoria"). Se tali probabilità non variano nel tempo la catena di Markov si dice omogenea.

La matrice di transizione M di una catena di Markov omogenea può essere utilizzata per studiare l'evoluzione nel tempo di un sistema dinamico. Si parte da una distribuzione di probabilità iniziale $x_0$ (un vettore non negativo a somma 1) e le si applica la matrice M:

$$x_0' M = x_1' \qquad [4]$$

Il risultato è la distribuzione di probabilità al tempo 1. Proseguendo, si può applicare la matrice M al vettore $x_1$' ricavando la distribuzione di probabilità al tempo 2.

$$x_1' \, M = x_2' \qquad [5]$$

Naturalmente, la distribuzione al tempo t + 2 si può ricavare anche direttamente dalla distribuzione iniziale $x_0$', moltiplicandola due volte per la matrice M:

$$x_0' \, M \, M = x_2' \qquad [6]$$

ovvero:

$$x_0' \, M^2 = x_2' \qquad [7]$$

dove con $M^2$ si intende il prodotto della matrice M per se stessa.

Il processo può proseguire indefinitamente e permette di calcolare completamente la traiettoria del sistema:

$$x_0' \, M^t = x_t' \qquad [8]$$

dove $x_t$' è il vettore riga che contiene la distribuzione di probabilità al tempo t.

## 4. Il teorema di Perron-Frobenius per catene di Markov regolari

Fra le catene di Markov, particolare importanza assume il caso delle catene regolari, per cui vale la seguente definizione: una catena di Markov si dice regolare se è omogenea e inoltre la sua matrice di transizione è primitiva.

In tal caso il teorema di Perron-Frobenius si presenta in una versione più forte (talora qualificata come "teorema ergodico"), grazie al fatto che la matrice M, essendo non solo quadrata, non negativa e irriducibile, ma anche stocastica e primitiva, possiede alcune proprietà supplementari, che non valgono in generale per i casi considerati da Perron e Frobenius (matrici quadrate positive e matrici non negative irriducibili). Vediamo tali proprietà:

1. l'autovalore dominante è 1 ed è semplice (la sua molteplicità è 1);
2. l'autovettore destro d' è un vettore i cui elementi sono tutti pari a 1;
3. l'autovettore sinistro s' è una distribuzione di probabilità strettamente positiva (i suoi elementi sono tutti positivi, e sommano a 1);
4. l'autovettore sinistro s' rappresenta la distribuzione di equilibrio (o stazionaria), ossia quella in cui il sistema tende a permanere una volta che l'ha raggiunta[5];

5. per t che tende a infinito, la matrice M converge ad una matrice positiva M*:

$$\lim_{t \to \infty} M^t = M^*.$$

6. le righe di tale matrice M* sono identiche e contengono ciascuna il vettore di equilibrio s'.

7. data una distribuzione di probabilità iniziale $x_o$ qualsiasi, non necessariamente priva di zeri, il processo di Markov converge a s': il processo è equifinale.

Dal punto di vista operativo, questo significa che:

– se sappiamo che M è regolare, possiamo ricavare la distribuzione di equilibrio;
– la distribuzione di equilibrio s' può essere ricavata partendo da una distribuzione iniziale qualsiasi $x_o$, e applicando ripetutamente ad essa la matrice M finché la distribuzione risultante $x_t$ non si stabilizza;
– se non ci interessa la traiettoria del sistema verso l'equilibrio, ma solo il risultato finale, non ci occorre alcun vettore iniziale ma possiamo limitarci a calcolare la matrice M*, le cui righe contengono – ripetuta – la distribuzione di equilibrio s'.

In concreto, questo significa che – se possiamo dimostrare che la matrice di transizione di una determinata catena di Markov è primitiva – allora possiamo calcolare qual è la distribuzione finale cui il sistema tenderebbe se la sua matrice di transizione restasse costante nel tempo.

## 5. Ambiti di applicazione

Nel campo delle scienze sociali, forse l'ambito di applicazione più importante del teorema di Perron-Frobenius è stato quello delle tavole input-output, o matrici di Leontieff. Le tavole input-output descrivono, sotto ipotesi semplificatrici di linearità, la tecnologia di un apparato produttivo e permettono di analizzarne le proprietà.

Il caso più celebre ed elegante di uso del teorema è probabilmente quello compiuto negli anni Sessanta dagli economisti neo-ricardiani per descrivere le proprietà di un sistema economico a rendimenti costanti, sulla scia del celebre e fondamentale libro di Piero Sraffa, *Produzione di merci a mezzo di merci* (1960). In quel contesto l'autovalore dominante, o "radice di Perron", serve a determinare il "saggio di sovrappiù", mentre le proprietà formali della matrice dei coefficienti di produzione (ovviamente non negativa) servono a distinguere fra merci base e merci non base. L'autovettore sinistro[6], a sua volta, non è altro che il sistema dei prezzi, di cui – sotto opportune condizioni – è possibile dimostrare l'unicità e la natura non degenere (nessun prezzo è negativo o nullo). Questo tipo di applicazione occupa un posto cruciale nella storia del pensiero economico, perché ha consentito di dimostrare che, sotto ipotesi plausibili anche se non perfettamente realistiche (rendimenti costanti), il gioco della domanda e dell'offerta non svolge alcun ruolo nella determinazione dei prezzi di equilibrio, contrariamente a quanto per quasi un secolo aveva sostenuto la dominante teoria neoclassica (Lunghini, Ranchetti, 1998).

In analisi dei dati, nonostante il frequente uso di matrici di transizione, l'impiego del teorema di Perron-Frobenius è piuttosto raro. Eppure sono molto frequenti le situazioni che, sia pure con qualche assunto semplificatore[7], si presterebbero a sfruttare la versione forte del teorema, quella che si applica alle catene di Markov regolari.

In sociologia, l'indirizzo di ricerca empirico probabilmente più importante – quello degli studi di mobilità – lavora sistematicamente con matrici di transizione fra la struttura di classe dei padri e quella dei figli, generalmente analizzate con lo strumento dei modelli log-lineari. Gli studiosi, tuttavia, si concentrano quasi sempre sul problema di determinare il grado di mobilità sociale, senza riguardo alla distribuzione limite cui la struttura di classe potrebbe tendere. Esistono tuttavia delle eccezioni come Prais (1955), Matras (1961) e Bartholomew (1973).

In economia, le matrici di transizione fanno la loro comparsa per descrivere il funzionamento del mercato del lavoro, dove gli individui "circolano" sistematicamente fra almeno tre condizioni base: occupato, disoccupato, inoccupato. Anche qui l'uso delle catene di Markov è raro, benché tutto sommato più frequente che negli studi di mobilità.

In scienza politica, e in particolare in analisi elettorale, è piuttosto comune la stima di matrici di flusso fra due elezioni successive, talora in base a dati survey talora mediante il cosiddetto modello di Goodman (Schadee, Corbetta, 1984; Corbetta *et al.*, 1988) o qualche suo più o meno diretto discendente (De Sio, 2009). E tuttavia sono rarissimi i casi in cui, una volta determinata la matrice di flusso, si passa al calcolo dell'autovettore sinistro, che permetterebbe di sapere qual è la distribuzione dei consensi cui il sistema tenderebbe in assenza di cambiamenti nella struttura dei flussi.

È quello che, invece, proveremo a fare noi su dati reali, tratti dalle ultime consultazioni elettorali.

## 6. Dove andavano gli elettori: le tendenze nascoste nei flussi 2006-2008

Per la XV Legislatura, che copre il biennio 2006-08, sono disponibili diverse stime della matrice di flusso dei voti. Qui ne analizzeremo tre, dovute a Lorenzo De Sio, Paolo Feltrin e Luana Russo[8].

Per ciascuna di esse considereremo sia una versione ridotta, con 3 coalizioni più il non-voto, sia una versione analitica, con 8 partiti o gruppi di partiti più il non-voto[9]. Si tratta, in tutto, di 6 matrici di flusso, di cui 3 di formato 4 x 4 e altre 3 di formato 9 x 9 (riportate in Appendice B).

Prima di procedere al calcolo dell'autovettore di Frobenius, ovvero della distribuzione di equilibrio, conviene controllare che tale distribuzione esista o, equivalentemente, che le matrici di flusso siano equifinali. Ciò non è necessario quando le matrici di transizione sono interamente positive, ma lo diventa quando contengono qualche zero. In tali casi la matrice potrebbe non essere regolare e la distribuzione limite potrebbe non esistere.

Le nostre tre matrici "piccole" (4 x 4) sono certamente equifinali perché non contengono zeri. Delle restanti tre matrici nessuna è priva di zeri, quindi nulla assicura che siano primitive. E tuttavia bastano poche iterazioni per rendersi conto che sono tutte e tre primitive: dopo poche iterazioni diventano positive, il che garantisce l'equifinalità del processo.

Quanto al tempo di convergenza, abbiamo scoperto con sorpresa che è rapidissimo. Se stabiliamo convenzionalmente che il processo di Markov è "sostanzialmente converso" quando la distribuzione corrente è sufficientemente vicina a quella asintotica (finale), e come soglia di convergenza adottiamo una distanza media dell'1% fra distribuzione asintotica e distribuzione quasi-finale, constatiamo che in tutti e 6 i casi considerati (3 matrici piccole + 3 matrici grandi) il tempo di convergenza è appena 3. Poiché la legislatura 2006-08 è durata solo 2 anni, ciò significa che – se le sue tendenze fossero perdurate nel tempo – sarebbero bastate appena 3 legislature di 2 anni per osservare la stabilizzazione del sistema. E dal momento che il processo inizia nel 2006, 3 iterazioni di 2 anni significa 2006 + 6, ovvero 2012. Dunque, esattamente oggi.

Vediamo ora i risultati numerici, cominciando dalle matrici "piccole" (tab. 2).

*Tab. 2. Distribuzioni di equilibrio ricavate dall'applicazione del teorema di Perron-Frobenius alle tre matrici dei flussi sintetiche*

| | Distribuzione 2006 | Distribuzione 2008 | Distribuzioni di equilibrio | | | |
|---|---|---|---|---|---|---|
| | | | De Sio | Feltrin | Russo | Media |
| Centro destra | 35 | 38 | 49,0 | 49,2 | 49,4 | 49,2 |
| Udc | 5 | 4 | 3,4 | 4,7 | 3,2 | 3,8 |
| Centro sinistra | 40 | 33 | 19,9 | 20,0 | 16,3 | 18,7 |
| Non-voto | 20 | 24 | 27,7 | 26,1 | 31,1 | 28,3 |
| Totale | 100 | 100 | 100,0 | 100,0 | 100,0 | 100,0 |

Le prime due colonne contengono i risultati elettorali del 2006 e del 2008[10], la prima riporta dunque la distribuzione iniziale. Le tre colonne successive riportano le distribuzioni finali, o di equilibrio, associate alle matrici di transizione compatte (4 x 4) di De Sio, Feltrin, Russo. L'ultima colonna riporta la media delle tre distribuzioni di equilibrio.

I risultati delle tre analisi sono piuttosto simili, e precisamente: se le tendenze del 2006-08 fossero proseguite nell'attuale legislatura, avremmo assistito a uno sfondamento del centro destra, a un forte incremento del non-voto, a un lieve declino dell'Udc e ad un'implosione della sinistra. Più esattamente, il centro destra avrebbe raggiunto la maggioranza assoluta dei consensi, in quanto il 49% del corpo elettorale corrisponde ad oltre 2/3 dei voti validi; l'Udc sarebbe divenuto sempre meno rilevante; la sinistra avrebbe dimezzato i suoi consensi; la partecipazione elettorale sarebbe scesa drasticamente.

Naturalmente sappiamo che non è andata così, salvo forse per la partecipazione elettorale. I sondaggi ci dicono che oggi i rapporti di forza fra centro destra e centro sinistra non sono certo favorevoli alla destra: il centro destra raccoglie meno consensi del centro sinistra, mentre la distribuzione asintotica ci parla di un centro destra che ha quasi il triplo dei consensi del centro sinistra.

A che serve, dunque, l'analisi della distribuzione di equilibrio? Dipende. Se abbiamo ragione di pensare che una legislatura sia "normale", nel senso che esprime meccanismi e tendenze permanenti, la matrice di transizione può essere usata per fare caute previsioni. Se invece, come per lo più succede, la legislatura ha un suo profilo inconfondibile, che la rende diversa dalla maggior parte delle altre legislature, allora la distribuzione di equilibrio, nonché la traiettoria attraverso la quale la si raggiunge, servono a farci apprezzare meglio la natura e l'entità delle forze che hanno mosso il consenso entro tale legislatura. Il teorema di Perron-Frobenius è il grimaldello che ci consente di "aprire" la scatola della legislatura esaminandone le tendenze profonde. Proprio l'implausibilità della distribuzione di equilibrio rivela quanto potenti e anomale siano state le forze che hanno governato gli animi degli elettori nel corso di una determinato periodo. Nel caso da noi studiato, quello della legislatura del 2006-08, lo sviluppo della matrice di transizione ci rivela quanto radicale, intenso e profondo sia stato il rifiuto che l'elettorato ha maturato nei confronti del governo Prodi.

Un rifiuto che, se non fosse stato interrotto dalla vittoria del centro destra e dai nuovi meccanismi innescati dalla crisi economica e istituzionale, avrebbe cambiato radicalmente gli equilibri del sistema politico. Ma in che direzione? Verso destra, lo abbiamo visto. Ma in che modo? A favore del Pdl o della Lega? A danno del Pd o dei suoi alleati?

Per rispondere a queste domande può essere utile ripetere l'esercizio sulle versioni analitiche (9 x 9) delle nostre tre matrici (tab. 3). Ed ecco i risultati.

*Tab. 3. Distribuzioni di equilibrio ricavate dall'applicazione del teorema di Perron-Frobenius alle tre matrici dei flussi analitiche*

| | Distribuzione 2006 | Distribuzione 2008 | Distribuzioni di equilibrio | | | |
|---|---|---|---|---|---|---|
| | | | De Sio | Feltrin | Russo | Media |
| Pdl (Fi + An) | 30 | 29 | 30,1 | 30,8 | 27,4 | 29,4 |
| Lega + Mpa | 4 | 7 | 20,5 | 16,4 | 19,7 | 18,9 |
| Altri destra | 1 | 2 | 1,8 | 2,5 | 2,7 | 2,3 |
| Udc | 5 | 4 | 3,1 | 4,4 | 3,5 | 3,7 |
| Pd/Ulivo | 28 | 26 | 16,0 | 15,5 | 12,9 | 14,8 |
| Idv | 2 | 3 | 2,6 | 3,3 | 2,8 | 2,9 |
| Sinistra radicale | 7 | 3 | 1,0 | 1,2 | 1,3 | 1,2 |
| Altri sinistra | 3 | 1 | 0,7 | 0,4 | 1,0 | 0,7 |
| Non-voto | 20 | 24 | 24,2 | 25,5 | 28,7 | 26,1 |
| Totale | 100 | 100 | 100,0 | 100,0 | 100,0 | 100,0 |

Come si vede, nelle loro linee generali le distribuzioni analitiche confermano le tendenze delle distribuzioni sintetiche, con alcune specificazioni interessanti. Il rafforzamento del centro destra è interamente dovuto alla Lega, che tende a passare dal 4% del corpo elettorale a quasi il 20%, il che la farebbe diventare di gran lunga il primo partito del Nord. La crisi della sinistra sembra travolgere tutte le sue componenti, con il Pd che dimezza i suoi consensi e la Sinistra radicale che crolla intorno all'1%. Solo l'Italia dei valori fa segnare una leggera tendenza al rafforzamento, dal 2% al 3% circa. Quanto al non-voto si conferma la tendenza al suo aumento, anche se in modo leggermente meno accentuato di quanto risultava dalle matrici sintetiche.

Il teorema di Perron-Frobenius, naturalmente, va maneggiato con molta cura. Se la matrice di flusso non è stabile, o riflette un momento politico eccezionale, il teorema consente di mettere più chiaramente in luce tale eccezionalità, perché la distribuzione di equilibrio appare irrealistica. Se invece la matrice di flusso è stabile, o potrebbe stabilizzarsi intorno ai valori rilevati, il teorema consente di "tracciare" l'evoluzione del sistema in un modo accurato, che di norma diverge sensibilmente da una semplice estrapolazione delle tendenze osservate fra due elezioni.

## Appendice A. Programma Matrix per il calcolo dell'autovettore sinistro

Di seguito riportiamo le istruzioni, scritte in linguaggio Matrix di SPSS, di un semplice programma che, data una matrice di flusso, calcola:

*a*) la distribuzione iniziale;
*b*) il tempo di convergenza in prossimità della distribuzione di equilibrio;
*c*) la distribuzione quasi-finale o di quasi-equilibrio;
*d*) la distribuzione finale o di equilibrio.

Per matrice di flusso intendo una matrice quadrata contenente i flussi assoluti, oppure le probabilità (o le percentuali) di transizione totali. Non sono ammesse matrici con probabilità condizionali, di riga o di colonna.

La matrice di input deve essere quadrata, non negativa e primitiva. Il dataset che la contiene deve essere un file dati SPSS (formato: NOME.sav), organizzato ponendo sulle righe gli stati al tempo t e sulle colonne quelli al tempo t + 1.

L'utente può scegliere la soglia che definisce quando una distribuzione può essere considerata sufficientemente vicina a quella finale (parametro "soglia" all'inizio del programma).

### Istruzioni in Matrix

```
*Programma Perron_frobenius (per Polena).

set mxloops = 1.000.000.

matrix.

*Parametri di governo.
compute soglia = 0,01.
compute cicli = 1.000.
compute precis = 1/1.000.000.

*Lettura dei dati (matrice SQARE.sav, di flusso o di transizione).

get A /file='h:\markov\ SQARE.sav'.

*Controllo che l'input sia una matrice quadrata.
do if (ncol(A) ne nrow(A)).
print /Title "Matrice rettangolare!".
end if.

compute n = 0,5*(ncol(A)+nrow(A)).
```

```
*Trasformazione dell'input in una matrice di transizione.
compute Arighe = rsum (A).
compute x = t (Arighe)/msum(Arighe).
compute A = inv (mdiag(Arighe))*A.

*Calcolo della configurazione di equilibrio.
compute ZERO = make(n,n,0).
compute Bold = ZERO.
compute B = A.

***.
loop j = 1 to 100.
compute B = B*B.
compute xcorr = x*B.
compute stress = msum(abs(B-Bold))/n**2.
compute Bold = B.
end loop if stress < precis.
***.
compute xasint = B(1,:).

*Calcolo del tempo di convergenza.
compute C = A.
***.
loop k = 1 to 100.
compute C = C*C.
compute xcorr = x*C.
compute dist = msum(abs(xcorr-xasint))/n.
compute scarto = 100*dist.
compute Cold = C.
end loop if dist < soglia.
***.

*Stampa dei risultati.
print k /title = "Tempo di convergenza".
print x /title = "Distribuzione iniziale".
print xcorr /title = "Distribuzione quasi-finale".
print xasint /title = "Distribuzione finale (autovettore sinistro)".

end matrix.
```

## Appendice B. Matrici di flusso

*Tab. 1. Matrici dei flussi elettorali 2006-08 (in versione sintetica)*

| | 2008 – De Sio | | | | 2008 – Feltrin | | | | 2008 – Russo | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2006 | Cd | Udc | Cs | Nv | Cd | Udc | Cs | Nv | Cd | Udc | Cs | Nv |
| Cd | 29,2 | 2,3 | 3,4 | 3,3 | 28,6 | 1,2 | 3,3 | 5,1 | 30,8 | 2,2 | 3,8 | 2,1 |
| Udc | 0,5 | 2,3 | 0,9 | 0,6 | 1,0 | 1,3 | 1,4 | 0,7 | 0,6 | 2,3 | 1,2 | 0,4 |
| Cs | 1,8 | 0,4 | 29,2 | 1,9 | 1,2 | 0,3 | 27,2 | 3,6 | 1,8 | 0,7 | 30,7 | 0,5 |
| Nv | 3,2 | 0,5 | 6,8 | 13,8 | 4,9 | 0,7 | 8,6 | 11,0 | 1,9 | 0,4 | 4,5 | 16,7 |

*Nota*: per le fonti dei dati vedi nota 8.

*Tab. 2. Matrice dei flussi elettorali 2006-08 (in versione analitica)*

| | 2008 – De Sio | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2006 | Pdl | Ln/Mpa | Altri des. | Udc | Pd/Ulivo | Idv | Sin. rad. | Altri sin. | Nv |
| Pdl | 22,3 | 0,3 | 0 | 1,8 | 1,3 | 0,3 | 0,2 | 0,2 | 2,6 |
| Ln/Mpa | 2,2 | 3,1 | 0 | 0,4 | 0,3 | 0 | 0,3 | 0,7 | 0,4 |
| Altri des. | 1,3 | 0 | 0 | 0,1 | 0,1 | 0 | 0 | 0 | 0,3 |
| Udc | 0,5 | 0 | 0 | 2,3 | 0,9 | 0 | 0 | 0 | 0,6 |
| Pd/Ulivo | 0,4 | 0,1 | 0,4 | 0,2 | 18,8 | 0,3 | 2 | 2 | 1,4 |
| Idv | 0,1 | 0 | 0,3 | 0,1 | 1 | 1 | 0,5 | 0,2 | 0,2 |
| Sin. rad. | 0 | 0 | 0 | 0 | 0,3 | 0,1 | 2,6 | 0,1 | 0,3 |
| Altri sin | 0 | 0 | 0,5 | 0,1 | 0,1 | 0 | 0,1 | 0,1 | 0 |
| Nv | 2,4 | 0,2 | 0,6 | 0,5 | 2,5 | 0,5 | 2,6 | 1,2 | 13,8 |

*Nota*: per le fonti dei dati vedi nota 8.

*Tab. 3. Matrice dei flussi elettorali 2006-08 (in versione analitica)*

| | 2008 – Feltrin | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2006 | Pdl | Ln/Mpa | Altri des. | Udc | Pd/Ulivo | Idv | Sin. rad. | Altri sin. | Nv |
| Pdl | 21,4 | 0,5 | 0,1 | 1,0 | 1,8 | 0,1 | 0,2 | 0,2 | 3,7 |
| Ln/Mpa | 2,2 | 3,0 | 0,0 | 0,2 | 0,6 | 0,0 | 0,1 | 0,1 | 1,1 |
| Altri des. | 1,2 | 0,1 | 0,1 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,2 |
| Udc | 1,0 | 0,0 | 0,0 | 1,3 | 1,1 | 0,1 | 0,1 | 0,1 | 0,7 |
| Pd/Ulivo | 0,6 | 0,1 | 0,1 | 0,2 | 19,9 | 0,4 | 1,4 | 0,5 | 2,5 |
| Idv | 0,2 | 0,0 | 0,0 | 0,0 | 1,5 | 0,7 | 0,3 | 0,1 | 0,6 |
| Sin. rad. | 0,0 | 0,0 | 0,0 | 0,0 | 0,8 | 0,0 | 1,1 | 0,1 | 0,3 |
| Altri sin. | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,0 | 0,2 | 0,1 |
| Nv | 4,4 | 0,4 | 0,1 | 0,7 | 6,4 | 0,4 | 1,2 | 0,7 | 11,0 |

*Nota*: per le fonti dei dati vedi nota 8.

*Tab. 4. Matrice dei flussi elettorali 2006-08 (in versione analitica)*

| 2006 | 2008 – Russo | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Pdl | Ln/Mpa | Altri des. | Udc | Pd/Ulivo | Idv | Sin. rad. | Altri sin. | Nv |
| Pdl | 22,8 | 0,2 | 0,6 | 1,4 | 0,6 | 0,2 | 0,6 | 0,8 | 1,6 |
| Ln/Mpa | 2,3 | 3,0 | 0,4 | 0,5 | 0,2 | 0,1 | 0,1 | 0,3 | 0,2 |
| Altri des. | 1,0 | 0,1 | 0,3 | 0,3 | 0,3 | 0,1 | 0,3 | 0,2 | 0,2 |
| Udc | 0,3 | 0,1 | 0,2 | 2,3 | 0,5 | 0,1 | 0,2 | 0,4 | 0,4 |
| Pd/Ulivo | 0,5 | 0,0 | 0,1 | 0,4 | 20,4 | 0,3 | 2,5 | 1,1 | 0,1 |
| Idv | 0,5 | 0,1 | 0,1 | 0,1 | 1,1 | 0,7 | 0,5 | 0,3 | 0,1 |
| Sin. rad. | 0,1 | 0,0 | 0,1 | 0,1 | 0,4 | 0,0 | 2,2 | 0,2 | 0,2 |
| Altri sin. | 0,2 | 0,0 | 0,1 | 0,1 | 0,2 | 0,1 | 0,2 | 0,5 | 0,1 |
| Nv | 1,4 | 0,1 | 0,4 | 0,4 | 1,6 | 0,3 | 1,8 | 0,8 | 16,7 |

*Nota*: per le fonti dei dati vedi nota 8.

## Note

[1] Più esattamente, una matrice quadrata non negativa si dice irriducibile se non esiste una permutazione che permetta di scriverla come una matrice triangolare a blocchi (con un blocco di soli zeri al di sopra o al di sotto della diagonale principale). Equivalentemente, una matrice quadrata non negativa A è irriducibile se, data una qualsiasi coppia ordinata i e j di stati, esiste un intero k per cui l'elemento $a_{ij}(k)$ della matrice $A^k$ è maggiore di zero. Il significato sostanziale della irriducibilità è che il grafo associato alla matrice – costruito tracciando una freccia ogniqualvolta un numero positivo connette il nodo i al nodo j – risulta connesso, ovvero strutturato in modo tale che da ogni elemento sia possibile, seguendo le frecce del grafo, arrivare a qualsiasi altro.

[2] Un autovalore si dice semplice se il suo autovettore sinistro e il suo autovettore destro sono unici.

[3] Come di consueto, la lettera minuscola (u, v ecc.) indica un vettore colonna, mentre il corrispondente vettore riga si indica con la medesima lettera seguita da un apice (u', v' ecc.).

[4] Fra le probabilità di passaggio, naturalmente, dallo stato i allo stato j.

[5] Questa proprietà si deduce immediatamente dalla definizione di autovettore sinistro e dalla circostanza che l'autovalore dominante sia 1. Infatti l'espressione u' M = 1 u' indica precisamente che l'applicazione della matrice M alla distribuzione u' restituisce esattamente la medesima distribuzione.

[6] A seconda dei contesti e delle notazioni, quello che nella nostra analisi è l'autovettore sinistro può diventare l'autovettore destro (è il caso di alcune formalizzazioni del modello di Sraffa).

[7] Mi riferisco, in particolare, alla condizione di chiusura del sistema di stati. Quando la matrice descrive i passaggi di stato di un insieme di unità elementari (individui, famiglie, imprese ecc.), a rigore occorre assumere che nessuna unità elementare entri o esca dal sistema, perché la matrice di transizione non prevede scambi con l'esterno ma solo circolazione interna delle unità. Questa restrizione non è particolarmente grave quando l'intervallo temporale considerato è breve, ma può diventare critica se si considerano periodi lunghi.

[8] La matrice di De Sio si basa su dati *survey* ed è tratta dal volume ITANES, *Il ritorno di Berlusconi. Vincitori e vinti nelle elezioni del 2008*, p. 65. La matrice di Feltrin si basa su 23.661 interviste, effettuate con metodo CATI fra il 19 febbraio e il 12 marzo 2008 per conto della società "Tolomeo Studi e Ricerche". La matrice di Luana Russo è invece stimata in base ai dati delle sezioni elettorali (si veda l'articolo pubblicato in questo numero di "Polena").

[9] Per non-voto qui intendiamo il comportamento di quanti si astengono, votano scheda bianca o nulla, oppure scelgono una forza politica non riconducibile a nessuna delle principali coalizioni; nelle elezioni del 2006 e del 2008 queste formazioni erano di entità trascurabile, e la loro assimilazione al non-voto vero e proprio non altera sostanzialmente i risultati dell'analisi.

[10] Le due colonne contengono valori medi, ricavati a partire dalle distribuzioni marginali delle tre matrici di flusso, tra loro molto simili. Solo la matrice di Feltrin è un po' diversa, in quanto assegna un valore più basso all'Udc nel 2006. Le differenze fra matrici sono dovute al fatto che sono ottenute con metodi diversi (universo delle sezioni, campioni di sezioni e dati survey), e in questo lavoro non sono state preventivamente uniformate con un ciclo di IPF (*Iterative Proportional Fitting*).

## Riferimenti bibliografici


Bartholomew D. J.
1973 *Stochastic Models for Social Processes*, Wiley, New York (II ed.).

Corbetta P., Parisi A., Schadee H.
1988 *Elezioni in Italia: struttura e tipologia delle consultazioni politiche*, il Mulino, Bologna.

De Sio L.
2009 *Oltre il modello di Goodman: l'analisi dei flussi in base a dati aggregati*, in "Polena", 1, pp. 9-35.

Frobenius G. F.
1912 *Über Matrizen aus nicht negativen Elementen*, Sitzungsber. Königl. Preuss. Akad. Wiss., pp. 456-77.

Lunghini G., Ranchetti F.
1998 *Valore, teorie del valore*, in *Enciclopedia delle scienze sociali*, Istituto dell'enciclopedia italiana, Roma.

Matras J.
1961 *Differential Fertility, Intergenerational Occupational Mobility and Change in the Occupational Distribution: Some Elementary Interrelationships*, in "Population Studies", 15, pp. 187-97.

Minc H.
1988 *Non-negative Matrices*, Wiley, New York.

Perron O.
1907 *Zur Theorie der Matrices*, in "Mathematische Annalen", 64, 2, pp. 248-63.

Prais J.
1955 *The Formal Theory of Social Mobility*, in "Population Studies", IX, 1, pp. 72-81.

Schadee H. M. A., Corbetta P. G.
1984 *Metodi e modelli di analisi dei dati elettorali*, il Mulino, Bologna.

Seneta E.
1981 *Non-negative Matrices and Markov Chains*, Softcover Springer Series in Statistics, New York (II ed.).

Sraffa P.
1960 *Produzione di merci a mezzo di merci. Premesse a una critica della teoria economica*, Einaudi, Torino.