

Filtraggio della Matrice di Correlazione

Relatore: Prof. Stefano Marmi

Studente: Marco Miani

Scuola Normale Superiore

30 aprile 2019

Introduzione

Sia $(\Omega, \mathcal{P}, \mathcal{F})$ un qualsiasi spazio di probabilità e sia $X : \Omega \mapsto \mathbb{R}^n$ una V.A vettoriale su tale spazio.

$$X = (X_1, \dots, X_n)$$

Introduzione

Sia $(\Omega, \mathcal{P}, \mathcal{F})$ un qualsiasi spazio di probabilità e sia $X : \Omega \mapsto \mathbb{R}^n$ una V.A vettoriale su tale spazio.

$$X = (X_1, \dots, X_n)$$

Consideriamo t eventi $\omega_1, \dots, \omega_t$ e prendiamone le immagini tramite X e mettiamole come colonne di una matrice

$$M = \left(\begin{array}{c|c|c} X_1(\omega_1) & \cdots & X_1(\omega_t) \\ \vdots & & \vdots \\ X_n(\omega_1) & \cdots & X_n(\omega_t) \end{array} \right) \in \mathfrak{M}(n, t, \mathbb{R})$$

che chiameremo matrice delle realizzazioni.

Covarianza

$$(\Sigma)_{i,j} := \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] = \int_{\Omega} X_i(\omega) X_j(\omega) dP(\omega)$$

Introduzione

Covarianza

$$(\Sigma)_{i,j} := \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] = \int_{\Omega} X_i(\omega) X_j(\omega) dP(\omega)$$

Covarianza "discreta"

$$(C)_{i,j} := \frac{(MM^T)_{i,j}}{t} = \frac{1}{t} \sum_{s=1}^t M_{i,s} M_{j,s} = \frac{1}{t} \sum_{s=1}^t X_i(\omega_s) X_j(\omega_s)$$

Introduzione

Covarianza

$$(\Sigma)_{i,j} := \text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] = \int_{\Omega} X_i(\omega) X_j(\omega) dP(\omega)$$

Covarianza "discreta"

$$(C)_{i,j} := \frac{(MM^T)_{i,j}}{t} = \frac{1}{t} \sum_{s=1}^t M_{i,s} M_{j,s} = \frac{1}{t} \sum_{s=1}^t X_i(\omega_s) X_j(\omega_s)$$

La legge dei grandi numeri ci dice che

$$\lim_{t \rightarrow \infty} \frac{(MM^T)_{i,j}}{t} = (\Sigma)_{i,j}$$

E quindi per t abbastanza grande si ha $\Sigma \approx C$

Kelly e Markowitz

L'importanza della matrice Σ risulta ancor più evidente dopo aver visto alcuni risultati

L'importanza della matrice Σ risulta ancor più evidente dopo aver visto alcuni risultati

Criterio di Kelly

Data una certa scommessa e un certo capitale, ci chiediamo quale sia la percentuale f^* del capitale da investire che massimizza il guadagno.

$$f_{ottimale}^* = \arg \min \{ \mathbb{E}[\log(return(f^*))] \mid 0 \leq f^* \leq 1 \}$$

Kelly e Markowitz

L'importanza della matrice Σ risulta ancor più evidente dopo aver visto alcuni risultati

Criterio di Kelly

Data una certa scommessa e un certo capitale, ci chiediamo quale sia la percentuale f^* del capitale da investire che massimizza il guadagno.

$$f_{ottimale}^* = \arg \min \{ \mathbb{E}[\log(return(f^*))] \mid 0 \leq f^* \leq 1 \}$$

Proiettile di Markowitz

Dato un mercato, ci chiediamo quale sia il portfolio W di rischio minimo con un dato ritorno μ

$$W_{ottimale}(\mu) = \arg \min \left\{ W\Sigma W^T \mid \mu = MW \right\}$$

Filtraggio

Se Σ è ignota e conosciamo solo C ? (e.g. corse di cavalli, stock in borsa...)

Filtraggio

Se Σ è ignota e conosciamo solo C ? (e.g. corse di cavalli, stock in borsa...)

Definizione

Un filtraggio nella sua forma più generale è un qualsiasi operatore

$$(filt) : \mathfrak{M}(n, \mathbb{R}) \rightarrow \mathfrak{M}(n, \mathbb{R})$$

Noi considereremo solo filtraggi tra matrici simmetriche definite positive

Chiameremo:

-filtraggio nullo ($filt_{nullo}$) := $Id_{\mathfrak{M}(n, \mathbb{R})}$.

-filtraggio ideale quello tale che $C^{(filt_{ideale})} = \Sigma$ $\forall C$ realizzazione.

Filtraggio

Se Σ è ignota e conosciamo solo C ? (e.g. corse di cavalli, stock in borsa...)

Definizione

Un filtraggio nella sua forma più generale è un qualsiasi operatore

$$(filt) : \mathfrak{M}(n, \mathbb{R}) \rightarrow \mathfrak{M}(n, \mathbb{R})$$

Noi considereremo solo filtraggi tra matrici simmetriche definite positive

Chiameremo:

- filtraggio nullo ($filt_{nullo}$) := $Id_{\mathfrak{M}(n, \mathbb{R})}$.
- filtraggio ideale quello tale che $C^{(filt_{ideale})} = \Sigma$ $\forall C$ realizzazione.

Esamineremo 4 tipi di filtraggio:

SLCA

ALCA

su grafi

Rosenow

Potters

spettrali

Filtraggio con metodi su grafi

Si costruisce un grafo completo pesato (N, E, W) .

- I nodi $N = \{1, \dots, n\}$ corrispondono alle componenti della V.A. X
- Gli archi E sono tutte le possibili coppie di nodi (grafo completo)
- Come pesi assegnamo quelli naturali $W(i,j) := C_{i,j}$

Filtraggio con metodi su grafi

Si costruisce un grafo completo pesato (N, E, W) .

- I nodi $N = \{1, \dots, n\}$ corrispondono alle componenti della V.A. X
- Gli archi E sono tutte le possibili coppie di nodi (grafo completo)
- Come pesi assegnamo quelli naturali $W(i,j) := C_{i,j}$

Si selezionano gli archi più "significativi" costruendo il Maximum Spanning Tree relativo al grafo.

Filtraggio con metodi su grafi

Si costruisce un grafo completo pesato (N, E, W) .

- I nodi $N = \{1, \dots, n\}$ corrispondono alle componenti della V.A. X
- Gli archi E sono tutte le possibili coppie di nodi (grafo completo)
- Come pesi assegnamo quelli naturali $W(i,j) := C_{i,j}$

Si selezionano gli archi più "significativi" costruendo il Maximum Spanning Tree relativo al grafo.

Mentre procede l'algoritmo sul grafo, al passo s aggiorniamo la relativa matrice $C^{(s)}$ in modo opportuno.

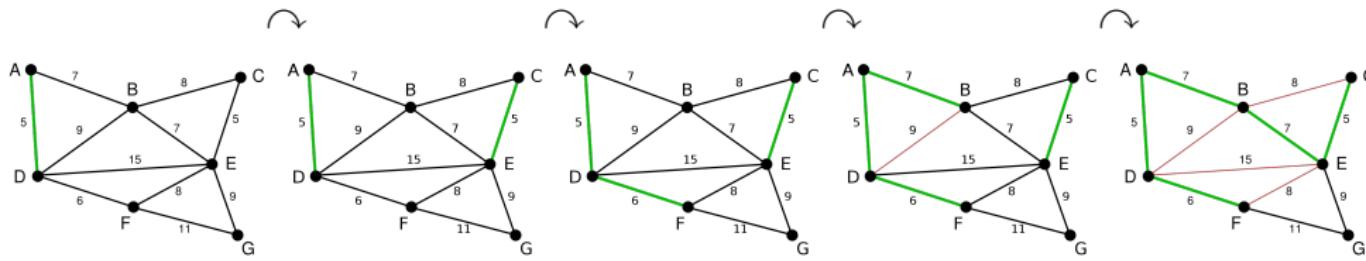
Filtraggio con metodi su grafi

MST - Algoritmo di Kruskal

All'inizio ci sono n cluster, ciascuno composto da un singolo nodo, e poi si procede iterativamente.

- Si seleziona l'arco di peso maggiore.
- Si uniscono i due cluster di nodi che sono collegati da tale arco.
- Si eliminano gli archi interni ai cluster.

Dopo $n - 1$ iterazioni si ha solo un cluster e l'algoritmo termina.



Filtraggio con metodi su grafi

Filtraggio 1 - SLCA (Med)

Al passo s , quando uniamo i cluster H e K nel cluster Q poniamo

$$C_{i,j}^{(s+1)} = \begin{cases} \frac{|H| C_{H,j}^{(s)} + |K| C_{K,j}^{(s)}}{|H| + |K|} & \text{se } i \in Q \quad j \notin H \cup K \\ C_{i,j}^{(s)} & \text{altrimenti} \end{cases}$$

Filtraggio con metodi su grafi

Filtraggio 1 - SLCA (Med)

Al passo s , quando uniamo i cluster H e K nel cluster Q poniamo

$$C_{i,j}^{(s+1)} = \begin{cases} \frac{|H| C_{H,j}^{(s)} + |K| C_{K,j}^{(s)}}{|H| + |K|} & \text{se } i \in Q \quad j \notin H \cup K \\ C_{i,j}^{(s)} & \text{altrimenti} \end{cases}$$

Filtraggio 2 - ALCA (Max)

Al passo s , quando uniamo i cluster H e K nel cluster Q poniamo

$$C_{i,j}^{(s+1)} = \begin{cases} \max\{C_{H,j}^{(s)}, C_{K,j}^{(s)}\} & \text{se } i \in Q \quad j \notin H \cup K \\ C_{i,j}^{(s)} & \text{altrimenti} \end{cases}$$

Filtraggio con metodi su grafi

Filtraggio 1 - SLCA (Med)

Al passo s , quando uniamo i cluster H e K nel cluster Q poniamo

$$C_{i,j}^{(s+1)} = \begin{cases} \frac{|H| C_{H,j}^{(s)} + |K| C_{K,j}^{(s)}}{|H| + |K|} & \text{se } i \in Q \quad j \notin H \cup K \\ C_{i,j}^{(s)} & \text{altrimenti} \end{cases}$$

Filtraggio 2 - ALCA (Max)

Al passo s , quando uniamo i cluster H e K nel cluster Q poniamo

$$C_{i,j}^{(s+1)} = \begin{cases} \max\{C_{H,j}^{(s)}, C_{K,j}^{(s)}\} & \text{se } i \in Q \quad j \notin H \cup K \\ C_{i,j}^{(s)} & \text{altrimenti} \end{cases}$$

Alla fine si pone $C^{(filt)} = C^{(n-1)}$

Filtraggio con metodi spettrali

Si diagonalizza C tramite matrici ortogonali $C = V^T D V$ e si modifica la matrice D , operando sugli autovalori più piccoli ($\leq \lambda_{max}$ fissato).

$$D \rightsquigarrow \tilde{D}$$

Filtraggio con metodi spettrali

Si diagonalizza C tramite matrici ortogonali $C = V^T D V$ e si modifica la matrice D , operando sugli autovalori più piccoli ($\leq \lambda_{max}$ fissato).

$$D \rightsquigarrow \tilde{D}$$

Si rimoltiplica, ottenendo $\tilde{C} := V^T \tilde{D} V$ e si modifica \tilde{C} per assicurarsi che il risultato sia ancora una matrice di correlazione.

$$\tilde{C} \rightsquigarrow C^{(filt)}$$

Filtraggio con metodi spettrali

Filtraggio 3 - Rosenow

Si sostituiscono gli elementi di D più piccoli di λ_{max} con 0

$$\tilde{D}_{i,i} = \begin{cases} D_{i,i} & \text{se } D_{i,i} \geq \lambda_{max} \\ 0 & \text{se } D_{i,i} < \lambda_{max} \end{cases}$$

Si forzano gli elementi della diagonale a 1

$$C_{i,j}^{(filt)} = \delta_{i,j} + (1 - \delta_{i,j}) \tilde{C}_{i,j}$$

Filtraggio con metodi spettrali

Filtraggio 3 - Rosenow

Si sostituiscono gli elementi di D più piccoli di λ_{max} con 0

$$\tilde{D}_{i,i} = \begin{cases} D_{i,i} & \text{se } D_{i,i} \geq \lambda_{max} \\ 0 & \text{se } D_{i,i} < \lambda_{max} \end{cases}$$

Si forzano gli elementi della diagonale a 1

$$C_{i,j}^{(filt)} = \delta_{i,j} + (1 - \delta_{i,j}) \tilde{C}_{i,j}$$

Filtraggio 4 - Potters

Si sostituiscono gli elementi di D più piccoli di λ_{max} con la loro media

$$\tilde{D}_{i,i} = \begin{cases} D_{i,i} & \text{se } D_{i,i} \geq \lambda_{max} \\ \mathbb{E}[\lambda_j : \lambda_j < \lambda_{max}] & \text{se } D_{i,i} < \lambda_{max} \end{cases}$$

Si pone scalano gli elementi di \tilde{C} con le radici degli elementi diagonali

$$C_{i,j}^{(filt)} = \frac{\tilde{C}_{i,j}}{\sqrt{\tilde{C}_{i,i} \tilde{C}_{j,j}}}$$

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$
- Calcoliamo le matrici di covarianza $C_1, \dots, C_{1000} \in \mathfrak{M}(n, \mathbb{R})$ associate alle rispettive realizzazioni

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$
- Calcoliamo le matrici di covarianza $C_1, \dots, C_{1000} \in \mathfrak{M}(n, \mathbb{R})$ associate alle rispettive realizzazioni
- Scegliamo un filtraggio e applichiamolo, ottenendo $C_1^{(filt)}, \dots, C_{1000}^{(filt)}$

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$
- Calcoliamo le matrici di covarianza $C_1, \dots, C_{1000} \in \mathfrak{M}(n, \mathbb{R})$ associate alle rispettive realizzazioni
- Scegliamo un filtraggio e applichiamolo, ottenendo $C_1^{(filt)}, \dots, C_{1000}^{(filt)}$

Σ

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$
- Calcoliamo le matrici di covarianza $C_1, \dots, C_{1000} \in \mathfrak{M}(n, \mathbb{R})$ associate alle rispettive realizzazioni
- Scegliamo un filtraggio e applichiamolo, ottenendo $C_1^{(filt)}, \dots, C_{1000}^{(filt)}$

$$\begin{array}{ccc} & C_1 & \\ \nearrow & \vdots & \\ \Sigma & \rightarrow & C_i \\ \searrow & \vdots & \\ & C_{1000} & \end{array}$$

Stima bontà del filtraggio

- Fissiamo una V.A. vettoriale con matrice di covarianza $\Sigma \in \mathfrak{M}(n, \mathbb{R})$ simmetrica definita positiva
- Costruiamo tante (~ 1000) sue realizzazioni indipendenti $M_1, \dots, M_{1000} \in \mathfrak{M}(n, t, \mathbb{R})$
- Calcoliamo le matrici di covarianza $C_1, \dots, C_{1000} \in \mathfrak{M}(n, \mathbb{R})$ associate alle rispettive realizzazioni
- Scegliamo un filtraggio e appliciamolo, ottenendo $C_1^{(filt)}, \dots, C_{1000}^{(filt)}$

$$\begin{array}{ccc} C_1 & \rightarrow & C_1^{(filt)} \\ \nearrow & \vdots & \vdots \\ \Sigma & \rightarrow & C_i & \rightarrow & C_i^{(filt)} \\ \searrow & \vdots & \vdots \\ & & C_{1000} & \rightarrow & C_{1000}^{(filt)} \end{array}$$

Stima bontà del filtraggio

Fissiamo una distanza tra matrici $d : \mathfrak{M}(n, \mathbb{R}) \times \mathfrak{M}(n, \mathbb{R}) \rightarrow \mathbb{R}^+$.

Stima bontà del filtraggio

Fissiamo una distanza tra matrici $d : \mathfrak{M}(n, \mathbb{R}) \times \mathfrak{M}(n, \mathbb{R}) \rightarrow \mathbb{R}^+$.

Per capire quanto un $(filt) \approx (filt_{ideale})$ possiamo misurare:

- ➊ Quanto $C^{(filt)}$ si avvicina a Σ ?

$$\langle d(\Sigma, C^{(filt)}) \rangle$$

- ➋ Quanta informazione di C viene persa/trattenuta in $C^{(filt)}$?

$$\langle d(C, C^{(filt)}) \rangle$$

- ➌ Quanto è stabile il filtraggio?

$$\langle d(C^{(filt)}, C^{(filt)}) \rangle$$

Stima bontà del filtraggio

Fissiamo una distanza tra matrici $d : \mathfrak{M}(n, \mathbb{R}) \times \mathfrak{M}(n, \mathbb{R}) \rightarrow \mathbb{R}^+$.

Per capire quanto un $(filt) \approx (filt_{ideale})$ possiamo misurare:

- ➊ Quanto $C^{(filt)}$ si avvicina a Σ ?

$$\langle d(\Sigma, C^{(filt)}) \rangle := \frac{1}{1000} \sum_{i=1}^{1000} d(\Sigma, C_i^{(filt)})$$

- ➋ Quanta informazione di C viene persa/trattenuta in $C^{(filt)}$?

$$\langle d(C, C^{(filt)}) \rangle := \frac{1}{1000} \sum_{i=1}^{1000} d(C, C_i^{(filt)})$$

- ➌ Quanto è stabile il filtraggio?

$$\langle d(C^{(filt)}, C^{(filt)}) \rangle := \frac{1}{499500} \sum_{\substack{i,j=1 \\ i \neq j}}^{1000} d(C_i^{(filt)}, C_j^{(filt)})$$

Stima bontà del filtraggio

$$\bullet^{\Sigma}$$

Stima bontà del filtraggio

• C_1

• Σ

Stima bontà del filtraggio

• C_1

• C_2

• C_3

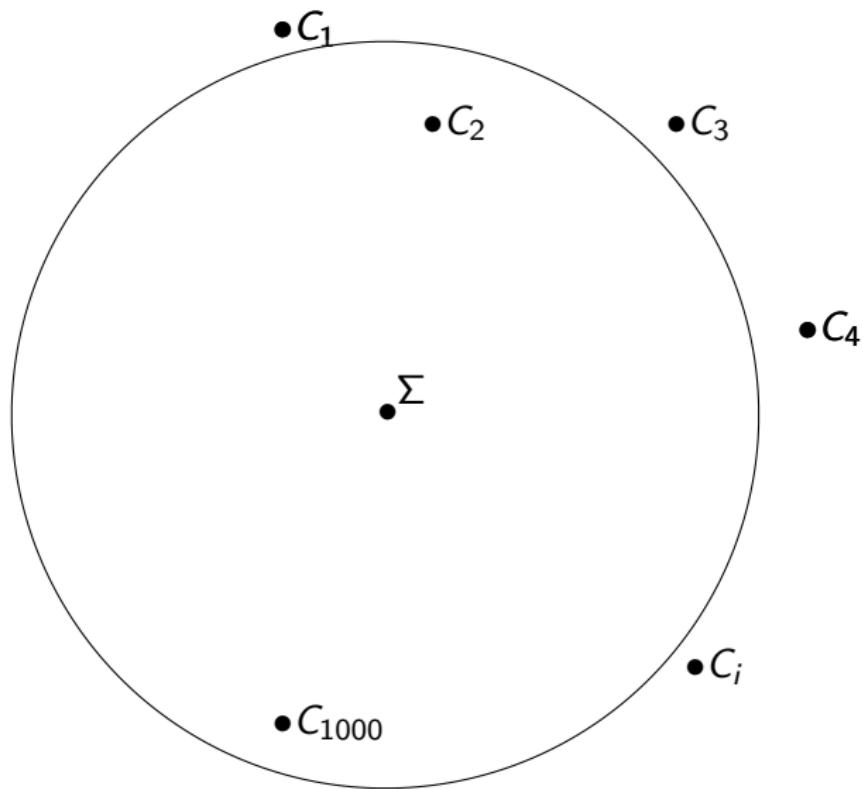
• C_4

• Σ

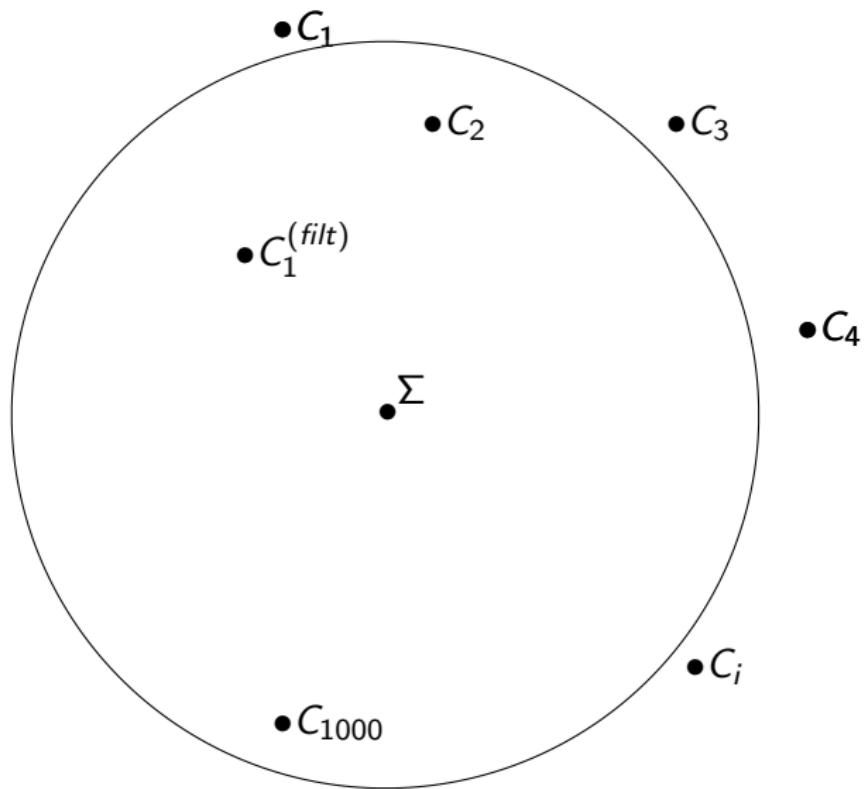
• C_i

• C_{1000}

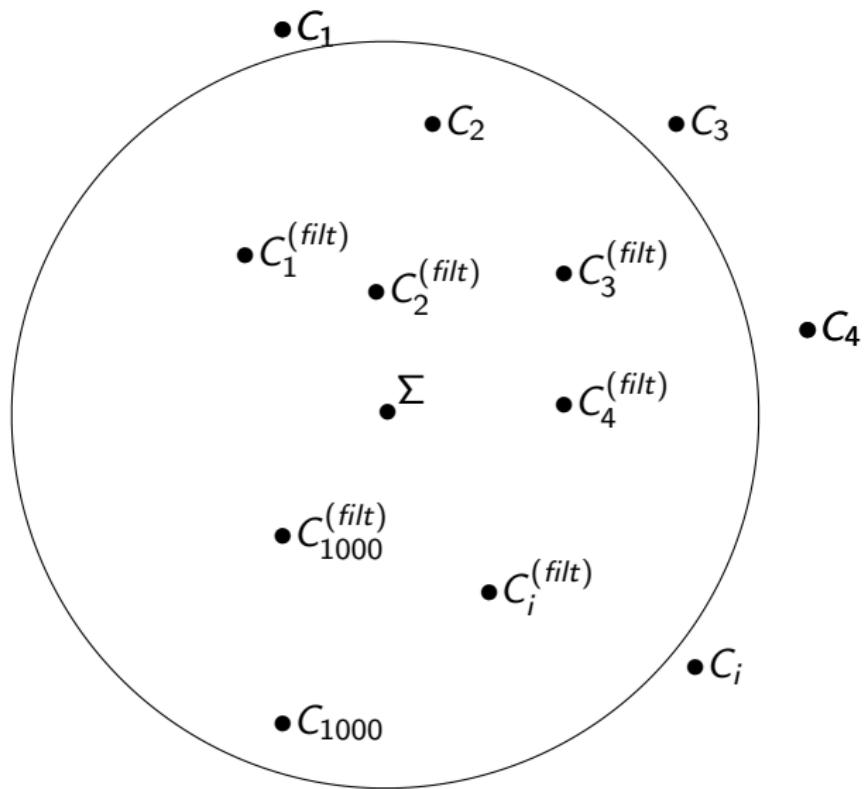
Stima bontà del filtraggio



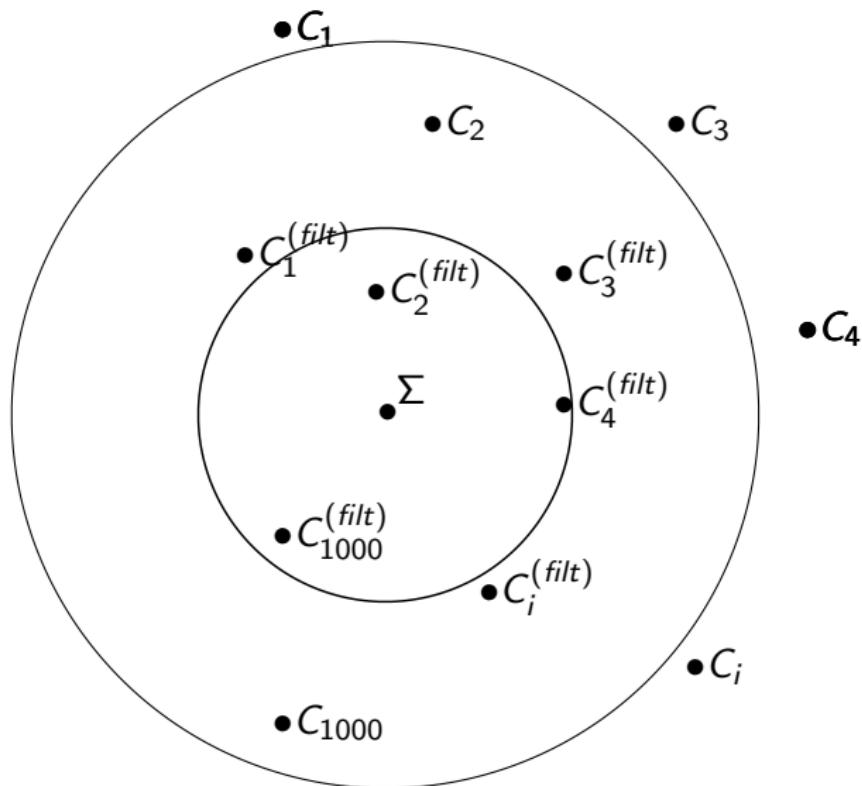
Stima bontà del filtraggio



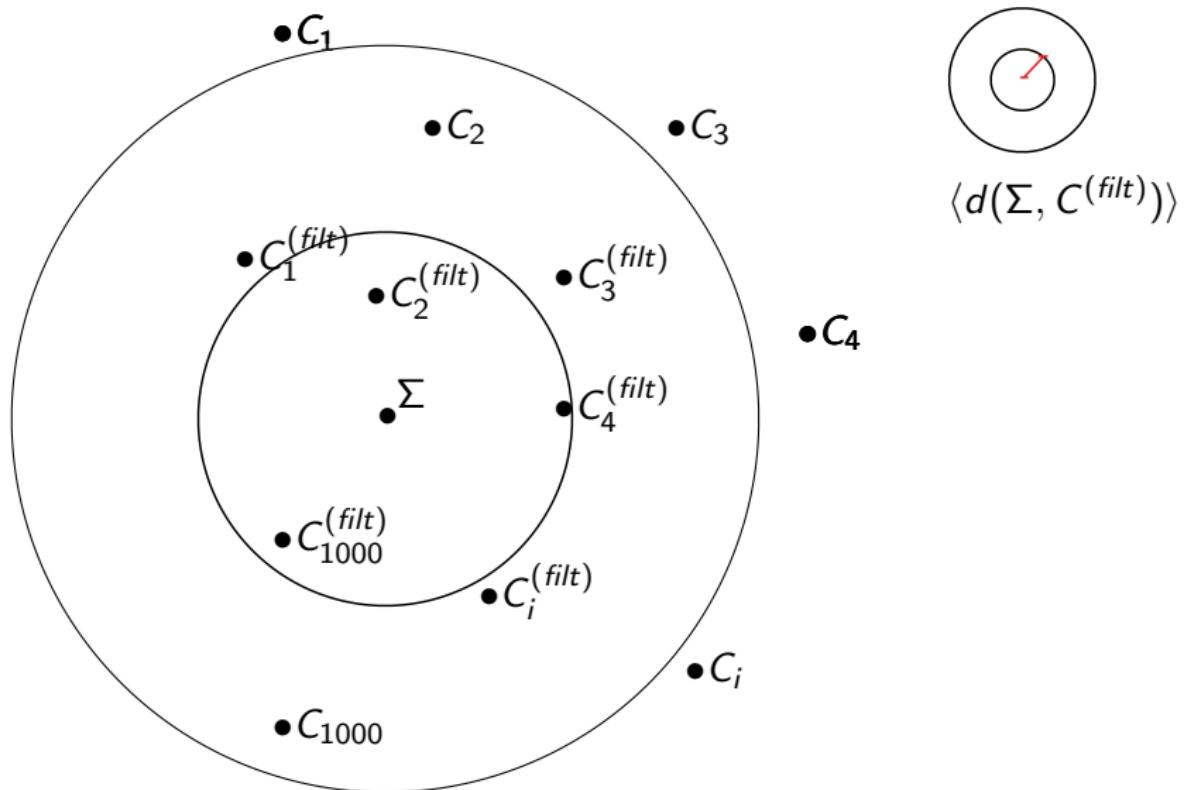
Stima bontà del filtraggio



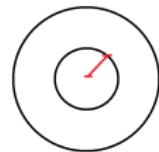
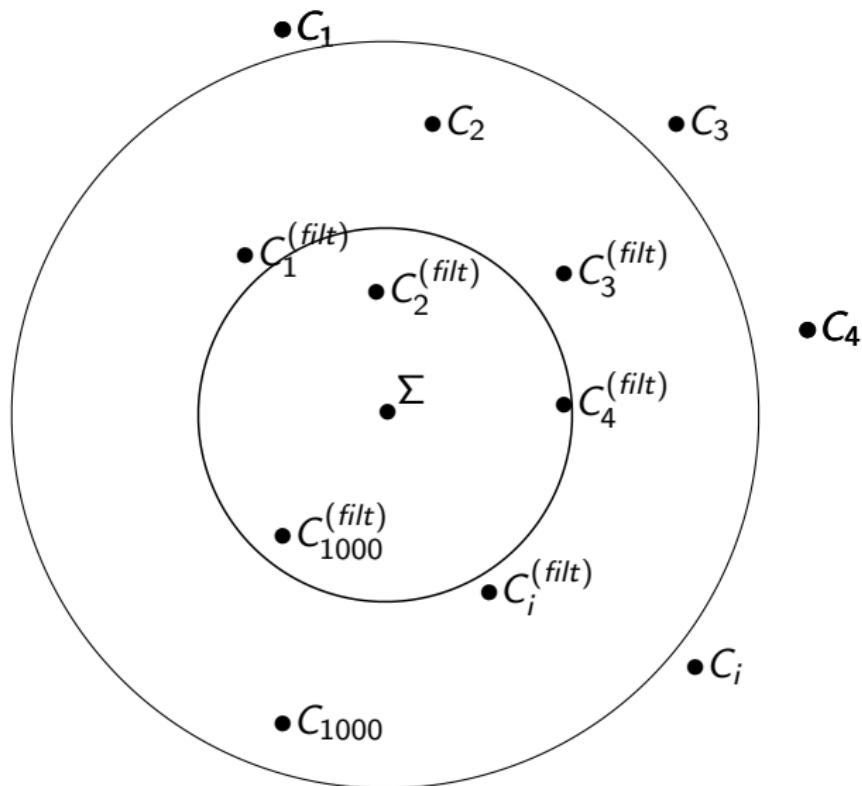
Stima bontà del filtraggio



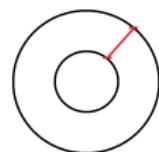
Stima bontà del filtraggio



Stima bontà del filtraggio

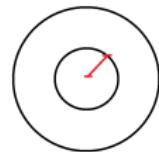
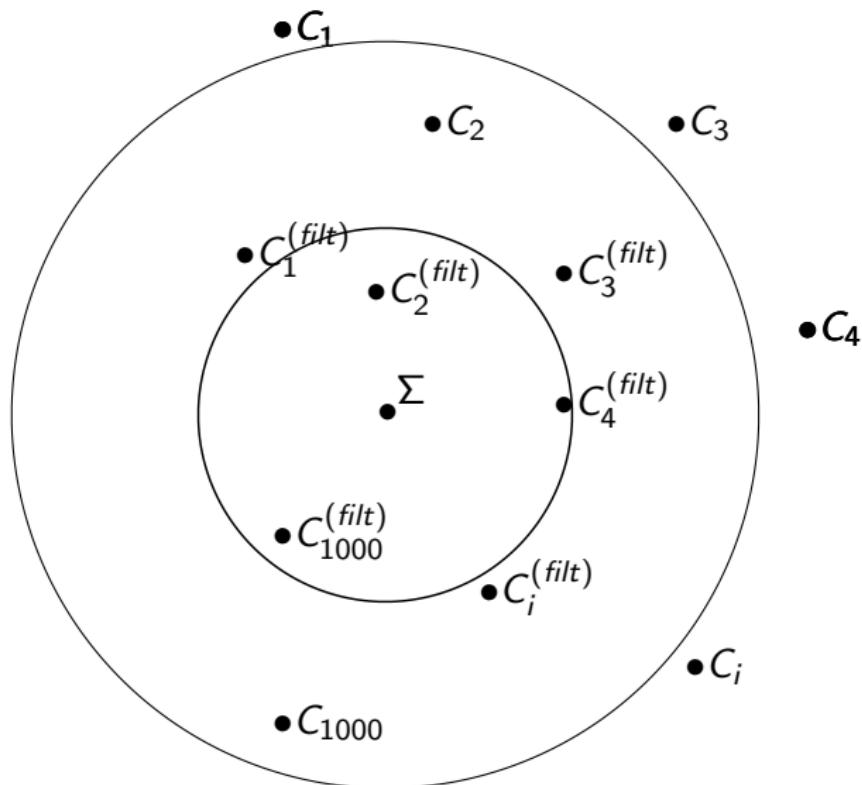


$$\langle d(\Sigma, C^{filt}) \rangle$$

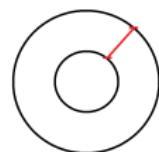


$$\langle d(C, C^{filt}) \rangle$$

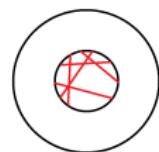
Stima bontà del filtraggio



$$\langle d(\Sigma, C^{filt}) \rangle$$



$$\langle d(C, C^{filt}) \rangle$$



$$\langle d(C^{filt}, C^{filt}) \rangle$$

Stima bontà del filtraggio

Se la matrice Σ ci è ignota:

① $\langle d(\Sigma, C^{(filt)}) \rangle$ NON calcolabile

② $\langle d(C, C^{(filt)}) \rangle$ calcolabile

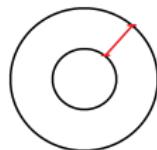
③ $\langle d(C^{(filt)}, C^{(filt)}) \rangle$ calcolabile

Stima bontà del filtraggio

Se la matrice Σ ci è ignota:

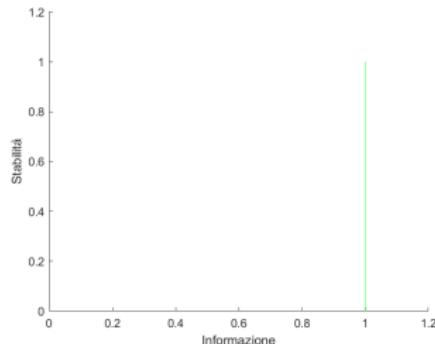
① $\langle d(\Sigma, C^{(filt)}) \rangle$ **NON calcolabile**

② $\langle d(C, C^{(filt)}) \rangle$ **calcolabile**



$$0 = \underbrace{\langle d(C, C) \rangle}_{\text{filtraggio nullo}} \leq \langle d(C, C^{(filt)}) \rangle \leq \underbrace{\langle d(C, \Sigma) \rangle}_{\text{filtraggio perfetto}} = ?$$

③ $\langle d(C^{(filt)}, C^{(filt)}) \rangle$ **calcolabile**

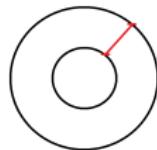


Stima bontà del filtraggio

Se la matrice Σ ci è ignota:

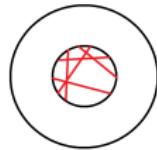
① $\langle d(\Sigma, C^{(filt)}) \rangle$ **NON calcolabile**

② $\langle d(C, C^{(filt)}) \rangle$ **calcolabile**

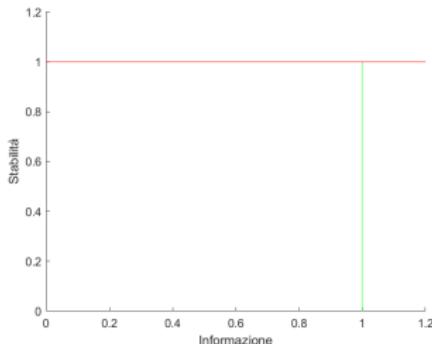


$$0 = \underbrace{\langle d(C, C) \rangle}_{\text{filtraggio nullo}} \leq \langle d(C, C^{(filt)}) \rangle \leq \underbrace{\langle d(C, \Sigma) \rangle}_{\text{filtraggio perfetto}} = ?$$

③ $\langle d(C^{(filt)}, C^{(filt)}) \rangle$ **calcolabile**



$$0 = \underbrace{\langle d(\Sigma, \Sigma) \rangle}_{\text{filtraggio perfetto}} \leq \langle d(C_i^{(filt)}, C_j^{(filt)}) \rangle \leq \underbrace{\langle d(C_i, C_j) \rangle}_{\text{filtraggio nullo}} = ?$$



Distanza di Kullback-Leibler

Definizione

Date due densità di probabilità P e Q si definisce:

$$K(P, Q) := \mathbb{E} \left[\ln \left(\frac{P}{Q} \right) \right] = \int \ln \left(\frac{P(x)}{Q(x)} \right) dP(x)$$

la distanza di Kullback-Leibler.

Distanza di Kullback-Leibler

Definizione

Date due densità di probabilità P e Q si definisce:

$$K(P, Q) := \mathbb{E} \left[\ln \left(\frac{P}{Q} \right) \right] = \int \ln \left(\frac{P(x)}{Q(x)} \right) dP(x)$$

la distanza di Kullback-Leibler.

Proprietà:

- $K(P, P) = 0$
- $K(P, Q) \geq 0$
 $-K(P, Q) = \int \log\left(\frac{Q}{P}\right)P dx \leq \log\left(\int \frac{Q}{P}P dx\right) = \log(1) = 0$
- **Non** è simmetrica
 $K_{sim}(P, Q) := K(P, Q) + K(Q, P) = \int \log\left(\frac{P}{Q}\right)(P - Q) dx$
- **Non** soddisfa la diseguaglianza triangolare

Distanza di Kullback-Leibler

Definizione

Date due densità di probabilità P e Q si definisce:

$$K(P, Q) := \mathbb{E} \left[\ln \left(\frac{P}{Q} \right) \right] = \int \ln \left(\frac{P(x)}{Q(x)} \right) dP(x)$$

la distanza di Kullback-Leibler.

Nata originariamente per definire l'**informazione mutuale** tra Variabili Aleatorie X e Y come

$$\begin{aligned} I(X, Y) &:= K(\mathbb{P}(X, Y), \mathbb{P}(X)\mathbb{P}(Y)) \\ &= \int \ln \left(\frac{\mathbb{P}(X, Y)}{\mathbb{P}(X)\mathbb{P}(Y)} \right) d\mathbb{P}(X, Y) \end{aligned}$$

Distanza di Kullback-Leibler

Consideriamo due variabili aleatorie gaussiane di media nulla e varianza Σ :

$$X_1 \sim N(0, \Sigma_1) \quad X_2 \sim N(0, \Sigma_2)$$

e le relative densità di probabilità associate

$$P_i(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma_i)}} e^{-\frac{1}{2}x^T \Sigma_i^{-1} x}$$

Distanza di Kullback-Leibler

Consideriamo due variabili aleatorie gaussiane di media nulla e varianza Σ :

$$X_1 \sim N(0, \Sigma_1) \quad X_2 \sim N(0, \Sigma_2)$$

e le relative densità di probabilità associate

$$P_i(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma_i)}} e^{-\frac{1}{2}x^T \Sigma_i^{-1} x}$$

Possiamo quindi definire una distanza tra matrici

Definizione

Date due matrici simmetriche definite positive Σ_1 e Σ_2 si definisce:

$$K(\Sigma_1, \Sigma_2) := K(P_1, P_2)$$

la distanza di Kullback-Leibler tra matrici.

Distanza di Kullback-Leibler

$$K(\Sigma_1, \Sigma_2) = \int \ln \left(\frac{P_1(x)}{P_2(x)} \right) dP_1(x)$$

Distanza di Kullback-Leibler

$$K(\Sigma_1, \Sigma_2) = \int \ln \left(\frac{P_1(x)}{P_2(x)} \right) dP_1(x)$$

$$= \frac{1}{2} \left(\ln \left(\frac{\det(\Sigma_2)}{\det(\Sigma_1)} \right) + \text{TR}(\Sigma_2^{-1} \Sigma_1) - n \right)$$

Distanza di Kullback-Leibler

$$\begin{aligned} K(\Sigma_1, \Sigma_2) &= \int \ln \left(\frac{P_1(x)}{P_2(x)} \right) dP_1(x) \\ &= \int \ln \left(\sqrt{\frac{\det(\Sigma_2)}{\det(\Sigma_1)}} e^{-\frac{1}{2}(x^T \Sigma_1^{-1} x - x^T \Sigma_2^{-1} x)} \right) dP_1(x) \\ &= \int \ln \left(\sqrt{\frac{\det(\Sigma_2)}{\det(\Sigma_1)}} \right) dP_1(x) - \frac{1}{2} \int (x^T \Sigma_1^{-1} x - x^T \Sigma_2^{-1} x) dP_1(x) \\ &= \frac{1}{2} \left(\ln \left(\frac{\det(\Sigma_2)}{\det(\Sigma_1)} \right) + \int x^T \Sigma_2^{-1} x dP_1(x) - \int x^T \Sigma_1^{-1} x dP_1(x) \right) \\ &= \frac{1}{2} \left(\ln \left(\frac{\det(\Sigma_2)}{\det(\Sigma_1)} \right) + \text{TR}(\Sigma_2^{-1} \Sigma_1) - n \right) \end{aligned}$$

calcolabile esplicitamente in funzione delle sole matrici Σ_1 e Σ_2

Distanza di Kullback-Leibler

Delle C così costruite seguono la distribuzione di Wishart

$$C \sim W_n(\Sigma, t)$$

Distanza di Kullback-Leibler

Delle C così costruite seguono la distribuzione di Wishart

$$C \sim W_n(\Sigma, t)$$

con densità

$$f(C) = \frac{1}{\Gamma_n\left(\frac{t}{2}\right) \left(\sqrt{2^n \det(\Sigma)}\right)^t} (\det(C))^{\frac{t-n-1}{2}} e^{-\frac{1}{2} \text{TR}(\Sigma^{-1} C)}$$

dove $\Gamma_n\left(\frac{t}{2}\right) = \pi^{\frac{n(n-1)}{4}} \prod_{j=0}^{n-1} \Gamma\left(\frac{t-j}{2}\right)$ è la Multivariate Gamma Function.

Distanza di Kullback-Leibler

Grazie a questa possiamo calcolare

$$\mathbb{E}_W[K(\Sigma, C)]$$

$$\mathbb{E}_W[K(C, \Sigma)]$$

$$\mathbb{E}_{W \otimes W}[K(C, C)]$$

Distanza di Kullback-Leibler

Grazie a questa possiamo calcolare

$$\mathbb{E}_W[K(\Sigma, C)] = \frac{1}{2} \left(n \ln \left(\frac{2}{t} \right) + \sum_{p=t-n+1}^t \left(\frac{\Gamma'(p/2)}{\Gamma(p/2)} \right) + \frac{n(n+1)}{t-n-1} \right)$$

$$\mathbb{E}_W[K(C, \Sigma)] = \frac{1}{2} \left(n \ln \left(\frac{t}{2} \right) - \sum_{p=t-n+1}^t \left(\frac{\Gamma'(p/2)}{\Gamma(p/2)} \right) \right)$$

$$\mathbb{E}_{W \otimes W}[K(C, C)] = \frac{1}{2} \frac{n(n-1)}{t-n-1}$$

Distanza di Kullback-Leibler

Grazie a questa possiamo calcolare

$$\mathbb{E}_W[K(\Sigma, C)] = \frac{1}{2} \left(n \ln \left(\frac{2}{t} \right) + \sum_{p=t-n+1}^t \left(\frac{\Gamma'(p/2)}{\Gamma(p/2)} \right) + \frac{n(n+1)}{t-n-1} \right)$$

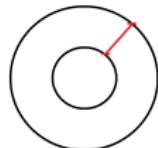
$$\mathbb{E}_W[K(C, \Sigma)] = \frac{1}{2} \left(n \ln \left(\frac{t}{2} \right) - \sum_{p=t-n+1}^t \left(\frac{\Gamma'(p/2)}{\Gamma(p/2)} \right) \right)$$

$$\mathbb{E}_{W \otimes W}[K(C, C)] = \frac{1}{2} \frac{n(n-1)}{t-n-1}$$

che **NON** dipendono da Σ

Applicazione

- $\langle d(C, C^{(filt)}) \rangle$

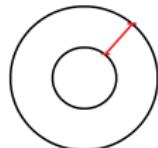


Quanta informazione di C viene persa/trattenuta in $C^{(filt)}$?

$$0 = \underbrace{\langle K(C, C) \rangle}_{\text{filtraggio nullo}} \leq \langle K(C, C^{(filt)}) \rangle \leq \underbrace{\langle K(C, \Sigma) \rangle \approx \mathbb{E}_W[K(C, \Sigma)]}_{\text{filtraggio perfetto}}$$

Applicazione

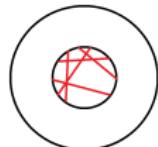
- $\langle d(C, C^{(filt)}) \rangle$



Quanta informazione di C viene persa/trattenuta in $C^{(filt)}$?

$$0 = \underbrace{\langle K(C, C) \rangle}_{\text{filtraggio nullo}} \leq \langle K(C, C^{(filt)}) \rangle \leq \underbrace{\langle K(C, \Sigma) \rangle}_{\text{filtraggio perfetto}} \approx \mathbb{E}_W[K(C, \Sigma)]$$

- $\langle d(C^{(filt)}, C^{(filt)}) \rangle$



Quanto è stabile il filtraggio?

$$0 = \underbrace{\langle K(\Sigma, \Sigma) \rangle}_{\text{filtraggio perfetto}} \leq \langle K(C_i^{(filt)}, C_j^{(filt)}) \rangle \leq \underbrace{\langle K(C_i, C_j) \rangle}_{\text{filtraggio nullo}} \approx \mathbb{E}_{W \otimes W}[K(C, C)]$$

Applicazione - Test Teorici

- Fissiamo n e t

Applicazione - Test Teorici

- Fissiamo n e t
- Generiamo una Σ casuale $\in \mathfrak{M}(n, n, \mathbb{R})$

Applicazione - Test Teorici

- Fissiamo n e t
- Generiamo una Σ casuale $\in \mathfrak{M}(n, n, \mathbb{R})$
- Costruiamo **tante** sue realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$

Applicazione - Test Teorici

- Fissiamo n e t
- Generiamo una Σ casuale $\in \mathfrak{M}(n, n, \mathbb{R})$
- Costruiamo **tante** sue realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i

Applicazione - Test Teorici

- Fissiamo n e t
- Generiamo una Σ casuale $\in \mathfrak{M}(n, n, \mathbb{R})$
- Costruiamo **tante** sue realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i
- Applichiamo i filtraggi

$$C_i \rightsquigarrow [C_i^{(filt_{Max})}, C_i^{(filt_{Med})}, C_i^{(filt_{Potter})}, C_i^{(filt_{Rosenow})}]$$

Applicazione - Test Teorici

- Fissiamo n e t
- Generiamo una Σ casuale $\in \mathfrak{M}(n, n, \mathbb{R})$
- Costruiamo **tante** sue realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i
- Applichiamo i filtri

$$C_i \rightsquigarrow [C_i^{(filt_{Max})}, C_i^{(filt_{Med})}, C_i^{(filt_{Potter})}, C_i^{(filt_{Rosenow})}]$$

- Calcoliamo le medie delle distanze e plottiamo i 4 punti

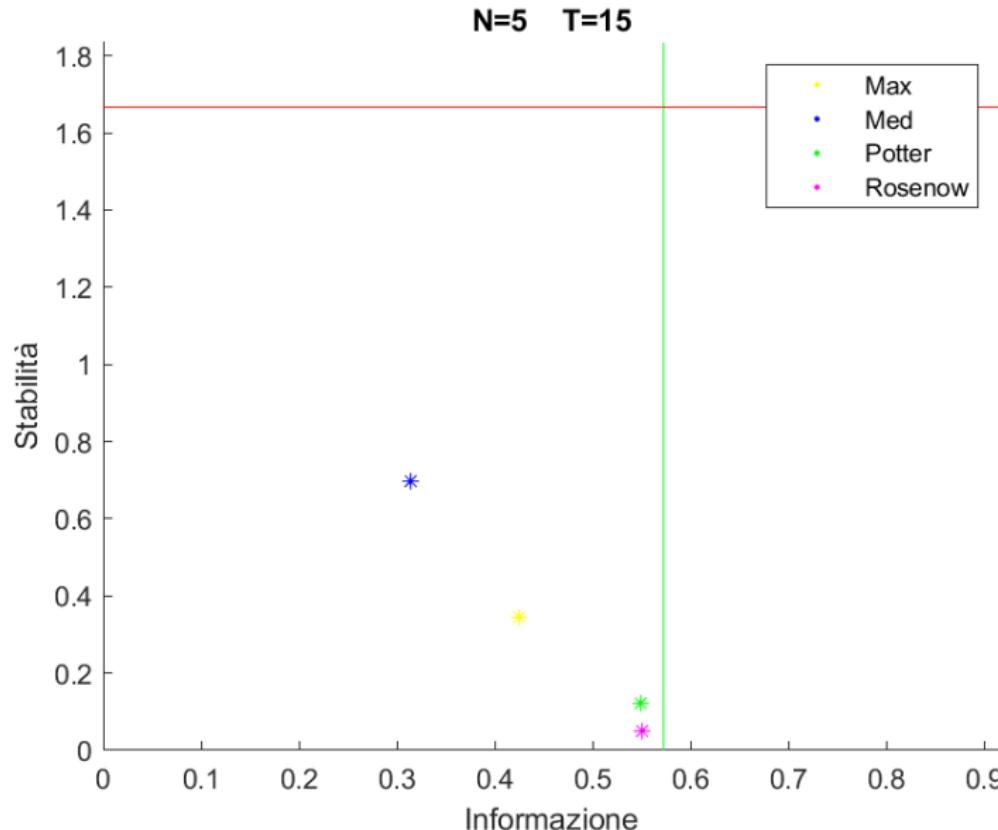
$$\left(\langle K(C, C^{(filt_{Max})}) \rangle, \langle K(C^{(filt_{Max})}, C^{(filt_{Max})}) \rangle \right) \in \mathbb{R}^2$$

$$\left(\langle K(C, C^{(filt_{Med})}) \rangle, \langle K(C^{(filt_{Med})}, C^{(filt_{Med})}) \rangle \right) \in \mathbb{R}^2$$

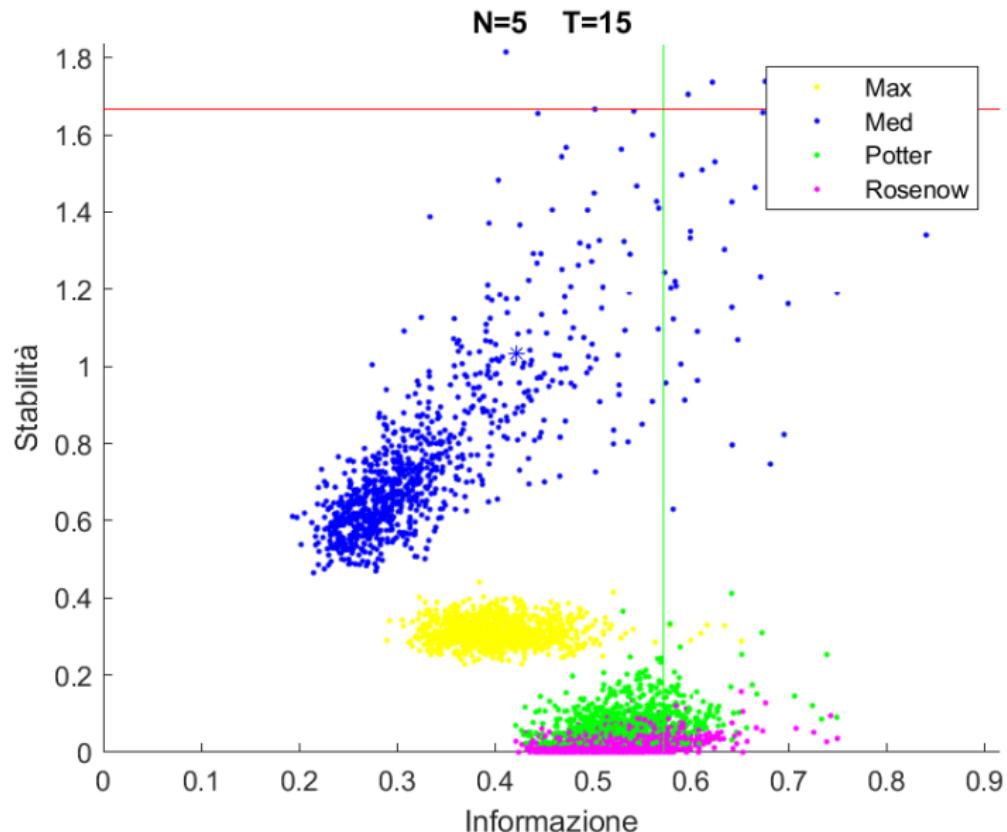
$$\left(\langle K(C, C^{(filt_{Potter})}) \rangle, \langle K(C^{(filt_{Potter})}, C^{(filt_{Potter})}) \rangle \right) \in \mathbb{R}^2$$

$$\left(\langle K(C, C^{(filt_{Rosenow})}) \rangle, \langle K(C^{(filt_{Rosenow})}, C^{(filt_{Rosenow})}) \rangle \right) \in \mathbb{R}^2$$

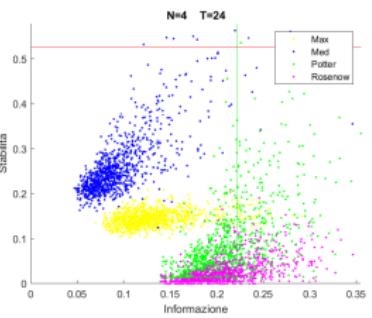
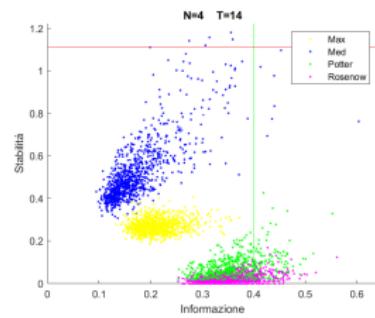
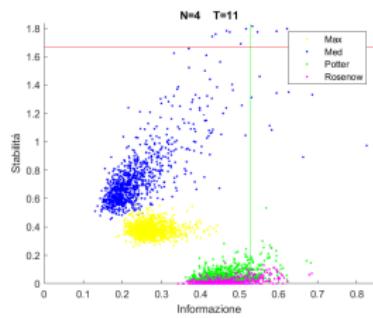
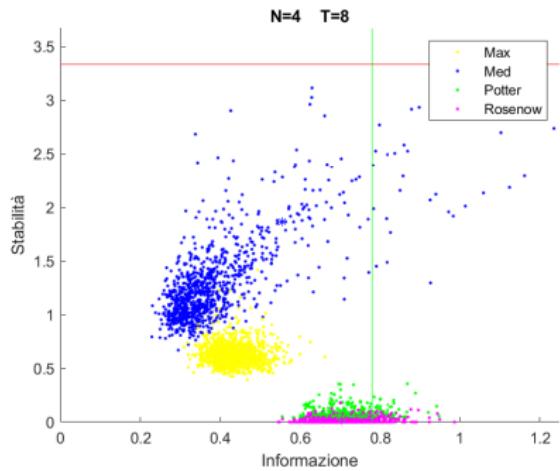
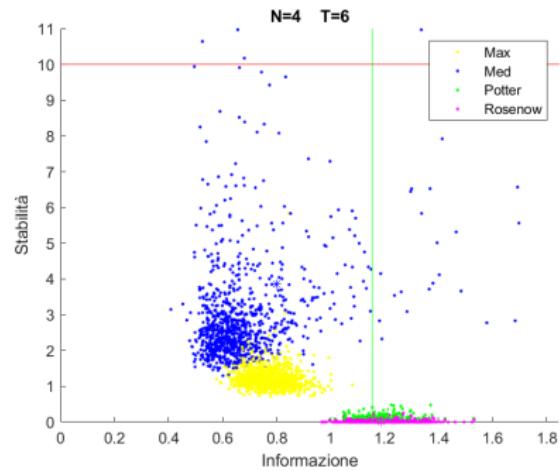
Applicazione - Test Teorici



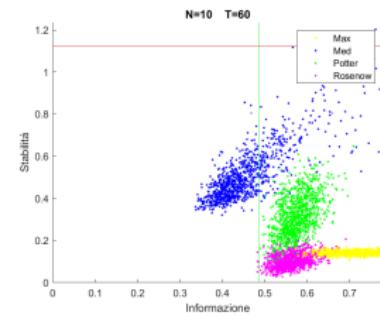
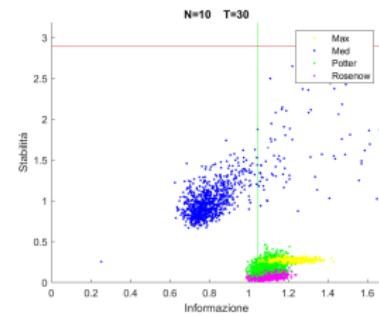
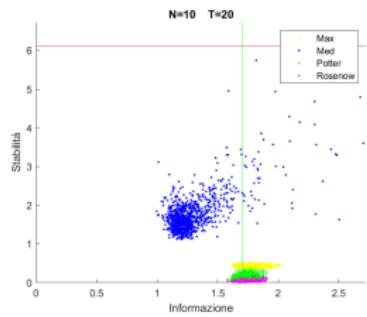
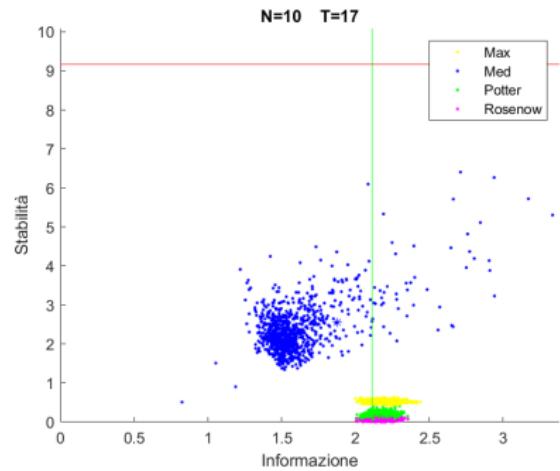
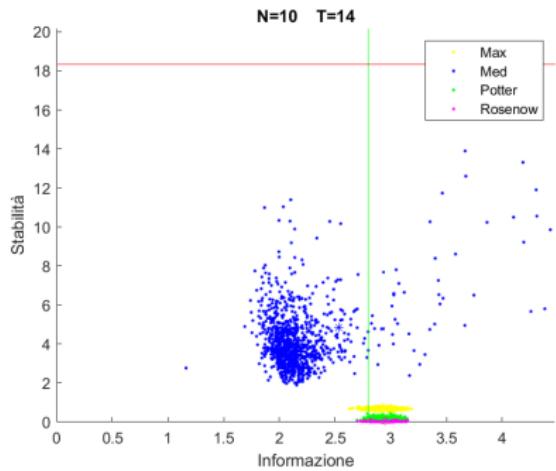
Applicazione - Test Teorici



Applicazione - Test Teorici ($n=4$)



Applicazione - Test Teorici (n=10)

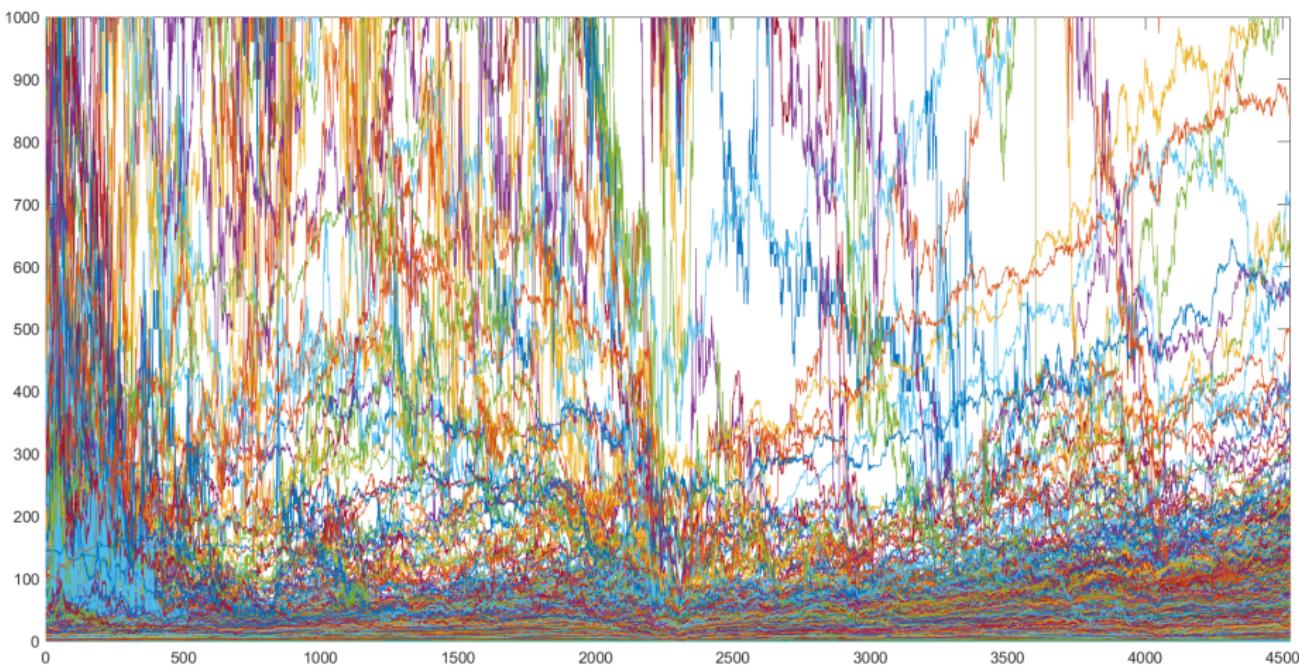


Applicazione - Test Pratici

I dati da cui sono partito sono i prezzi in chiusura giornalieri di 7227 stock.
Di questi ho selezionati i 2122 che avevano dati per almeno 20 anni

Applicazione - Test Pratici

I dati da cui sono partito sono i prezzi in chiusura giornalieri di 7227 stock.
Di questi ho selezionati i 2122 che avevano dati per almeno 20 anni



Applicazione - Test Pratici

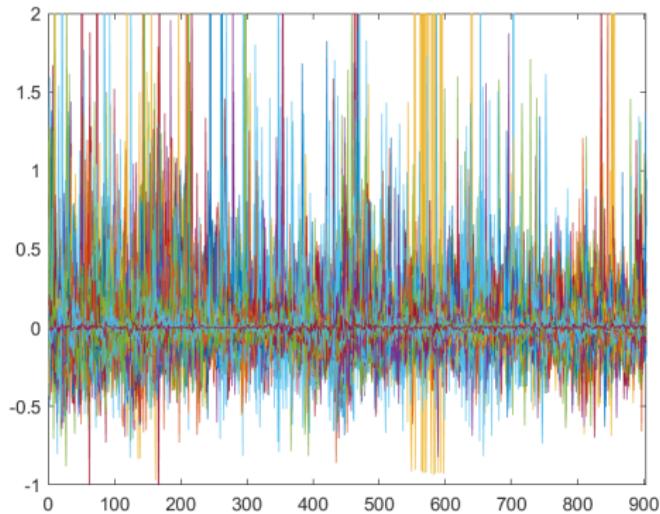
Dai prezzi giornalieri sono passato alle medie settimanali e da queste sono passato agli incrementi percentuali tra una settimana e la successiva.

$$incr(i) = \frac{prezzo(i) - prezzo(i - 1)}{|prezzo(i)|}$$

Applicazione - Test Pratici

Dai prezzi giornalieri sono passato alle medie settimanali e da queste sono passato agli incrementi percentuali tra una settimana e la successiva.

$$incr(i) = \frac{prezzo(i) - prezzo(i - 1)}{|prezzo(i)|}$$



Applicazione - Test Pratici

- Fissiamo n e t

Applicazione - Test Pratici

- Fissiamo n e t
- Scegliamo n stock quotati nella borsa di New York

Applicazione - Test Pratici

- Fissiamo n e t
- Scegliamo n stock quotati nella borsa di New York
- Prendiamo i dati relativi relativi a **tanti** intervalli di tempo lunghi t disgiunti (=realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$)

Applicazione - Test Pratici

- Fissiamo n e t
- Scegliamo n stock quotati nella borsa di New York
- Prendiamo i dati relativi relativi a **tanti** intervalli di tempo lunghi t disgiunti (=realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$)
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i

Applicazione - Test Pratici

- Fissiamo n e t
- Scegliamo n stock quotati nella borsa di New York
- Prendiamo i dati relativi relativi a **tanti** intervalli di tempo lunghi t disgiunti (=realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$)
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i
- Applichiamo i filtri

$$C_i \rightsquigarrow [C_i^{(filt_{Max})}, C_i^{(filt_{Med})}, C_i^{(filt_{Potter})}, C_i^{(filt_{Rosenow})}]$$

Applicazione - Test Pratici

- Fissiamo n e t
- Scegliamo n stock quotati nella borsa di New York
- Prendiamo i dati relativi relativi a **tanti** intervalli di tempo lunghi t disgiunti (=realizzazioni indipendenti $M_i \in \mathfrak{M}(n, t, \mathbb{R})$)
- Per ciascuna realizzazione M_i calcoliamo la matrice di covarianza C_i
- Applichiamo i filtri

$$C_i \rightsquigarrow [C_i^{(filt_{Max})}, C_i^{(filt_{Med})}, C_i^{(filt_{Potter})}, C_i^{(filt_{Rosenow})}]$$

- Calcoliamo le medie delle distanze e plottiamo i 4 punti

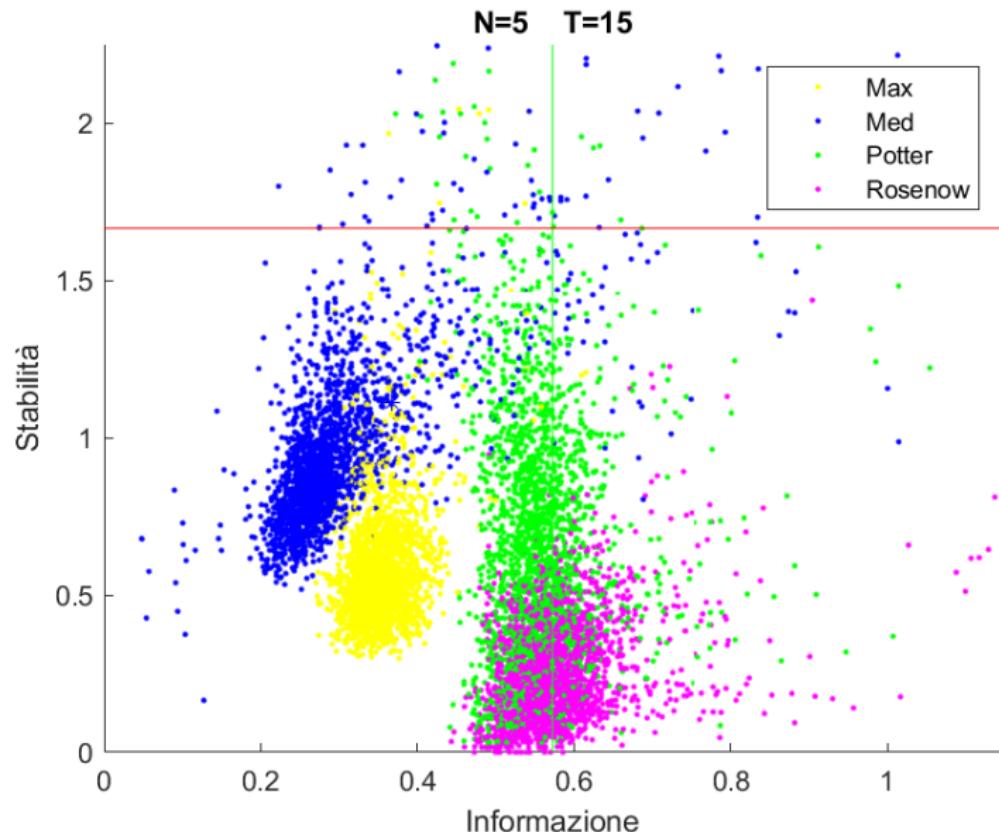
$$\left(\langle K(C, C^{(filt_{Max})}) \rangle, \langle K(C^{(filt_{Max})}, C^{(filt_{Max})}) \rangle \right) \in \mathbb{R}^2$$

$$\left(\langle K(C, C^{(filt_{Med})}) \rangle, \langle K(C^{(filt_{Med})}, C^{(filt_{Med})}) \rangle \right) \in \mathbb{R}^2$$

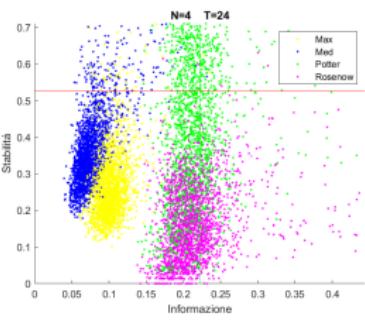
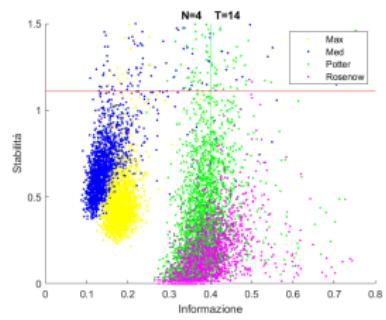
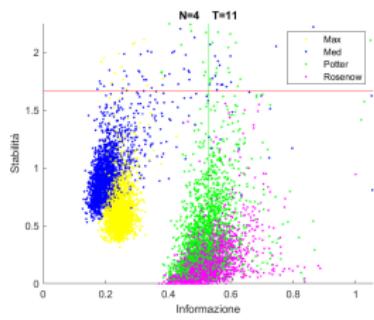
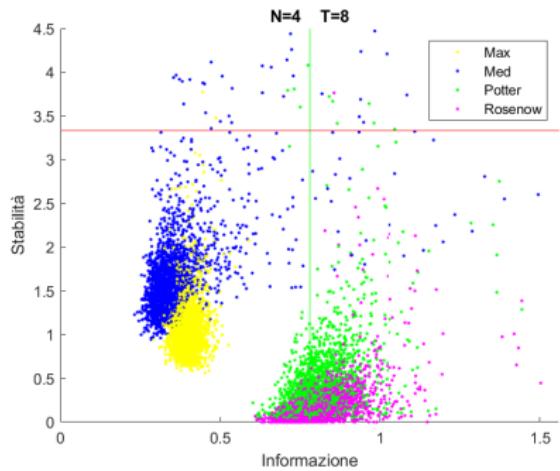
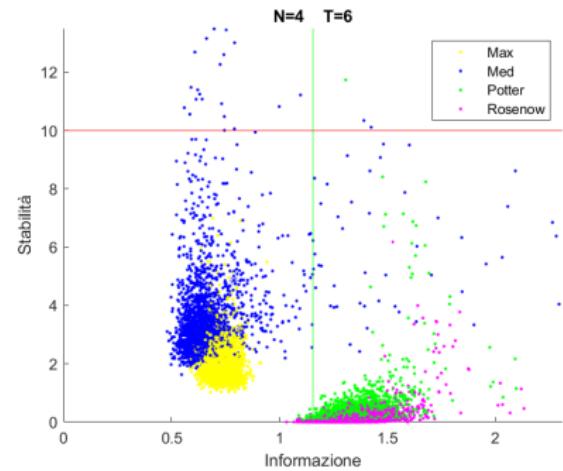
$$\left(\langle K(C, C^{(filt_{Potter})}) \rangle, \langle K(C^{(filt_{Potter})}, C^{(filt_{Potter})}) \rangle \right) \in \mathbb{R}^2$$

$$\left(\langle K(C, C^{(filt_{Rosenow})}) \rangle, \langle K(C^{(filt_{Rosenow})}, C^{(filt_{Rosenow})}) \rangle \right) \in \mathbb{R}^2$$

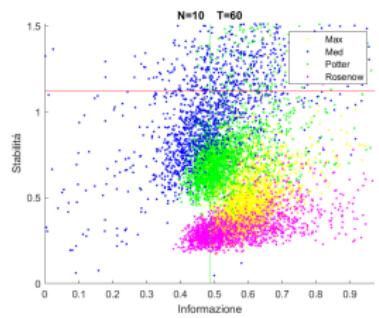
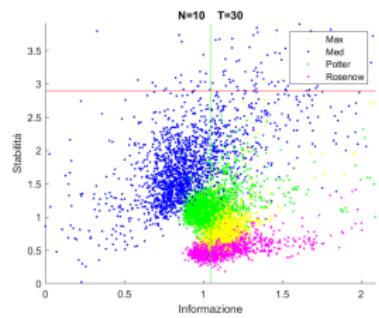
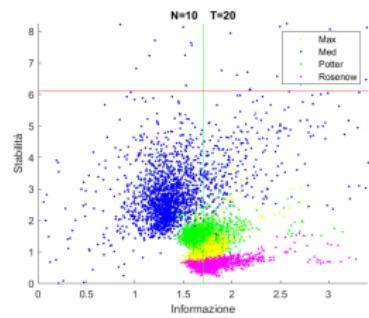
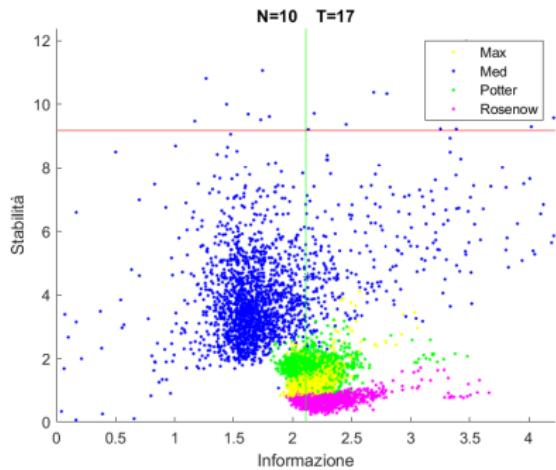
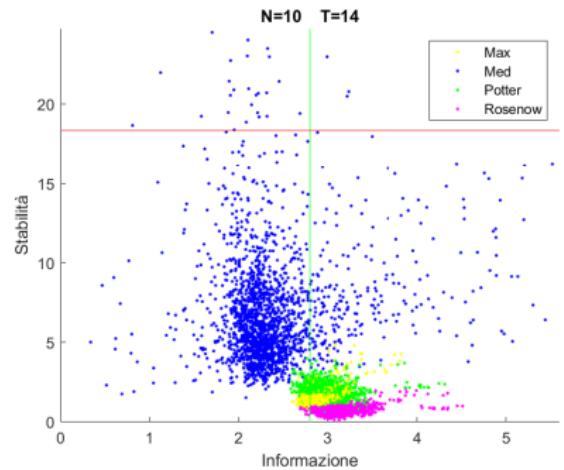
Applicazione - Test Pratici



Applicazione - Test Pratici ($n=4$)



Applicazione - Test Pratici ($n=10$)



Proiettile di Markowitz

Sia dato un mercato con vettore dei ritorni attesi M e covarianza Σ .

Definizione

Un **portfolio** è un vettore $W \in \{[0, 1]\}^n$ di pesi.

W_i è la percentuale del capitale che investiamo nello stock i . ($\sum W_i = 1$)

Il ritorno atteso è quindi $\mu = MW^T$ e la varianza è $\sigma^2 = W\Sigma W^T$.

C'è un ordinamento parziale nell'insieme \mathcal{W} dei portafogli.

$$W_1 \preceq W_2 \Leftrightarrow \mathbb{E}[W_1] \leq \mathbb{E}[W_2] \wedge \text{Var}(W_1) \geq \text{Var}(W_2)$$

Definizione

Un portfolio si dice **efficiente** se è un elemento massimale rispetto a \preceq

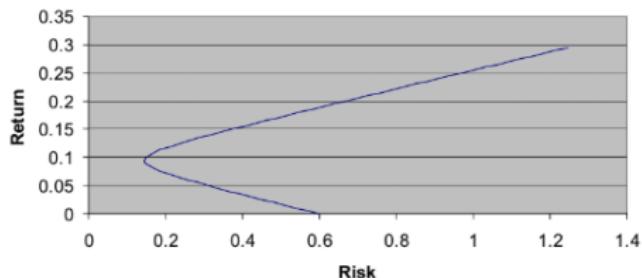
Proiettile di Markowitz

Markowitz

Dato $\mu \in \mathbb{R}^+$, il portfolio con ritorno atteso μ e varianza minima ha pesi:

$$W = \frac{\det \begin{pmatrix} \mu & M\Sigma^{-1}\mathbb{1}^T \\ 1 & \mathbb{1}\Sigma^{-1}\mathbb{1}^T \end{pmatrix} M\Sigma^{-1} + \det \begin{pmatrix} M\Sigma^{-1}M^T & \mu \\ \mathbb{1}\Sigma^{-1}M^T & 1 \end{pmatrix} \mathbb{1}\Sigma^{-1}}{\det \begin{pmatrix} M\Sigma^{-1}M^T & M\Sigma^{-1}\mathbb{1}^T \\ \mathbb{1}\Sigma^{-1}M^T & \mathbb{1}\Sigma^{-1}\mathbb{1}^T \end{pmatrix}}$$

Possiamo quindi tracciare la curva dei portafogli efficienti, il cosiddetto proiettile di Markowitz



Criterio di Kelly

Supponiamo che a un giocatore venga data la possibilità di scommettere una fissata frazione f del suo capitale (arbitrariamente divisibile) su lanci successivi di monete.

Ad ogni lancio con probabilità $p < \frac{1}{2}$ vince quanto ha scommesso, con probabilità $1 - p$ perde.

Quale f scegliere?

- $f = 1$ massimizza il valore atteso
- $f = 0$ minimizza la varianza

Criterio di Kelly

Se F_n è il capitale al passo n , definiamo il grow rate esponenziale come

$$G := \lim_{n \rightarrow \infty} \log \left(\left(\frac{F_n}{F_0} \right)^{\frac{1}{n}} \right)$$

Se il giocatore dopo n passi ha vinto W volte e perso L volte,
 $F_n = (1 + f)^W (1 - f)^L F_0$, e quindi

$$G = \lim_{n \rightarrow \infty} \left(\frac{W}{n} \log(1 + f) + \frac{L}{n} \log(1 - f) \right) = p \log(1 + f) + (1 - p) \log(1 - f)$$

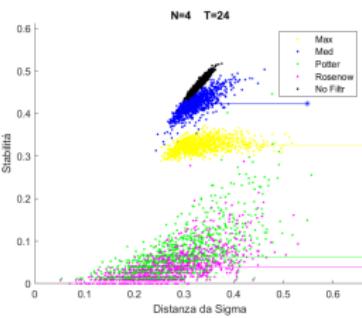
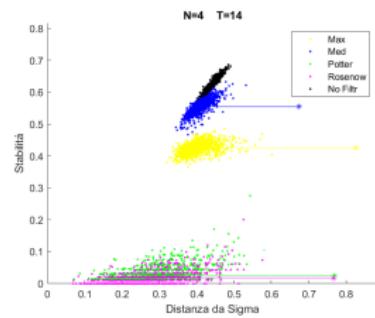
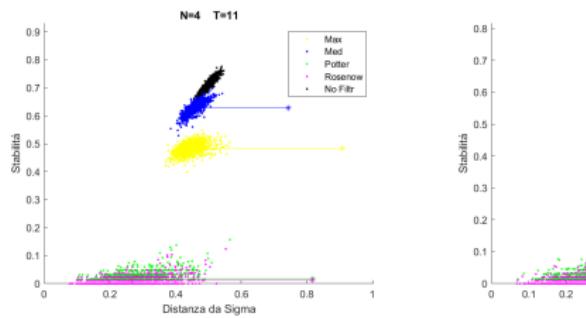
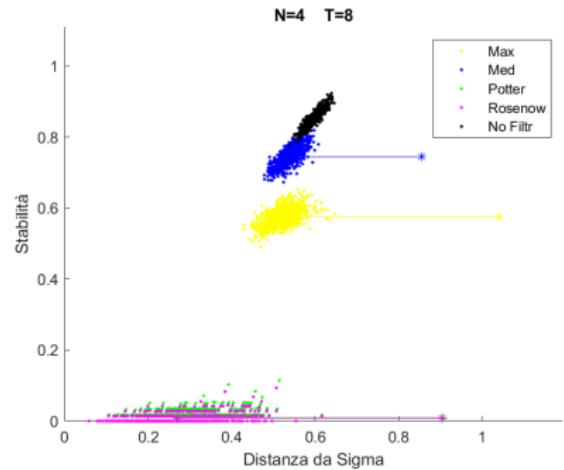
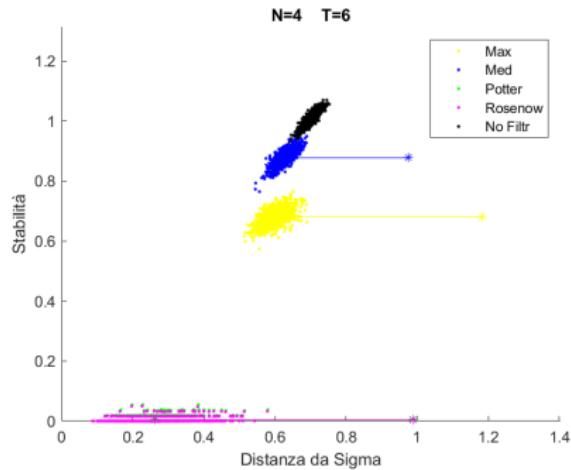
Il criterio di Kelly

Se definiamo $f_{ottimale}^* := \arg \min \{ \mathbb{E}[\log(return(f^*))] : 0 \leq f^* \leq 1 \}$

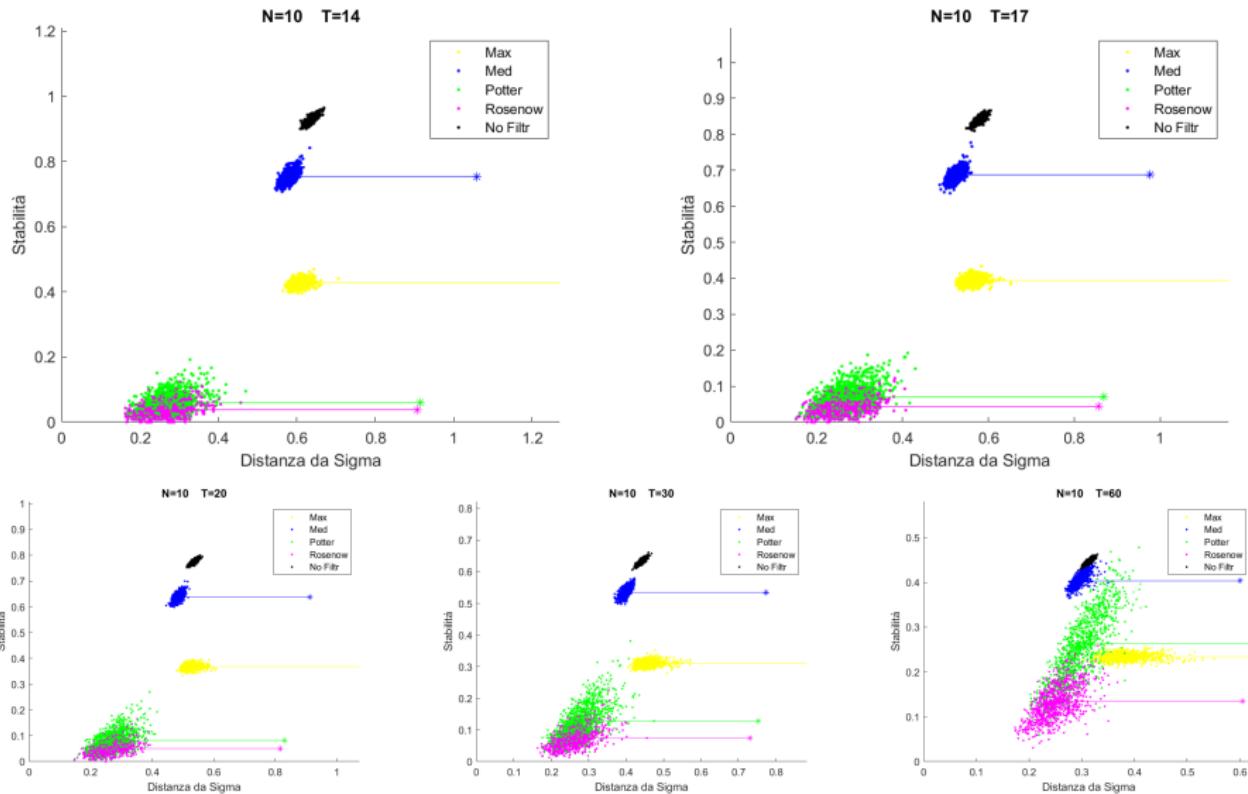
Per scommesse semplici con due possibili risultati vale:

$$f^* = p - (1 - p)$$

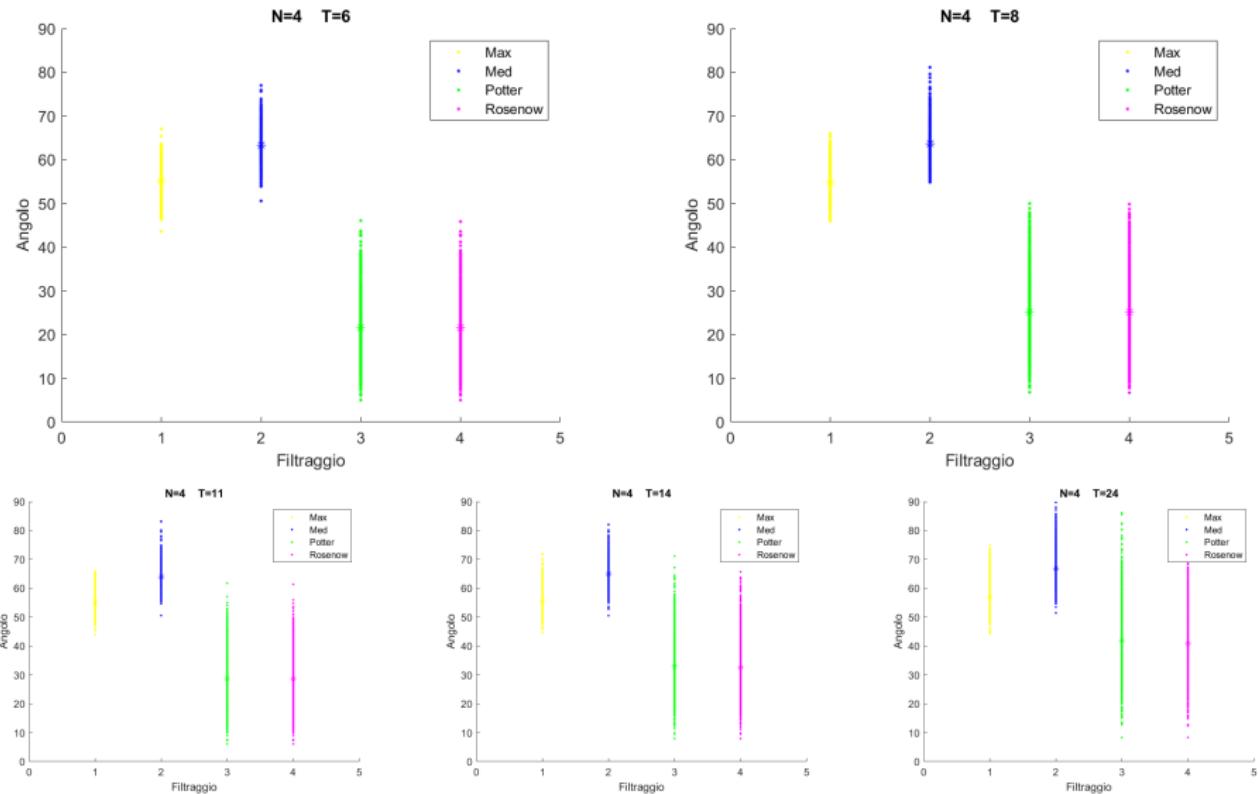
Check Risultati



Check Risultati



Check Risultati



Check Risultati

