IL NOSTRO PRIMO APPROCCIO ai PROCESSI di MARKOV

Riconsideriamo il nostro (vecchio...) sistema

|_|_|_| U3 |_|_| U2 con 3 Nere ● ● ●
e 2 Bianche ○ ○

Come ricorderete, abbiamo imparato a distinguere MACROSTATI e MICROSTATI, cosa che è possibile una volta che si sia fissato CHE TIPO di OSSERVAZIONE vogliamo fare sul nostro sistema.

La scelta di interessarci ai MACROSTATI di COLORE ci ha fatto concludere che (rispetto ad osservazioni del COLORE)

- il SISTEMA ha 3 (e solo 3...) CONFIGURAZIONI
- identificabili dal VALORE di n_3 = # di Nere in U3 : $n_3 = 1, 2, 3$
- ogni altra "osservabile di COLORE" risulta fissata dalle REGOLE di CONSERVAZIONE (n_2 = # N in U2 ; b_3 = # B in U3 ; b_2 = # B in U2)

$n_3 + n_2 = 3$ (conservazione del # totale di N)
$b_3 + b_2 = 2$ (conservazione del # totale di B)
$b_3 + n_3 = 3$ (conservazione del # totale di biglie in U3)
$b_2 + n_2 = 2$ (conservazione del # totale di biglie in U2)

→ Posso allora associare alla configurazione la etichetta n_3 e parlare di

CONFIG. 1 ($n_3 = 1$ ovvero ● ○ ○ / ● ●)
CONFIG. 2 ($n_3 = 2$ ovvero ● ● ○ / ● ○)
CONFIG. 3 ($n_3 = 3$ ovvero ● ● ● / ○ ○)

Avevamo calcolato le PROBABILITA' a PRIORI di REALIZZARE ciascuna CONFIGURAZ.

con il risultato $P(m_3=1)=\frac{3}{10}$ ($m_3=1 \Leftrightarrow m_2=2$ e estrarre la sequenza ●● da mettere in U2 ha Prob data da $\frac{3}{5}\cdot\frac{2}{4}=\frac{3}{10}$)

$P(m_3=2)=\frac{6}{10}$ ($m_3=2 \Leftrightarrow m_2=1$ e per realizzare ●○ in U2 devo estrarre ○ ●○ ○ ○● da cui la Prob. $\frac{3}{5}\cdot\frac{2}{4}+\frac{2}{5}\frac{3}{4}=\frac{6}{10}$)

$P(m_3=3)=\frac{1}{10}$ ($m_3=3 \Leftrightarrow m_2=0$ e dunque devo avere ○○ in U2 e allora la Prob. è $\frac{2}{5}\frac{1}{4}=\frac{1}{10}$)

Notare che (ovviamente) $\forall_{i=1}^{3} \quad P(m_3=i)\geq 0$

e $\sum_{i=1}^{3} P(m_3=i)=1$

... era stata la nostra PALESTRA per l'utilizzo delle regole per PROB. di EVENTI SOMMA e PRODOTTO ...

Notiamo subito che DATA la $P(m_3)$ posso CALCOLARE VALORI MEDI (su di essa!)

Così ad esempio $M[m_3]=\sum_{i=1}^{3} i\,P(m_3=i)=1\cdot\frac{3}{10}+2\cdot\frac{6}{10}+3\cdot\frac{1}{10}=1.8$

che scriveremo anche $\langle m_3\rangle_p=1{,}8$

$\hookrightarrow$ le $\langle\ldots\rangle$ dicono VALOR MEDIO e il PEDICE $\langle\ldots\rangle_p$ $\rightsquigarrow$ dice che la MEDIA è su P...

Osservazione banale: $\langle m_3\rangle_p$ NON è INTERO (ovvero non coincide con un valore REALIZZABILE)

Perché dovrebbe...? BASTA che SIA (questo sì...) $1 \leq \langle n_3 \rangle_p \leq 3$

(dove nel nostro caso gli eguali sono esclusi, perché ogni $n_3 = i$ ha $P(n_3 = i) > 0$ ...)

NB NOTARE che invece (banale, ma prestateci attenzione ...)

$$\langle b_3 \rangle_p = 1.2 \quad (0 \cdot 0.1 + 1 \cdot 0.6 + 2 \times 0.3)$$

così che $\langle n_3 \rangle_p + \langle b_3 \rangle_p = 3$ che è OVVIO perché

$$\langle n_3 \rangle_p + \langle b_3 \rangle_p = \langle n_3 + b_3 \rangle_p = \langle 3 \rangle_p = 3$$

(una costante c ha sempre
$M[c] = c$ ...
$M[c] = \sum_i c \cdot P_i = c \sum_i P_i = c$ ...)

e (verificatelo pure...) TUTTE le REGOLE di CONSERVAZIONE VALGONO anche come VALORI di ATTESA!

Ora FACCIAMO QUALCOSA di DIVERSO, che vi ricorderà quanto fatto con il PROCESSO di EHRENFEST ... ovvero STABILIAMO le REGOLE di un GIOCO di ESTRAZIONE che realizza un PROCESSO STOCASTICO

Al momento ci basta una DEFINIZIONE OPERATIVA:

realizziamo una SEQUENZA di CONFIGURAZIONI che si realizzano come CONSEGUENZA di EVENTI STOCASTICI ...

Per la precisione ...

0. Realizza una CONFIGURAZIONE del SISTEMA
(ovvero preparalo – tipicamente in MODO CASUALE – così da avere un DATO VALORE di n_3)

1. Estrai A CASO una BIGLIA da U_3 ed
estrai A CASO una BIGLIA da U_2 e

SCAMBIALE di POSTO

ad esempio ● ● ○   ○ ●   così che ci troveremo in ● ● ●   ○ ○

(ovvero abbiamo REALIZZATO la TRANSIZIONE $n_3 = 2 \longrightarrow n_3 = 3$ ...)

2. ITERA la PROCEDURA in 1. per un certo TEMPO T

NOTA BENE la unità di misura del tempo qui è un po' strana, ma ben definita: il tempo $t = K$ è quello a cui avviene il K-mo scambio di biglie...

Attenzione!

Notiamo subito che PER QUANTO IL PROCESSO AVVENGA NEL TEMPO, è tuttavia vero che SOLO IL TEMPO PRESENTE DETERMINA la TRANSIZIONE alla PROSSIMA CONFIGURAZIONE!

↓

Le chiameremo $\boxed{\text{CATENE di MARKOV}}$

↓

"Il futuro dipende dal passato solo attraverso il presente..."

Abbiamo provato a SIMULARE AL CALCOLATORE il PROCESSO

(cfr la funzione scritta da Enrico a lezione...) e ci siamo chiesti:

→ Se MEDIAMO su TEMPI LUNGHI il valore di M_3 calcolato sulle CONFIGURAZIONI ESTRATTE, possiamo REGISTRARE una qualche CONVERGENZA ad un DATO VALORE?

Possiamo dire che calcoliamo la MEDIA ARITHMETICA dei VALORI di M_3, che in questo contesto possiamo facilmente INTERPRETARE come MEDIA TEMPORALE calcolata sulla EVOLUZIONE del PROCESSO

→ La CONCLUSIONE EMPIRICA alla luce dei nostri ESPERIMENTI NUMERICI è che APPARENTEMENTE $\frac{1}{T}\sum_{t=1}^{T} M_3(t) \xrightarrow[T\to\infty]{} \sim 1.8$

dove $M_3(t)$ è il VALORE di M_3 OSSERVATO al TEMPO t, ovvero (in accordo a quanto detto...) DOPO la t-ma ESTRAZIONE

... PERCHE' ... ?

Facciamo un poco di ordine e notiamo che

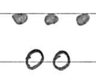
LE REGOLE del GIOCO FISSANO AD OGNI ISTANTE le STESSE PROBABILITA' di TRANSIZIONE da una (qualsiasi!) CONFIG. j a una (qualsiasi) CONFIGURAZ. i

Usiamo la NOTAZIONE $\boxed{P(i \leftarrow j)}$ = PROBAB. di TRANSIZIONE dalla CONFIGURAZIONE j alla CONFIGUR. i

Siccome i e j possono assumere 3 VALORI (per il nostro sistema che ha 3 configur.) avrò $3 \times 3 = 9$ PROB. di TRANSIZ. (ovvero 9 numeri)

Facciamo qualche esempio
(a lezione le abbiamo calcolate tutte)

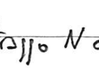
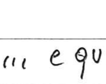
✓ $P(3 \leftarrow 2)$ ovvero $P(M_3 = 3 \longleftarrow M_3 = 2)$

ovvero devo passare da  a 

é necessario ESTRARRE una N da U2 e una B da U3 ...

e dunque $P(3 \leftarrow 2) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$

✓ analogamente $P(2 \leftarrow 2)$ ovvero $P(M_3 = 2 \longleftarrow M_3 = 2)$

cioé da  (ancora ...) a 

evidentemente
- o estraggo N da U3 e N da U2
- o estraggo B da U3 e B da U2

... e quindi $P(2 \leftarrow 2) = \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \frac{1}{2} = \frac{3}{6} = \frac{1}{2}$

✓ Notare che QUALCHE TRANSIZIONE è ESCLUSA, ad esempio

$$P(3 \leftarrow 1) = 0$$ (Non posso andare da a ...) ...

✓ ... e UNA è CERTA, ovvero $P(2 \leftarrow 3) = 1$

(da vado con certezza in)...

Ri-calcolatevele TUTTE! e ricordate che le ABBIAMO ASSEMBLATE in una MATRICE W

$$\boxed{W_{ij} = P(i \leftarrow j)}$$

così che in fin dei conti

$$W = \begin{pmatrix} 1/3 & 1/3 & 0 \\ 2/3 & 1/2 & 1 \\ 0 & 1/6 & 0 \end{pmatrix}$$

dove VALGONO le REGOLE

$$\forall_{i,j=1}^{3} \quad W_{ij} \geq 0$$ $\longrightarrow$ ovvio... è una PROBABILITA'

$$\forall_{j=1}^{3} \quad \sum_{i=1}^{3} W_{ij} = 1$$ $\longrightarrow$ ovvio ... $\sum_i P(i \leftarrow j) = 1$

"da qualche parte devo pur finire!"

... che è una CONSERVAZIONE delle PROBABILITA'

Ogni MATRICE W $N \times N$ che soddisfi $\boxed{W_{ij} \geq 0 \quad \forall_{i,j=1}^{N} \qquad \sum_{i=1}^{N} W_{ij} = 0 \quad \forall_{j=1}^{N}}$

è detta $\boxed{\text{MATRICE STOCASTICA}}$

✓ Notate che possiamo descrivere tramite W il nostro processo!

→ con la $W = \begin{pmatrix} 1/3 & 1/3 & 0 \\ 2/3 & 1/2 & 1 \\ 0 & 1/6 & 0 \end{pmatrix}$

Andiamo con ordine...

- la "MOSSA 0." (realizza una configurazione del sistema a caso) equivale a DEFINIRE una $P^{(0)}$ da intendersi come $P^{(t=0)}$ ovvero $P_i^{(t=0)}$ = Prob. di essere in $m_3 = i$ al tempo $t=0$

$P^{(0)}$ è un vettore a 3 entrate, che scriveremo come VETTORE COLONNA $\begin{pmatrix} \\ \\ \end{pmatrix}$

Nota che è anche possibile prendere (ad esempio) $P^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

↓

vuol dire: PARTI (con certezza...) da $m_3 = 1$

- Ora ci chiediamo: QUANTO VALE $P^{(1)}$ dove $P^{(1)}$ è $P^{(t=1)}$ e

$P_i^{(1)}$ = PROB. di avere $m_3 = i$ al tempo $t=1$

Ovviamente vale (SOLITE REGOLE!!!) $P_i^{(1)} = \sum_{j=1}^{3} P_j^{(0)} P(i \leftarrow j) \qquad (*)$

ovvero devo sommare su tutte le possibili j di partenza (a $t=0$) la prob. di avere $m_3=j$ al tempo $t=0$ $\times$ la probabilità di transire da $m_3=j \longrightarrow m_3=i$
(che notiamo $P(i \leftarrow j)$)

e SICCOME $W_{ij} = P(i \leftarrow j)$ vale la <u>RELAZIONE MATRICIALE</u>

$$P^{(1)} = W \, P^{(0)}$$

vettore 3x1 — matrice 3x3 — vettore 3x1

che IN COMPONENTI SCRIVO $P_i^{(1)} = \sum_{j=1}^{3} W_{ij} \, P_j^{(0)}$ (che è la (*) di pag 8...)

Ovviamente vale che $P_i^{(2)} = \sum_{j=1}^{3} P_j^{(1)} \, P(i \leftarrow j)$ (stesso ragionamento di prima per la SECONDA TRANSIZIONE...)

ovvero $P_i^{(2)} = \sum_{j=1}^{3} W_{ij} \, P_j^{(1)}$

ma da quanto visto (non offendetevi... occhio agli indici!)

$$P_i^{(2)} = \sum_{j=1}^{3} W_{ij} \sum_{k=1}^{3} W_{jk} P_k^{(0)}$$

non offendetevi... moltiplicazione righe per colonne

$$= \sum_{k=1}^{3} \left(\sum_{j=1}^{3} W_{ij} W_{jk} \right) P_k^{(0)}$$

$$= \sum_{k=1}^{3} W^2_{ik} \, P_k^{(0)}$$ ovvero

$$P^{(2)} = W P^{(1)} = W^2 P^{(0)}$$

– Vale EVIDENTEMENTE $P^{(N)} = W^N P^{(0)}$

Tutto questo non ci sorprende... chiediamoci infatti...

QUAL E' IL SIGNIFICATO di W^2_{ij}?

E' facile capirlo... $W^2_{ij} = \sum_K W_{iK} W_{Kj} = \sum_K W_{Kj} W_{iK} = \sum_K P(K \leftarrow j)\, P(i \leftarrow K)$

perdonatemi ancora una volta la banalità... questi sono NUMERI COMMUTANTI e per chiarezza espositiva li riscriviamo così...

ovvero è il prodotto della PROB. di transire (prima... potremmo dire...) da j a K e della PROB. di transire (poi...) da K a i, SOMMATO su tutte le CONFIG. (intermedie) K

IN SOSTANZA... $W^2_{ij} = P(i \leftarrow j, \text{in 2 PASSI})$ ... consistente con $P^{(2)} = W^2 P^{(0)}$ ...

N.B. W STOCASTICA ovvero
(1) $W_{ij} \geq 0$
(2) $\sum_i W_{ij} = 1$

$\Longrightarrow$ W^2 STOCASTICA, in quanto

(a) $W^2_{ij} = \sum_K W_{iK} W_{Kj} \geq 0$ (banale da (1)...)

(b) $\sum_i W^2_{ij} = \sum_{iK} W_{iK} W_{Kj} = \sum_K \left(\sum_i W_{iK}\right) W_{Kj}$

da (2)... $= \sum_K W_{Kj} = 1$

ancora da (2)...

Quindi $W^2_{ij} \geq 0$

e $\sum_i W^2_{ij} = 1$ e dunque W STOCASTICA $\Longrightarrow$ W^2 STOCASTICA

ESERCIZIO

1. che dire di $W^{N>2}$?

2. se W_1 STOCASTICA e W_2 STOCASTICHE, che dire di $W_3 \equiv W_1 W_2$?

OK! Ora il formalismo è chiaro, ma che dire della domanda che ci eravamo posti, ovvero PERCHE' APPARENTEMENTE $\frac{1}{T}\sum_{t=1}^{T} m_3(t) \xrightarrow[T\to\infty]{} \text{cost}$

(costante che peraltro abbiamo ragione di sospettare valga 1.8 ...)

AUTOVALORI e AUTOVETTORI di una W STOCASTICA

"Dimmi chi è il tuo SPETTRO e ti dirò chi sei..."

Scriviamo il PROBLEMA agli AUTOVALORI e AUTOVETTORI per una W STOCASTICA

(ovvero tale che (1) $W_{ij} \geq 0$
(2) $\sum_i W_{ij} = 1$)

In NOTAZIONE MATRICIALE

$$W \xi^{(\lambda)} = \lambda \xi^{(\lambda)}$$

AUTOVALORE λ

AUTOVETTORE CORRISPONDENTE a λ

IN COMPONENTI

$$\sum_k W_{ik} \xi_k^{(\lambda)} = \lambda \xi_i^{(\lambda)}$$

che ora SOMMO su i (vale $\forall i$...), RICORDANDO la (2)

$$\sum_{ki} W_{ik} \xi_k^{(\lambda)} = \sum_k \Big(\sum_i W_{ik}\Big) \xi_k^{(\lambda)} \underset{\text{cfr (2)}}{=} \sum_k \xi_k^{(\lambda)}$$ 1° MEMBRO

mentre a 2° membro ho semplicemente $\lambda \sum_i \xi_i^{(\lambda)}$

IN CONCLUSIONE (Attenti! e perdonate ancora... gli INDICI sono MUTI!)

$$\sum_i \xi_i^{(\lambda)} = \lambda \sum_i \xi_i^{(\lambda)}$$ $\longrightarrow$ 2 SOLUZIONI (possibilità...)

- $\boxed{\lambda = 1}$
- $\boxed{\lambda \neq 1 \Rightarrow \sum_i \xi_i^{(\lambda\neq 1)} = 0}$

Approfondiamo il PROBLEMA agli AUTOVALORI/AUTOVETTORI per una W MATRICE STOCASTICA, studiando separatamente i 2 CASI: $\lambda = 1$ o $\lambda \neq 1$.

Per $\underline{\lambda=1}$ abbiamo $W \xi^{(1)} = \xi^{(1)}$ (ricordiamo che $\xi^{(1)}$ è AUTOVETT. corrisp. allo AUTOVALORE λ ...)

In componenti $\sum_j W_{ij} \xi_j^{(1)} = \xi_i^{(1)}$

che in modulo è $\left|\sum_j W_{ij} \xi_j^{(1)}\right| \leq \sum_j W_{ij} |\xi_j^{(1)}|$ ovvero $\left|\sum_j W_{ij} \xi_j^{(1)}\right| = |\xi_i^{(1)}| \leq \sum_j W_{ij} |\xi_j^{(1)}|$

il modulo di una SOMMA è $\leq$ della SOMMA dei MODULI ed in più $|W_{ij}| = W_{ij} \geq 0$ ...

relazione che vale $\forall i \Rightarrow$ posso SOMMARE su tutti gli $i \Longrightarrow \sum_i |\xi_i^{(1)}| \leq \sum_{ij} W_{ij} |\xi_j^{(1)}|$

ma ATTENZIONE! $\sum_{ij} W_{ij} |\xi_j^{(1)}| = \sum_j \left(\sum_i W_{ij}\right) |\xi_j^{(1)}| = \sum_j |\xi_j^{(1)}| = \sum_i |\xi_i^{(1)}|$

$\sum_i W_{ij} = 1 \;\; \forall j$ — perdonate la banalità... indice muto...

da cui ottengo che $\sum_i |\xi_i^{(1)}| \leq \sum_i |\xi_i^{(1)}|$ ... che è manifestamente VERA per il segno = ...

MORALE In $\leq$ valeva = fin dall'inizio ... ovvero $\left|\sum_j W_{ij} \xi_j^{(1)}\right| = \sum_j W_{ij} |\xi_j^{(1)}|$

ma IL MODULO di una SOMMA è UGUALE alla SOMMA dei MODULI quando TUTTI GLI ADDENDI sono CONCORDI in SEGNO

Dato allora il vettore $\xi^{(1)}$ posso DEFINIRE $\pi \equiv \xi^{(1)} / \sum_i \xi_i^{(1)}$ che è tale che

$\boxed{\forall i \;\; \pi_i \geq 0 \;\; \text{e} \;\; \sum_i \pi_i = 1}$ essendo sempre AUTOVETT. di AUTOVALORE 1 ovvero $W\pi = \pi$

NB In generale è sempre vero che $W \xi^{(\lambda)} = \lambda \xi^{(\lambda)} \Rightarrow W (c\, \xi^{(\lambda)}) = \lambda (c\, \xi^{(\lambda)})$ con c COSTANTE
(l'autovettore individua una DIREZIONE! il modulo è arbitrario...)

Abbiamo imparato che l'AUTOVETTORE corrispondente all'AUTOVALORE 1 può essere RISCALATO e INTERPRETATO come una distribuzione di PROBABILITA'!

Per $\lambda \neq 1$ viceversa $W \xi^{(\lambda \neq 1)} = \lambda \xi^{(\lambda \neq 1)}$

e in componenti $\sum_j W_{ij} \xi_j^{(\lambda \neq 1)} = \lambda\, \xi_i^{(\lambda \neq 1)}$

ancora MOD. di SOMMA è $\leq$ SOMMA dei MODULI...

e ancora $\left| \sum_j W_{ij} \xi_j^{(\lambda \neq 1)} \right| = |\lambda| \, |\xi_i^{(\lambda \neq 1)}| \leq \sum_j W_{ij} \, |\xi_j^{(\lambda \neq 1)}|$

ancora posso SOMMARE su i $\Rightarrow$ $|\lambda| \sum_i |\xi_i^{(\lambda \neq 1)}| \leq \sum_{ji} W_{ij} |\xi_j^{(\lambda \neq 1)}| = \sum_j |\xi_j^{(\lambda \neq 1)}|$

come prima $\sum_i W_{ij} = 1$...

ma allora $|\lambda| \sum_i |\xi_i^{(\lambda \neq 1)}| \leq \sum_i |\xi_i^{(\lambda \neq 1)}|$ che mi dice $\boxed{|\lambda| \leq 1}$

RICAPITOLANDO

$W \xi^{(\lambda)} = \lambda \xi^{(\lambda)}$ può avere SOLUZIONI

– $\boxed{\lambda = 1}$ con $\Pi = \dfrac{\xi^{(1)}}{\sum_i \xi_i^{(1)}}$ t.c. $\boxed{\Pi_i \geq 0 \text{ e } \sum_i \Pi_i = 1 \quad W\Pi = \Pi}$

Π interpretabile come DISTRIBUZ. di PROBABILITA'

– se $\lambda \neq 1 \rightarrow \boxed{|\lambda| < 1}$ con $\boxed{\sum_i \xi_i^{(\lambda \neq 1)} = 0}$ e dunque NON riscalabile ad avere una DISTRIBUZ. di PROB.

DOMANDA Ora ci concentreremo su un'altra proprietà: $W^N_{ij} \Rightarrow \Pi_i$
Questa sembra una proprietà non troppo parente di quanto visto fino ad ora... Provate allora a riconsiderare la nostra matrice W.
Che succede se calcolate (con Matlab) la potenza W^N, per N >> 1?

Il risultato generale è che

- SPAZIO delle CONFIGURAZIONI FINITO-DIMENSIONALE $\Rightarrow$ $\exists\, \pi : W\pi = \pi$ e $W^N_{ij} \xrightarrow[N\to\infty]{} \pi_i \;\; \forall j$

  interpretabile come PROBABILITA'

  se non ci sono altri autovalori $|\lambda| = 1$...

- SPAZIO delle CONFIGURAZIONI INFINITO-DIMENSIONALE $\Rightarrow$ se $\exists\, \pi : W^N_{ij} \xrightarrow[N\to\infty]{} \pi_i \;\; \forall j$

  tale π è UNICO e $W\pi = \pi$

  interpretabile come PROB.

ESERCIZIO

Provate a mostrare che $W\pi = \pi$ e $|\lambda| < 1$ per $\lambda \neq 1$ implica $W^N_{ij} \xrightarrow[N\to\infty]{} \pi_i$

NOTA BENE Abbiamo già imparato la

INTERPRETAZIONE di W^N_{ij} = Prob. di transizione $i \leftarrow j$ in N PASSI

ma allora $\forall j \;\; W^N_{ij} \xrightarrow[N\to\infty]{} \pi_i$ (*) vuol dire che per $N\to\infty$ INDIPEND. da dove parto ($\forall j$...) arrivo su i con PROBABILITA' π_i

π è detta allora PROBABILITA' ASINTOTICA o STAZIONARIA

(*) questo vuol anche dire che $\forall N \geq \tilde{N} \quad W^N_{ij} \simeq \pi_i$ ovvero da un certo punto in poi il processo è distribuito come π! Dopo un certo tempo visito le configurazioni in accordo a π!

$\rightarrow$ Vediamo ora questa cosa da un altro punto di vista, ricordando che

$$p^{(N)} = W^N p^{(0)}$$

Consideriamo il nostro caso ... abbiamo trovato $\lambda = 1$ e poi tutti i $\lambda \neq 1$ erano $|\lambda| < 1$.
In più: abbiamo trovato una BASE di AUTOVETTORI (che è come dire che W è risultata essere DIAGONALIZZABILE)

Siccome l'insieme $\{\Pi, \xi^{(\lambda \neq 1)}\}$ è una BASE, ogni vettore è esprimibile come...

(attenzione! intendo l'insieme di TUTTI gli AUTOVETTORI, fra cui metto in evidenza Π, che corrisponde a $\lambda \neq 1$.

... UNA COMBINAZIONE LINEARE $c_1 \Pi + \sum_{\lambda \neq 1} c_\lambda \xi^{(\lambda)}$, ed in particolare

$$p^{(0)} = c_1 \Pi + \sum_{\lambda \neq 1} c_\lambda \xi^{(\lambda)}$$ ... in componenti $$p^{(0)}_i = c_1 \Pi_i + \sum_{\lambda \neq 1} c_\lambda \xi^{(\lambda)}_i$$

ma $\sum_i p_i^{(0)} = 1$ ($p^{(0)}$ è una probabilità, ... per la precisione quella al tempo 0 ...)

ovvero $$\sum_i p_i^{(0)} = 1 = c_1 \sum_i \Pi_i + \sum_{\lambda \neq 1} c_\lambda \sum_i \xi_i^{(\lambda \neq 1)}$$

ma ATTENZIONE! $\sum_i \Pi_i = 1$ mentre avevamo visto che $\sum_i \xi_i^{(\lambda \neq 1)} = 0$!

MORALE $1 = c_1$ ovvero OGNI PROB. INIZIALE ha COMPONENTE 1 lungo Π!

e dunque $$p^{(0)} = \Pi + \sum_{\lambda \neq 1} c_\lambda \xi^{(\lambda)}$$

Ora $$W^N p^{(0)} = p^{(N)} = W^N \Pi + \sum_{\lambda \neq 1} c_\lambda W^N \xi^{(\lambda)}$$

... ma quanto fa $W^N \xi^{(\lambda)}$ se $W \xi^{(\lambda)} = \lambda \xi^{(\lambda)}$?

Vale $$W^2 \xi^{(\lambda)} = W(W \xi^{(\lambda)}) = W(\lambda \xi^{(\lambda)}) = \lambda W \xi^{(\lambda)} = \lambda^2 \xi^{(\lambda)}$$

ma allora $$W^3 \xi^{(\lambda)} = W(W^2 \xi^{(\lambda)}) = W \lambda^2 \xi^{(\lambda)} = \lambda^2 W \xi^{(\lambda)} = \lambda^3 \xi^{(\lambda)}$$

e in generale $$\underline{W^N \xi^{(\lambda)} = \lambda^N \xi^{(\lambda)}}$$

$1^N = 1$ ...

così che $$p^{(N)} = W^N p^{(0)} = W^N \Pi + \sum_{\lambda \neq 1} c_\lambda W^N \xi^{(\lambda)} = \Pi + \sum_{\lambda \neq 1} c_\lambda \lambda^N \xi^{(\lambda)}$$

Ma per noi tutti i $\lambda \neq 1$ erano $|\lambda| < 1$ e dunque per $\lambda \neq 1$ $\lambda^N \underset{N \to \infty}{\longrightarrow} 0$!

Questo vuol dire $\boxed{P^{(N)} \underset{N \to \infty}{\longrightarrow} \Pi}$ Cioè il PROCESSO è ASINTOTICAMENTE DISTRIBUITO IN ACCORDO a Π !

... e Π merita il nome di DISTRIBUZIONE STAZIONARIA ...

... il che ci fa anche capire PERCHE' $\frac{1}{N} \sum_{m=1}^{N} m_3(m) \simeq 1{,}8$ per $N \gg 1$

dove questa è la MEDIA ARITMETICA dei VALORI di m_3 (# N in U3 ...) "lungo il PROCESSO" ($m_3(m)$ è il valore di m_3 al "tempo" m ...)

Siccome asintoticamente $P^{(N)} \to \Pi$ convergiamo a $\sum_{i=1}^{3} i \cdot \Pi_i = 1{,}8$ ...

NB Ricorda $\Pi_1 = 0{,}3$
$\Pi_2 = 0{,}6$
$\Pi_3 = 0{,}1$

Questa convergenza è dunque null'altro che l'ennesima conferma (numerica...) che le medie aritmetiche su osservazioni di un fenomeno aleatorio convergono ai valori medi rispetto alla distribuzione di probabilità dello stesso (mentre le frequenze - che sono medie della variabile caratteristica - convergono alle probabilità ...)